Population Health: Behavioral and Social Science Insights

Edited by Robert M. Kaplan, Michael L. Spittel, Daryn H. David

AHRQ Agency for Healthcare Research and Quality

NIH National Institutes of Health
Office of Behavioral and Social Sciences Research

This report has been jointly produced by the Agency for Healthcare Research and Quality (AHRQ) and the Office of Behavioral and Social Sciences Research, National Institutes of Health (NIH). The findings and conclusions presented herein are those of the authors, who are responsible for the content; the findings and conclusions do not necessarily represent the views of AHRQ, NIH, or the U.S. Department of Health and Human Services.

Suggested citation:

Kaplan R, Spittel M, David D (Eds). Population Health: Behavioral and Social Science Insights. AHRQ Publication No. 15-0002. Rockville, MD: Agency for Healthcare Research and Quality and Office of Behavioral and Social Sciences Research, National Institutes of Health; July 2015.

Population Health: Behavioral and Social Science Insights

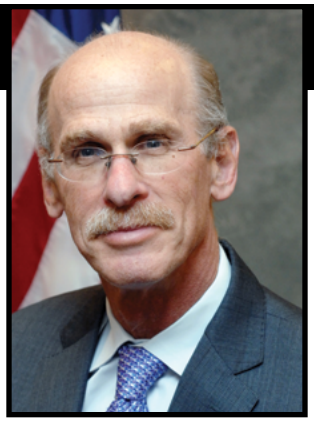
Co-edited by:

Robert M. Kaplan
Agency for Healthcare Research and Quality

Michael L. Spittel
Daryn H. David
Office of Behavioral and Social Sciences Research
National Institutes of Health

AHRQ Publication No. 15-0002
July 2015
ISBN 978-1-58763-444-4

Foreword



Richard Kronick

Health care is an important factor in improving health, but health outcomes are also strongly influenced by factors outside of the health care system. This volume brings together contributions from leading behavioral, social, and medical scientists to summarize what is known about the determinants of health outcomes and to outline directions for future research. I am gratified that the Agency for Healthcare Research and Quality (AHRQ) has been able to partner with the National Institutes of Health (NIH) in this effort.

AHRQ's mission is to produce evidence to make health care safer, improve health care quality, and make it more accessible, equitable, and affordable; a parallel goal is to work within the U.S. Department of Health and Human Services and with other partners to make sure that the evidence is understood and used. As an agency we have a strong commitment to disseminate the best available evidence to improve health outcomes from the patient perspective. AHRQ publishes reports that are relevant to improving population health, including the National Healthcare Quality and Disparities Report. This report focuses on factors within the health care delivery system, including where health disparities are improving and where they have stayed the same or gotten worse.

There are at least three ways to improve population health outcomes: first, by developing new treatments and therapies that will improve health for people suffering from illness (e.g., a cure for cancer; a treatment to slow or stop disease progression in people with multiple sclerosis [MS], a therapy to prevent the onset of Alzheimer's disease, and so on). Second, by assuring the technologies and therapies that are known to produce benefit are more uniformly delivered to all Americans. And third, we need to tackle the systemic problems that continue to negatively affect health, such as poverty and low educational attainment. AHRQ concentrates on the second of the three mechanisms for improving population health. NIH and the industry tend to focus on the first, and the chapters in this volume remind us of the importance of paying attention to the third.

Health care is limited in its capacity to remedy the national challenge of reducing racial and ethnic disparities in health outcomes. Population Health: Behavioral and Social Science Insights broadens our perspective by helping us understand the context in which health problems develop and the environments in which health care is delivered. As more information becomes available to clinicians, patients, and communities alike, AHRQ's goal is to help improve health and well-being and contribute to better health outcomes for the Nation overall.

Richard Kronick, PhD
Director
Agency for Healthcare Research and Quality

William Riley

My good colleague, Russ Glasgow, taught us that Reach X Efficacy = Impact. The vast majority of behavioral and social science research efforts to improve health have targeted the individual. These efforts have resulted in numerous efficacious interventions to change behaviors and improve health, but the field has struggled to implement these interventions with sufficient scalability or reach to have a large population impact. In contrast, this book focuses on population health and the targeting of large population groups with common systemic risks to their health. Although we may have less control of the levers of change in population health than in individually-based interventions, even small systemic shifts or "nudges" have considerable reach that produce large population impacts.

Over the last generation, we have witnessed the impact of increased tobacco taxes and indoor smoking bans on the prevalence of smoking. Of course, these population-based interventions do not occur in a vacuum, and the availability of viable individual cessation interventions, as well as changing attitudes regarding smoking, provided the basis for acceptance of these population-based interventions. As evident from the chapters in this book, tobacco use is but one of many population-based targets for improving health, and population health extends beyond policies that influence behavioral risk factors to include demographic and social influences (e.g., education, income), health care access and quality, and the interaction of these population influences with genetics and neurobiology.

Population health is a growing and diverse research area. Much of the early work in this area was primarily descriptive, focusing on the distributions of various health indexes within certain populations. Much of the recent research has focused on the mechanisms that produce these distributions of population health. This research is complex and is often limited by the mechanisms that can be gleaned from epidemiologic studies. The final section of this book on emerging tools for studying population health is an important contribution for improving the methodology to advance this science. Advancing population health methods is critical if we are to understand and target these mechanisms to develop population-based interventions that are empirically-based and have the potential to impact population health.

The National Institutes of Health's Office of Behavioral and Social Sciences Research (OBSSR) has an extensive history of advancing research on population health. This book is an outgrowth of work begun at OBSSR when Robert Kaplan, the co-editor of this book, was Director of OBSSR. Along with his co-editors Michael Spittel and Daryn David, Dr. Kaplan has drawn together in one book the work of many of the leaders in population health. It is the hope of OBSSR that this book stimulates rigorous and relevant research in population health that will improve the health of the Nation.

William T. Riley, PhD
Director, Office of Behavioral and Social Sciences Research

Preface

Christine A. Bachrach

The United States spends more on health care than any other advanced economy but still lags behind other countries on most measures of population health. This is a familiar refrain to devotees of health statistics but a surprising and troubling revelation to many who take pride in the outstanding health care U.S. citizens enjoy. The truth is, as reports from the Institute of Medicine and the National Research Council document, current U.S. approaches to safeguarding the health of the American people are not working as they should.

It's not that policymakers have neglected the problem. The Patient Protection and Affordable Care Act (ACA) has been a promising step towards overhauling the country's health care system. If it remains in place, it should help to bring down health care costs and extend coverage to many vulnerable groups who previously could not afford needed care. However, the primary focus of the ACA on reforming the health care system points to a fundamental problem in our societal approach to health: we tend to equate health with health care. We pour millions of dollars into finding cures for disease, both through taxpayer investments in the National Institutes of Health and charitable contributions to the many nonprofit organizations that also fund the search for biomedical advances. Faced with problems such as smoking and obesity, we turn to physicians to advise patients about behavior change and think that will solve the problem, despite the limited time available for clinical visits and the proven inefficacy of information alone in producing lasting behavior change. Only a small proportion of our national health expenditures goes to prevention.

Beyond medical care, we link health to personal behavior. In the images of pharmaceutical advertising, healthy people are exercising, getting sleep, and planting gardens. Media and public health messages extoll healthy behaviors, urging people to gyms, farmers markets, and organic food stores. We lionize the superfit in our admiration of extreme sports, marathoners, and football players, even if such pursuits take an often negative toll on health. Overwhelmingly, our investments in encouraging healthy behavior are targeted toward the individual. Efforts to facilitate healthy behaviors at a structural level – for example, by limiting portion sizes for sugary soft drinks – are a hard sell in the context of cultural values that put personal responsibility and personal freedoms first.

What's missing from these medicalized and individual-focused approaches to health is the recognition that health is as much the product of the social and physical environments people occupy as it is of their biology and behavior. Although recognized centuries ago, this fact is now supported by a large body of scientific evidence that shows not only the alignment of health with supportive environments, but also causal mechanisms that translate poor environmental conditions into poor health outcomes. The successful reduction of tobacco use in the last decades of the 20th century provides a compelling example of how structural approaches, such as taxation and smoke-free public spaces, can combine with pharmaceutical and behavior-change interventions to produce lasting health improvements.

This approach is now gaining widespread support. Reports from the World Health Organization and the Institute of Medicine have urged action targeting the "upstream" determinants of health. The Robert Wood Johnson Foundation's Commission to Build a Healthier America pointed to specific steps that can be taken by local, State, and Federal governments, as well as by businesses, schools, health care providers, and local community groups. Institutions as diverse as the Federal Reserve Bank, the National Institutes of Health, and Best Buy have implemented initiatives and policies targeted at the community and workplace levels, while the movement for Health in All Policies by the National Association of County & City Health Officials draws attention to the potential consequences of all policies, not just health care system policies, for improving or diminishing health.

The promise of developing multi-level solutions to population health problems depends on continuing to build the scientific evidence that informs these efforts. Understanding and improving population health requires that scientists from widely differing disciplines combine their knowledge about the societal, behavioral, and biological causes of health and work towards an integrated science that can identify the many and complex mechanisms through which health and health disparities are produced. It requires an understanding of how individual-level processes translate to population-level processes, and vice versa. It also requires scientists to examine common health determinants across different diseases and conditions. In short, it requires the new integrative approach to the science of health that is coming to be known as population health science.

Population health science is advancing rapidly, thanks to investments by the U.S. Department of Health and Human Services through agencies such as the Agency for Healthcare Research and Quality (AHRQ) and the National Institutes of Health (NIH). NIH's Office of Behavioral and Social Sciences Research has devoted considerable effort to coordinating these activities. Advances in population health science are the product of leadership, creativity, and collaboration among scientists from a wide range of disciplines and professions. Because of their ability to address human behavior and social systems, the behavioral and social sciences have played a prominent role in advancing this field, but the most exciting advances are emerging out of science that transcends the biological, behavioral, and social sciences. The contents of this volume testify both to how far we have come in understanding the production of health from a multi-level perspective to the many opportunities that are available for increasing the rigor, reach, and impact of population health science.

Christine A. Bachrach, PhD
Research Professor, Department of Sociology and
Maryland Population Research Center
University of Maryland, College Park
Co-Director, Health & Society Scholars Program
Robert Wood Johnson Foundation

Innovations in Population Health Research: The Challenge

Robert M. Kaplan, Daryn H. David, and Michael L. Spittel

The National Institutes of Health (NIH) is the largest supporter of biomedical research in the world. The NIH provides stewardship for medical and behavioral research for the United States and for populations throughout the world. The well-known basic science mission of the NIH is to pursue fundamental knowledge about the nature and behavior of living systems. Perhaps less appreciated is the second clause of the mission statement which emphasizes the application of knowledge to enhance health, lengthen life, and reduce illness and disability.

The goal of enhancing health, lengthening life, and reducing illness and disability is consistent with the health objectives for the United States and for the overall mission of the U.S. Department of Health and Human Services (HHS). Specifically, since the introduction of the Healthy People reports in 1990,[1] the overarching national goal has been to extend life expectancy and improve the quality of life for people living in America. The HHS mission is to enhance the health and well-being of Americans by providing effective health and social services and by fostering sound, sustained advances in the sciences underlying medicine, public health, and social services.

It is easy to resonate with the aim of lengthening human life and improving quality of life during the years in which people survive. In order to work toward these objectives, it is necessary to understand the factors associated with longer life and reduced illness and disability. It is typically assumed that the best mechanism for increasing life expectancy and reducing the burden of illness and disability is investment in medical care. However, a variety of different analyses using several different research methodologies have shown that medical care accounts for only a small portion of the variation in health outcomes.[2,3] The traditional biomedical model is thus limited in its ability to foster health and well-being, in part because the biomedical model often focuses exclusively on measures of biological process rather than the more global goal of helping people live longer and higher quality lives. Understanding how to improve overall health and well-being thus requires explorations beyond the health care system.

NIH's mission is to seek fundamental knowledge about the nature and behavior of living systems and to apply that knowledge to enhance health, lengthen life, and reduce illness and disability.

The purpose of this book is to gain a better understanding of the multitude of factors that determine longer life and improved quality of life in the years a person is alive. While the emphasis is primarily on the social and behavioral determinants that largely impact the health and well-being of individuals, this publication also addresses quality of life factors and determinants more broadly. This book originated with the Office of Behavioral and Social Sciences Research (OBSSR) in NIH's Office of the Director. As part of long-term strategic priority setting, we sought to identify the most important factors that influence the length and quality of human life. We were given the opportunity to invite a group of the most distinguished scholars to contribute chapters summarizing current research. The authors were asked to summarize what is currently known and to suggest directions for future scientific research. Each chapter in this book considers an area of investigation and ends with suggestions for future research and implications of current research for policy and practice. To contextualize these chapters and to highlight the pressing nature of the questions this book tackles, we first provide a summary of the state of Americans' health and well-being in comparison to our international peers and present more background concerning the limitations of current approaches to improving health and well-being.

U.S. Life Expectancy in International Perspective

Extending life and improving the quality of life may require greater attention to factors beyond of the current health care system. In Figure 1, data from the Organization for Economic Co-operation and Development (OECD) are used to graph total health care expenditures for various countries (x-axis) against life expectancy (y-axis). The United States is an extreme outlier in terms of expenditures. We spend more than 17 percent of our gross domestic product on health care, while most of our economic competitors spend about 10 percent. If the United States reduced its health care expenditures to the level of most European countries, we would save over $1 trillion per year—

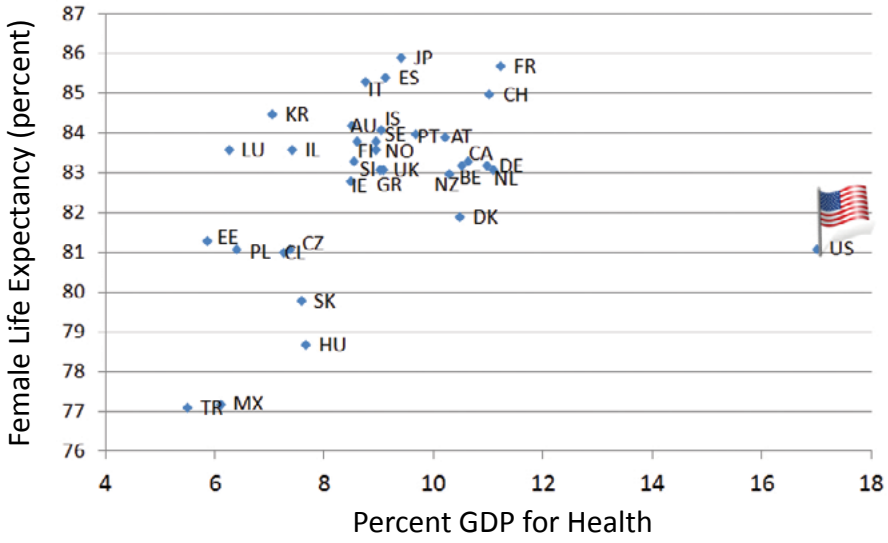


Figure 1. Relationship between percent of GDP spent on health care and female life expectancy in OECD countries

Source: Created using data from the Organization for Economic Co-operation and Development (OECD), 2011. When missing, life expectancy data estimates were imputed from prior year.

or about half the amount as the total public debt held by the U.S. Federal Reserve. Given our very high expenditures on health care, it is appropriate to explore how well the United States is doing in achieving its goal of longer life and improved quality of life.

International studies of life expectancy have gained particular attention in the last few years. These studies tend to show that the life expectancy advantage experienced by American citizens in comparison to other countries has been on the decline. One study from the National Research Council considered current life expectancy for 50-year-old women between the years 1955 and 2010.[4] Current life expectancy is the number of years of life on average remaining once a milestone age has been reached. So, current life expectancy for 50-year-old women is the median number of years of life remaining following the 50th birthday. In 1955 women in America ranked about 12th in the world on this indicator. By 2006, they had slipped to about the 26th position, just below Korea and Malta. In a life expectancy comparison of 10 wealthy countries, women in America were 3rd out of 10 in 1955, but they were 9th out 10 in 2006. Among the many countries with more rapid increases in life expectancy in the 50 years following 1955 were Japan, France, and Spain. Japan, for example, was considerably below the United States in 1955 and now is many years ahead.[4]

In response to these findings, the Office of Behavioral and Social Sciences Research, along with the National Institute on Aging, sponsored another study that compared life expectancy in the United States against 17 peer countries.[5] These comparison countries were primarily in Western Europe, but they also included Australia, Japan, and Canada. The results of the comparison are quite disturbing. Among the 17 countries studied, the United States had the second highest mortality rate from non-communicable diseases. Mortality from communicable diseases was fourth from the bottom for the United States. The United States had the third highest AIDS rates, exceeded only by Brazil and South Africa. And, the AIDS incidence in the United States was 122 per million, which is about nine times the average of countries in the OECD.

We have known for some time that U.S. life expectancy at birth is not keeping pace with other developed countries. Although our life expectancies are increasing, the rate of increase is much slower than it has been for our economic competitors. This trend has been developing over the course of several decades. Perhaps the most surprising finding from the Institute of Medicine (IOM) study[5] concerned years of life lost prior to age 50. The IOM committee considered international differences in the probability of celebrating a 50th birthday. On this indicator, the United States was last among the 17 comparison countries for both men and women. U.S. losses in life expectancy prior to age 50 are about double the rate observed in Sweden. Perhaps most disturbing is that this problem profoundly affects women. Figure 2 shows the trend in years of life lost in 21 high-income countries between the years 1980 and 2006. For men, the United States started at the low end of the distribution and worked its way to the bottom. For women, the United States started near the bottom and now has gone off the scale in relation to the comparison countries.[4]

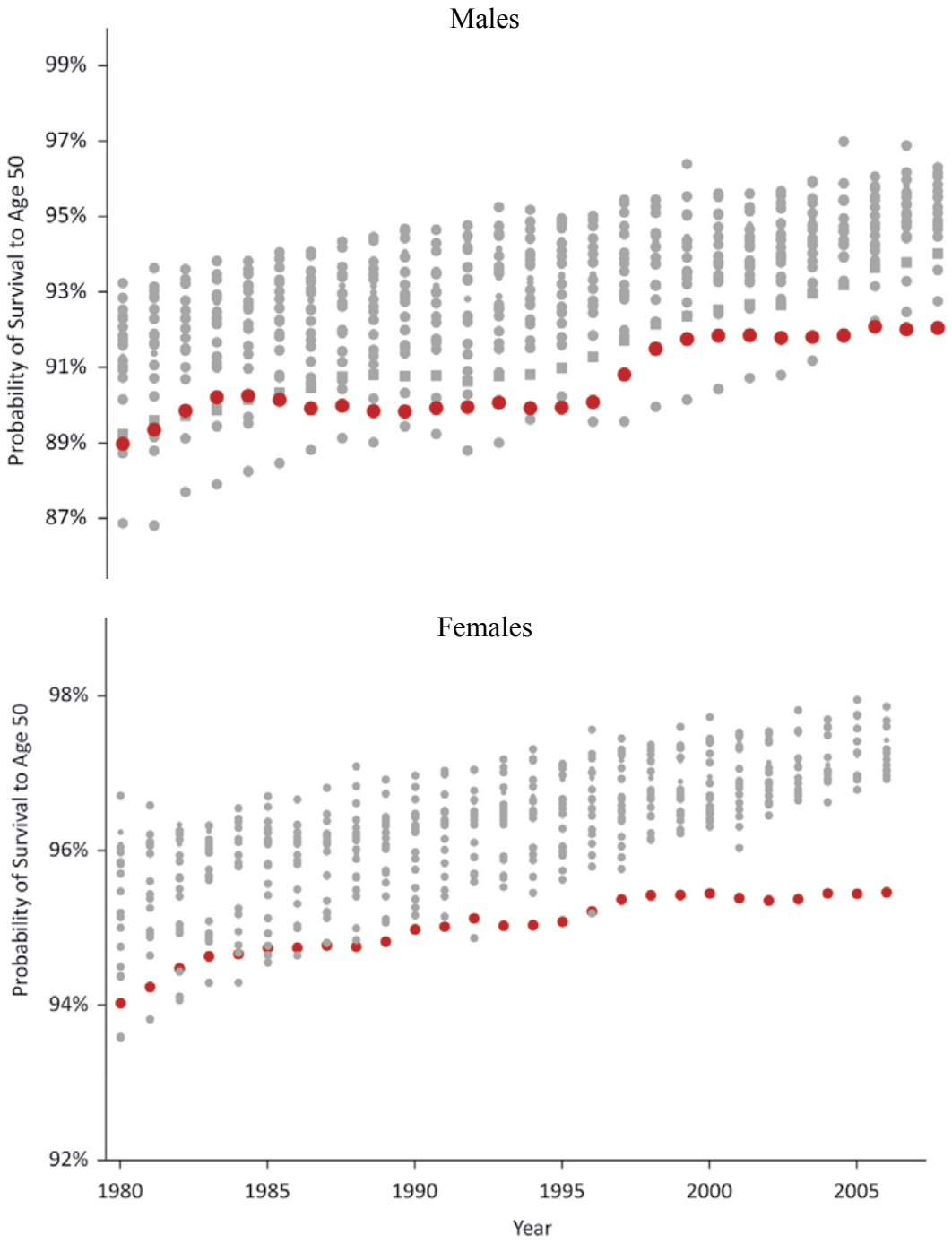


Figure 2. Probability of survival to age 50 in 21 high-income countries, 1980-2006

Source: National Research Council: Explaining divergent levels of longevity in high income countries, 2010. Used with permission.

Advances in Medical Care

It is often argued that the United States has the very best medical care in the world. So, we would expect advances in medical therapies to address many of our health care problems. The reality is that recent clinical trials often have not demonstrated the level of benefit that the public expects from medical therapies. In fact, most recent large randomized clinical trials have failed to show the expected benefit of medical and surgical therapies.[6]

To more accurately assess the benefits and limitations of current medical interventions for health maintenance and prevention, it is crucial to understand the "Patient Centered Outcomes Research (PCOR)" perspective.[7] The PCOR perspective argues that physiological measures are only important if they relate to life duration or life quality. Blood pressure, for example, is a meaningful biological measure because it is highly predictive of early death or disability associated with myocardial infarction or stroke. In contrast, other measures and outcome variables less clearly relate to the twin objectives of improved life quality or lengthened life expectancy. One example is catecholamine variations in response to acute stress, which are less directly linked to the objectives and outcomes that may concern health science researchers.

A second key perspective arising from PCOR is the focus on all-cause mortality as opposed to disease-specific mortality.[8] A variety of large clinical trials in medicine have demonstrated reductions in one cause of death but compensatory increases in other causes of death.[9] One example that helps justify the PCOR perspective is illustrated by the Physicians Health Trial. In this study, approximately 22,000 physicians were randomly assigned to take either 325 mg of aspirin every other day or a placebo. When the data were first analyzed, significantly fewer physicians in the aspirin component had died of myocardial infarction. However, considering all causes of cardiovascular death, the number of physicians who had died was exactly the same in the aspirin and placebo groups (see Figure 3). All of these deaths were in the study period, and all were considered premature deaths.[10] In essence, aspirin had altered the course of one possible cause of death (myocardial infarction), but the medication ultimately did not extend participants' life expectancy[11]; it merely changed what was recorded on the death certificate. From a patient's perspective, we would argue that people and their families are most concerned with an individual's vital status and less concerned with a specific cause of death.

With these PCOR concepts of longer life and higher quality of life in mind, it is helpful to consider the ACCORD trial of aggressive therapy for the treatment of non-insulin-dependent diabetes mellitus.[12] Patients were randomly assigned to standard therapy or intensive therapy. The intensive therapy significantly changed biological outcomes in the expected direction. Specifically, those assigned to intensive therapy had significantly lower levels of glycosylated hemoglobin. From a traditional perspective, the treatment achieved its goal. However, long-term followup considered total mortality and deaths from cardiovascular disease. Considering all-cause mortality, those assigned to intensive therapy had a higher probability of cardiovascular death in comparison to the standard therapy condition.

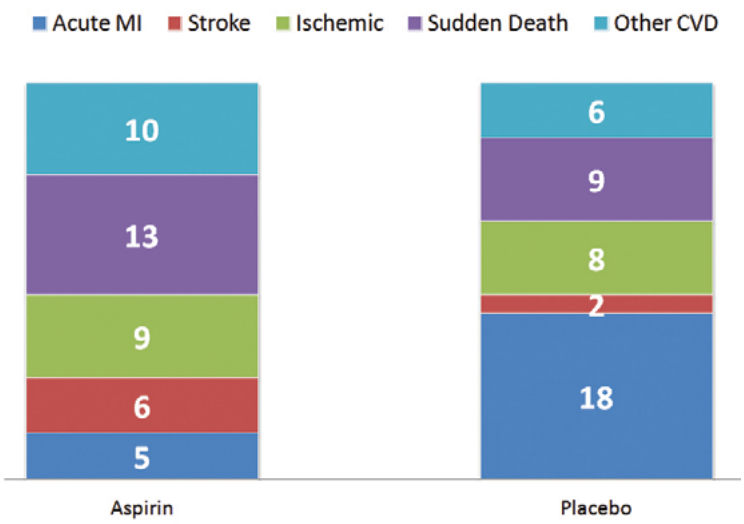


Figure 3. Total mortality in the aspirin component of the Physician's Health Study
Source: Adapted from the Final Report on the Aspirin Component of the Physicians' Health Study. N Engl J Med 1989;321:129-35.

The ACCORD trial is just one of many randomized clinical trials with similar results. Trials considering intensive therapy for anemia suggest that agents that increase red blood cells do their job and bring hemoglobin counts toward normal. Yet patients in these conditions have a higher probability of renal failure requiring dialysis and other adverse outcomes.[13] Large studies on hormone replacement therapy usually show that estrogen levels are raised toward normal premenopausal levels. Yet the consequences for patients, from a PCOR perspective, are often poorer rather than better.[14]

Certainly, we need to continually evaluate promising new therapies; however, we also need to devote more attention to nontraditional determinants of health outcome. For instance, we now know that medicines and surgery, despite their value and importance, do not explain most of the variation in human health outcomes. For example, Schroeder[15] estimated the contribution of a variety of different factors to premature death. His analysis suggested that health care contributes about 10 percent of the variation in outcomes, while environmental exposures contribute about 5 percent. Genetic predisposition may account for about 30 percent, but behavioral patterns contribute about 40 percent, and another 15 percent is contributed by social circumstances. Other methodologies lead to similar conclusions. For example, epidemiologists typically look for the relationship between risk factors and specific diseases. But, when they look at the relationship between behavioral risk factors and death from all causes, estimates typically suggest that behavioral factors, including tobacco smoking, physical activity, and diet are by far the largest contributors to poorer outcomes. These factors usually are much stronger than biochemical measures, including cholesterol, inflammation, and even early detection of breast cancer.[16]

These findings highlight the crucial importance of considering behavioral factors and social context when attempting to understand the elements impacting patient-centered outcomes, such as quality of life, health maintenance and improvement, and life expectancy.

Behavioral and Social Factors Underlying Disease

Although the exact number varies across different methods, most analyses suggest that behavioral and social factors account for at least half of the variation in health outcomes. One interesting example is provided by the Marmot report from the United Kingdom.[17] An expert panel reviewed a variety of factors associated with life expectancy. For example, Marmot and colleagues in the United Kingdom studied the relationship between neighborhood income deprivation and life expectancy and disability-free life expectancy. They found that there is a systematic relationship between these variables, with individuals from the most deprived neighborhoods having the shortest life expectancies.[17]

A growing body of evidence reinforces the observation that social factors have a profound impact on life expectancy. One example comes through the Los Angeles County Public Health Department, which reports on life expectancy by sex and racial/ethnic background.[18] The difference between Asian women with an average life expectancy of nearly 89 years and African-American men with a life expectancy of about 70 years is a full 19 years! There have been significant improvements in the life expectancy of black males in recent years, but on average, white women still live about 9 years longer than black men.[19] We do not fully understand the mechanisms underlying these enormous differences in longevity, but we do recognize that the effects are profound.

While people with infections benefit from rapid diagnosis and treatment, we now live in an age in which most of the burden of illness and disability is associated with multiple chronic conditions and non-communicable diseases, including coronary heart disease, cancer, and diabetes. These conditions are expensive to treat, and their outcomes vary greatly by socioeconomic status and lifestyle. Health behaviors are the biggest risk factor for these conditions.[20] Modification of risky health behaviors is central to the successful management of chronic conditions, and social factors—including income, social support, and access to information—play a crucial role in health outcomes.

Another very interesting finding that has emerged from recent research is the impact of behavioral and social factors in relation to diagnoses. Wennberg and colleagues studied the wide variations in health care costs across 306 hospital service areas in the United States.[21] The investigators used Medicare claims data to estimate what factors account for the variations in cost and in mortality. They used three different factors to explain this variation: one, a medical hierarchical conditions categories (HCC) index that adjusts for medical diagnoses; two, a poverty index; and three, a behavioral health index based simply on the number of people with hip fractures or strokes and behavioral responses including self-rated health, obesity, and smoking. The analysis was adjusted for demographic factors including age, sex, and race. Figure 4 summarizes the results of the analysis. The adjustment factor based on clusters of diagnoses does the best job in explaining health care costs. However, it does the very worst job in explaining mortality. Conversely, the behavioral index does a much better job than the medical index of explaining variation in mortality. The most striking finding, consistent with several other lines of evidence, is that medical care explains only a small amount of variation in health mortality rates.

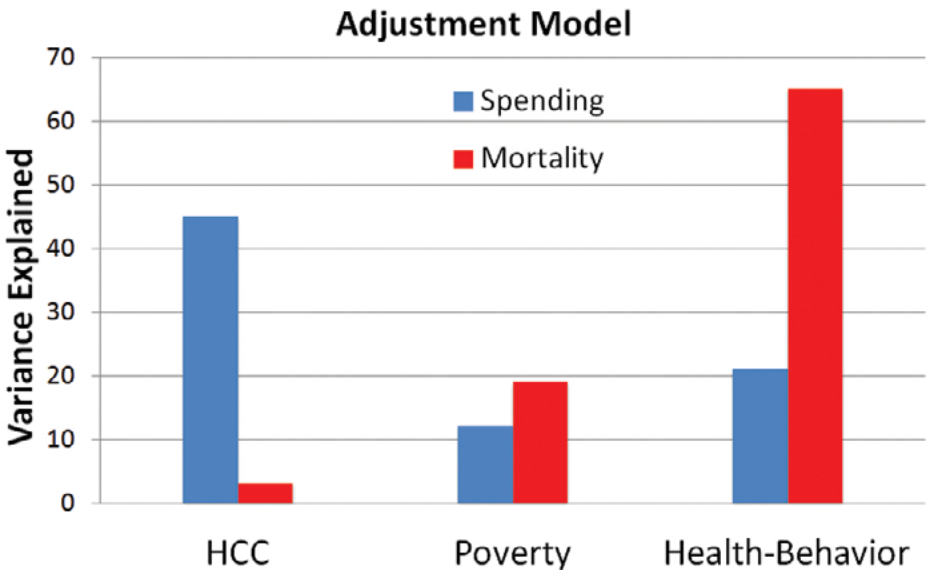


Figure 4. Proportion of variance in spending and mortality using models that adjust for medical conditions (HCC), poverty, or health and behavioral indexes

Source: Data adapted from Wennberg DE, Sharp SM, Bevan G, et al. A population health approach to reducing observational intensity bias in health risk adjustment: cross sectional analysis of insurance claims. Br Med J 2014;348:g2392.

The United States now spends more on health care than any other country in the world. In fact, on a per capita basis, no other country comes close. Brush[a] estimates that the total expenditure on health services in the United States is approximately $2.9 trillion/year. Given this huge expenditure, we need to ask the question, "What is the rate of return on our investment?" With only about 10 percent of the variability in life expectancy associated with health care,[3] and with about half of health outcomes associated with health behaviors and another other 40 percent as explained by a variety of other nonmedical factors, we must consider the cost side of the equation.

Of the $2.9 trillion in annual health-related expenditures, about 97 percent is devoted to health care, while only 3 percent is devoted to factors outside of the health care system. In other words, 97 percent of the investment is chasing the potential for 10 percent of the benefit, while as little as 3 percent of the expenditure is devoted to factors that may explain 50 to 90 percent of the potential benefit.[b] We believe that it is time to re-examine what is known about the relationship between factors outside of the health care system and health outcomes. Further, we need to develop a new research agenda directed toward maximizing the use of our resources to produce health. Although behavioral and social factors are likely to explain up to five times as much of the variation in health outcomes in comparison to medical care, basic biological mechanisms remain the almost exclusive

[a] Unpublished data courtesy of Rick Brush, Collective Health, 2014.

[b] We recognize that this analogy is oversimplified. For example, we have substantial investments in water safety, pollution control, and public safety. Even though these expenditures are designed to protect health, they typically are not included in estimates of health resources.

focus of biomedical research. While we strongly recognize the value of basic biomedical research and of advances in biomedical technology, there remains a pressing need to take a broader view in our approach to research and medical care.

Scope and Importance of Current Effort

Clearly we have an inadequate understanding of the non-medical care influences that help to explain health outcomes. Many of the chapters in this book focus on specific aspects of quality of life and mortality that are associated with behavioral and social variables. Explored in detail are factors that contribute to premature mortality and/or lower quality of life, including cigarette smoking (Abrams), unintentional injuries (Sleet and Gielen), insufficient physical activity (Sallis and Carlson), low educational attainment (Zimmerman and Woolf), challenging and/or unequal social circumstances (Pickett and Wilkinson; Williams and Purdie-Vaughn), ongoing HIV infection and prevalence (Holtgrave), and workplace policies (Berkman). In addition to highlighting the impact that these factors and behaviors can have on quality of life, each of the above chapters also provides recommendations for relevant policies, practices, and interventions that could help improve overall life expectancy and well-being.

Other chapters explore the rich public health dimensions to quality of life and life expectancy. For instance, Preston's piece provides an elegant overview of various social, behavioral, and public health factors that have contributed to improvements in health since the mid-19th century in different areas of the world. Stewart and Cutler highlight six distinct behavioral factors (smoking, obesity, heavy alcohol use, and unsafe use of motor vehicles, firearms, and poisonous substances) and illustrate how each has influenced health-related quality of life in the U.S. context. Baldwin's chapter asserts the grave impact that the emerging use of tobacco by youth in middle and lower income countries could have on the incidence of non-communicable illnesses in the developing world.

As noted earlier, the intersection of biological and social factors must also be considered when studying the determinants of quality of life and life expectancy. McEwen provides a nuanced overview of how neurological and endocrinological responses to varying degrees of social, environmental, and physical stress can affect lifelong well-being, while Boyce highlights the long-term impact that early exposure to adversity or trauma can have on health and quality of life. Adler and Prather pull the various contributions together with a strong case for integrating biological, biomedical, behavioral, and social science research perspectives to most effectively promote and advance individual and population-level health and longevity.

Several chapters in this book raise questions about the most promising systemic and methodological approaches to achieving longer life and to reducing the burden of illness and disability for populations. Some argue that our established methods of medical training (Satterfield and Carney) and mental health service delivery (Kazdin) are not sufficient for addressing current and emerging population health needs. Instead, cost-effectiveness and cost-benefit analyses can help reveal those modes of intervention and health care delivery that can have a strong impact on public health without depleting resources that could be used for other valuable expenditures (Russell), and health

impact assessments can provide prospective insight into the potential health effects of public health and social service interventions (Teutsch, Butler, Simon, and Fielding). Still other chapters shine light on the importance of cross-disciplinary approaches and advanced behavioral and social science methodologies for improving behavioral health policy and practice (Frank and Glied), containing the spread of emerging infectious diseases (Orr and colleagues), and developing novel population health interventions (Marteau). Finally, we conclude the volume with a summary of some of the lessons learned and suggestions for future research investigation.

Our major institutions have set the goal of improving human health by extending life expectancy and improving health-related quality of life. Usually the preferred tool for improving health is greater investment in medical care. Yet, despite the accomplishments of modern medicine, health care is limited in its potential to achieve the goals of longer and higher quality life. Many of the factors that determine health outcome are beyond the reach of the traditional biomedical model. Our ultimate goal in producing this book is to encourage the development of better evidence to inform medical and public health practice and to develop public policies that will result in better health for our populations. We hope this book initiates these important new directions.

Acknowledgments

Portions of this chapter are similar to an article by Robert M. Kaplan, Behavior change and reducing health disparities. Prev Med 2014;68:5-10. The opinions presented herein are those of the authors and do not necessarily represent the official position of the Agency for Healthcare Research and Quality, the National Institutes of Health, or the U.S. Department of Health and Human Services.

Authors' Affiliations

Robert M. Kaplan, PhD, Agency for Healthcare Research and Quality; formerly Office of Behavioral and Social Sciences Research, National Institutes of Health, Office of the Director.
Daryn H. David, PhD, and Michael L. Spittel, PhD, Office of Behavioral and Social Sciences Research, National Institutes of Health, Office of the Director.

Address correspondence to: Robert M. Kaplan, PhD, Office of the Director, Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850;
email Robert.Kaplan@ahrq.hhs.gov.

References

1. Healthy people 2000. Atlanta, GA: Centers for Disease Control and Prevention; September 1990. Available at http://www.cdc.gov/nchs/healthy_people/hp2000 htm. Accessed January 26, 2015.
2. Murray CJ, Vos T, Lozano R, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012;380:2197-223.
3. Schroeder SA. Shattuck Lecture. We can do better--improving the health of the American people. New Engl J Med 2007;357:1221-8.
4. Crimmins EM, Preston SH, Cohen B. Explaining divergent levels of longevity in High-income countries. Washington, DC: National Academies Press; 2011.
5. Woolf SH, Laudan A, Panel on Understanding Cross-National Health Differences Among High-Income Countries. U.S. health in international perspective: shorter lives, poorer health. Washington, DC: National Academy of Sciences, Institute of Medicine; 2013. Available at http://tinyurl.com/nnldgwg. Accessed January 22, 2015.
6. Gordon D, Taddei-Peters W, Mascette A, et al. Publication of trials funded by the National Heart, Lung, and Blood Institute. New Engl J Med 2013;369:1926-34.
7. Kaplan RM. Behavior change and reducing health disparities. Prev Med 2014;68:5-10.
8. Kaplan RM. Behavior as the central outcome in health care. Am Psychol 1990;45:1211-20.
9. Kaplan RM. The Ziggy theorem: toward an outcomes-focused health psychology. Health Psychol 1994;13:451-60.
10. Final report on the aspirin component of the ongoing Physicians' Health Study. Steering Committee of the Physicians' Health Study Research Group. N Engl J Med 1989;321:129-35.
11. Kaplan RM. Health outcome models for policy analysis. Health Psychol 1989;8:723-35.
12. Action to Control Cardiovascular Risk in Diabetes Study Group, Gerstein HC, Miller ME, et al. Effects of intensive glucose lowering in type 2 diabetes. Nw Engl J Med 2008;358:2545-59.
13. Drueke TB, Locatelli F, Clyne N, et al. Normalization of hemoglobin level in patients with chronic kidney disease and anemia. New Engl J Med 2006;355:2071-84.
14. Chlebowski RT, Hendrix SL, Langer RD, et al. Influence of estrogen plus progestin on breast cancer and mammography in healthy postmenopausal women. JAMA 2003;289:3243-53.
15. Schroeder SA. Shattuck Lecture. We can do better—improving the health of the American people. N Engl J Med 2007;357(12):1221-8.
16. Mokdad AH, Marks JS, Stroup DF, et al. Actual causes of death in the United States, 2000. JAMA 2004;291(10):1238-45.
17. Marmot M, Allen J, Goldblatt P, et al. Fair society, healthy lives: the Marmot review. London: Department of Health, Strategic Review of Health Inequalities in England Post-2010; 2010. Available at http://www.instituteofhealthequity.org/projects/fair-society-healthy-lives-the-marmot-review. Accessed January 22, , 2015.
18. Health LACDoP. Annual report, County of Los Angeles Public Health. Los Angeles: Department of Public Health; 2012-2013. Available at http://publichealth.lacounty.gov/docs/ar2012-2013.pdf. Accessed February 19, 2015.
19. Hoyert DL, Xu J. Deaths: preliminary data for 2011. Nat Vital Stat Rep 2012;61(6).
20. Stevens D, Siegel K, Smith R. Global interest in addressing non-communicable disease. Lancet 2007;370(9603):1901-2.
21. Wennberg DE, Sharp SM, Bevan G, et al. A population health approach to reducing observational intensity bias in health risk adjustment: cross-sectional analysis of insurance claims. Br Med J 2014;348:g2392.

Robert M. Kaplan, PhD, is Chief Science Officer, Agency for Healthcare Research and Quality (AHRQ). Formerly, he was Associate Director for Behavioral and Social Sciences and Director of the Office of Behavioral and Social Sciences Research in the Office of the Director, National Institutes of Health; Distinguished Professor of Health Services at the University of California Los Angeles (UCLA); and Distinguished Professor of Medicine at the UCLA David Geffen School of Medicine. In 2005, Dr. Kaplan was elected to the Institute of Medicine of the National Academy of Sciences.

Daryn H. David, PhD, is a researcher and clinical psychologist who has expertise in developing and evaluating innovative social science, psychological, and educational programs. Currently, she is a Clinical Instructor in the Department of Psychiatry at the Yale School of Medicine. She gained national science policy experience during a 2-year American Association for the Advancement of Science (AAAS) Science & Technology Policy Fellowship in the Office of Behavioral and Social Sciences Research at the National Institutes of Health.

Michael L. Spittel, PhD, is a health scientist administrator in the Office of Behavioral and Social Sciences Research, Office of the Director, National Institutes of Health. Previously, he was a program officer at the Demographic and Behavior Sciences Branch (DBSB), Eunice Kennedy Shriver National Institute of Child Health and Human Development. He also served as the program officer/scientist for the National Longitudinal Study of Adolescent to Adult Health (Add Health), Data Sharing for Demographic Research, Community Child Health Network, and co-managed the DBSB training program which supports pre- and postdoctoral researchers in demography.

Demographic and Social Epidemiological Perspectives on Population Health

Income Inequality and Health: A Causal Review

Kate E. Pickett and Richard G. Wilkinson

Abstract

There is a very large literature examining income inequality in relation to health. Early reviews came to different interpretations of the evidence, though a large majority of studies reported that health tended to be worse in more unequal societies. More recent studies, not included in those reviews, provide substantial new evidence. Our purpose in this chapter is to assess whether or not wider income differences play a causal role leading to worse health. We conducted a literature review within an epidemiological causal framework and inferred the likelihood of a causal relationship between income inequality and health (including violence) by considering the evidence as a whole. The body of evidence strongly suggests that income inequality affects population health and well-being. The major causal criteria of temporality, biological plausibility, consistency, and lack of alternative explanations are well supported. Of the small minority of studies that found no association, most can be explained by income inequality being measured at an inappropriate scale, the inclusion of mediating variables as controls, use of subjective rather than objective measures of health, or followup periods that were too short. The evidence that large income differences have damaging health and social consequences is strong, and in most countries, inequality is increasing. Narrowing the gap will improve the health and well-being of populations.

Introduction

World leaders, including the U.S. President, the U.K. Prime Minister, the Pope, and leaders at the International Monetary Fund, the United Nations, the World Bank, and the World Economic Forum have all described income inequality as one of the most important problems of our time, and several have emphasized its social costs.[1-7] Inequality is increasing in most regions of the world, rapidly in most rich countries over the past three decades.[8,9] There is a very large literature examining income inequality in relation to health. Early reviews came to different interpretations of the evidence, though a majority of studies reported that health tended to be worse in more unequal societies.[10-14] More recent studies, not included in those reviews, provide substantial new evidence.

There is also growing evidence that a wide range of social outcomes, associated with disadvantage within societies, are more common in societies with bigger income differences between rich and poor. Although our objective in this chapter is to assess whether or not wider income differences

play a causal role leading to worse health (including the public health issue of violence), we consider studies of other social outcomes where they affect interpretation of the health data.

The first task is to clarify the causal hypothesis and how it has developed as research has progressed. Research was initially focused simply on whether health was worse in more unequal societies, but there is now growing evidence to suggest that this should be seen as part of a wider tendency for a broad range of outcomes with negative social gradients (i.e. more prevalent where social status is lower) to be more common in societies with bigger income differences between rich and poor. Rather than this pattern being confined to physical health, it may apply also to mental health and public health issues such as violence, teenage births, child well-being, obesity, and more.

Whether causality is tested in relation to a hypothesis confined to a relationship between inequality and physical health, or whether the hypothesis extends to problems with social gradients more generally, has important implications for understanding possible causal mechanisms, mediators and confounders.

In this chapter, we will focus on the strongest and most important claim underpinning an effect of inequality on health: that large income differences between rich and poor lead to an increasing frequency of most of the problems associated with low social status within societies. Figure 1 provides an illustration of the relationships with which this chapter is concerned. It shows a cross-sectional association between income inequality in developed countries and an index that combines data on life expectancy, mental illness, obesity, infant mortality, teenage births, homicides, imprisonment, educational attainment, distrust, and social mobility. Raw scores for each variable were converted to z-scores, and each country was given its average z-score.[15]

History

The hypothesis that problems (including poor health) associated with low social status are more common in more unequal societies can be traced back to independent roots in papers on homicide rates and mortality rates. The research literature on homicide and inequality goes back at least 40 years, to a demonstration that they were positively associated among U.S. States.[16] The earliest paper on mortality and income inequality – some 35 years ago – showed a cross-sectional association between Gini coefficients of income inequality and both infant mortality and life expectancy at age 5 among a group of 56 developed and developing countries.[17] By 1993, a meta-analysis of some 34 studies concluded that there was a robust tendency for violence to be more common where income differences were larger.[18] The research on income inequality and health expanded rapidly after the first papers were published in journals of epidemiology and public health.[19,20] By 2006, our review of papers on income inequality and health identified 168 analyses, the overwhelming majority of which showed a positive association.[13] The two literatures – in criminology and sociology on the one hand, and epidemiology and public health on the other – developed independently and unaware of each other until the late 1990s.[21,22]

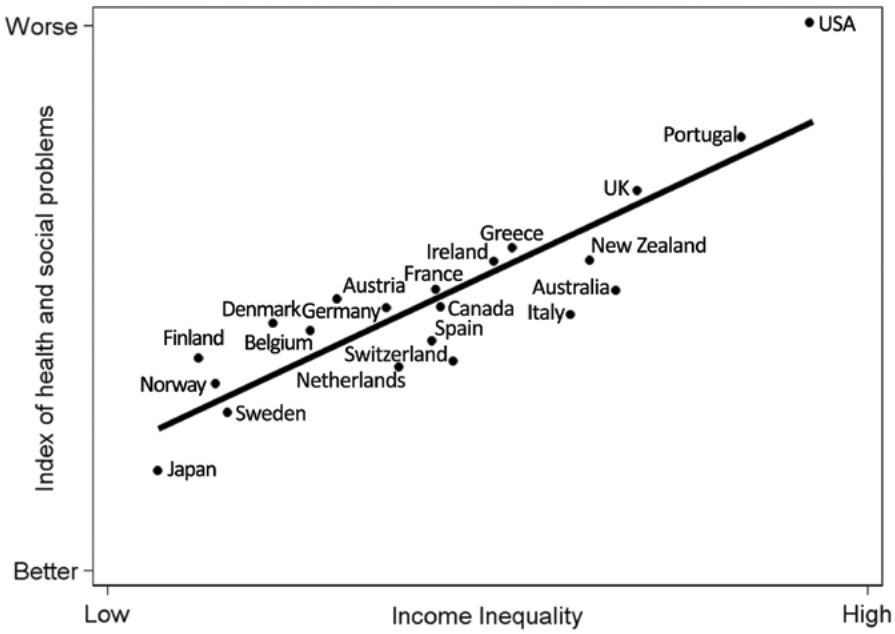


Figure 1. Index of health and social problems in relation to income inequality in rich countries
Source: Wilkinson R, Pickett K. The spirit level. London: Allen Lane; 2009. Used with permission.

Income inequality is measured by the ratio of incomes among the richest compared with the poorest 20 percent in each country. The index combines data on: life expectancy, mental illness, obesity, infant mortality, teenage births, homicides, imprisonment, educational attainment, distrust, and social mobility. Raw scores for each variable were converted to z-scores, and each country was given its average z-score.[15]

It was only in 2005 and 2006,[23-25] as researchers began to show that the correlates of inequality included teenage birth rates, obesity, and mental illness, that it started to look as if a more general explanatory hypothesis was needed than those which had addressed only physical health and violence. On the assumption that social gradients were often evidence that an outcome was sensitive to social status differentiation, we formed the hypothesis that greater inequality might act to strengthen the effects of socioeconomic status differentiation among outcomes with social gradients.

We tested this hypothesis by analyzing whether or not outcomes with steeper social gradients had stronger associations with societal inequality. We selected 10 different death rates, some with weaker and some with stronger social gradients, as measured by their correlation with county median income, among the 3,139 counties of the United States.[26] In a multilevel model controlling for the effects of county income, we then estimated the correlations of these death rates with State income inequality. The results, shown in Figure 2, provided strong confirmation of the hypothesis.

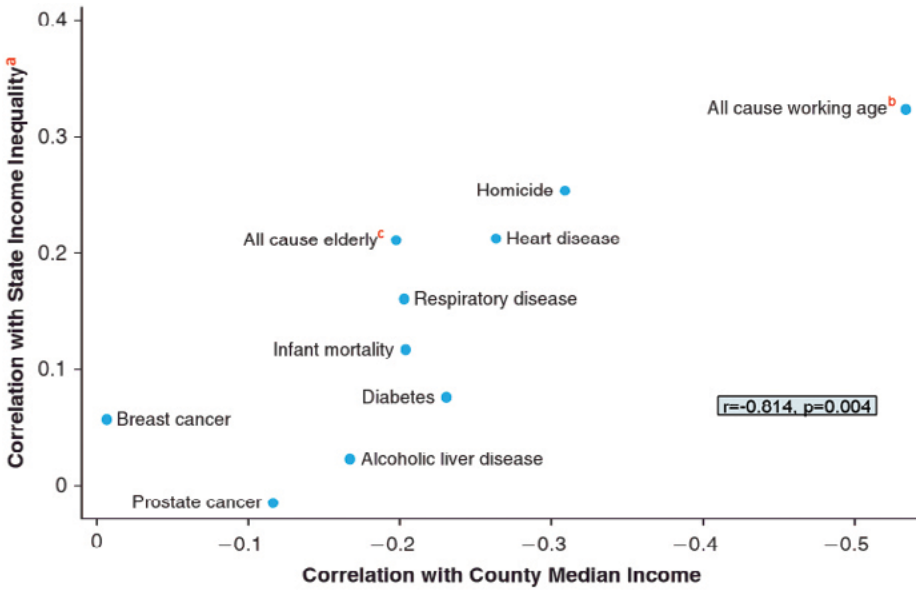


Figure 2. The effect of county-level median household income in relation to contextual effect of State-level income inequality

Source: Wilkinson RG, Pickett KE. Income inequality and socioeconomic gradients in mortality. Am J Public Health 2008;98(4):699-704. Used with permission.

Notes: Standardized beta coefficient from multilevel model, controlling for county-level income. r = -0.814; P = .004. [a] = Standardized parameter estimates (B) from multilevel model after county-level income was controlled. [b] = Aged 25-64 years. [c] = Aged ≥65 years.

This work was followed by studies that examined the association among rich developed countries between income inequality and a number of health and well-being outcomes, including the UNICEF (United Nations Children's Fund) Index of Child Wellbeing and the separate components of the Index of Health and Social Problems shown in Figure 1.[15,27,28]

Popperian Theory Testing

The philosopher of science, Sir Karl Popper, taught that the best evidence of the value of a theory is provided by testing its novel predictions.[29,30] A successful theory was "corroborated" (but could never be finally proven true) if it accurately predicted the results of scientific observations that had not previously been expected. The initial evidence of a relation between income inequality and population health using international data was first explicitly tested and confirmed in 1996 by two groups working independently at the universities of Harvard and Michigan, who looked to see if the same relationship could be found among the 50 U.S. States.[31,32] There are now very large numbers – hundreds – of replications of these findings in many different settings in societies at all levels of development. Even the more unequal provinces of China have been found to have significantly less good health.[33] There now can be no doubt that worse health is at least associated with greater inequality.

The tendency for more unequal societies to have higher homicide rates has also been replicated many times (for one recent review see Rufrancos et al.[34]). To suggest that a relationship is causal means predicting a subsidiary hypothesis about a mediating mechanism. A testable prediction of a causal mechanism was first suggested on the basis of qualitative impressions only. The hypothesis was that more equal societies were healthier because they were more cohesive and enjoyed better social relations.[35] A year later, that prediction was tested quantitatively: path analysis showed that the relationship between greater equality and lower death rates among the U.S. States was mediated by social capital (operationalized as group membership and social trust).[36]

Another way in which testable predictions have emerged is when new national data have become available that fit previously established relationships between inequality and an outcome measure. This happened when new data on social mobility and on mental illness rates became available for several additional countries, and the data were found to fit previously established relationships between those outcomes and inequality.[37]

Lastly, the first papers suggesting that mental illness was more common in more unequal societies used general measures of mental illness from the World Health Organization (WHO).[25,38] The picture has since been filled out by papers that show more specific forms of mental illness – including depression, schizophrenia, and psychotic symptoms – are all more common in more unequal societies.[39-41]

There have been a small proportion of negative findings throughout the development of this field. We will discuss some of the reasons for variations in findings at appropriate points later in this chapter.

Our aim in this review is to go beyond the "counting" methodology of previous major reviews; these mostly divided studies into supportive, mixed, and unsupportive of the income inequality-health relationship and counted them. Among mixed studies, which showed some but not all relationships to be significant, results within a study might vary by geographic scale or by health outcome, measure of inequality, sex or age of subjects, or other variables; merely counting these adds nothing to interpretation, even if we count more. Instead, we conduct a causal review to give a more structured and coherent framework to our examination of the literature. We have aimed to incorporate all new studies that illuminate relevant causal processes.

Epidemiological Criteria for Causality

In observational epidemiology, causality cannot be proven or disproven by any single study – there are no "black swans" – just because income inequality might not affect some health outcomes, or not in some times or places or for some populations, does not mean that it isn't a causal relationship in other contexts. Instead, in epidemiology, a body of evidence needs to be considered, usually including non-epidemiological studies, to judge whether or not an exposure-outcome relationship is causal.

Causal criteria were first proposed in the Surgeon General's Report of 1964, and then refined by Sir Austin Bradford Hill in 1965, in the context of examining the evidence linking cigarette smoking to lung cancer.[42,43] The use of causal criteria, indeed even the term "criteria," has been contentious, especially if the criteria are used as a simple checklist or algorithm; however, when used as a framework for thoughtful inference, and used to consider competing causal theories by focusing on crucial observations, they offer a useful organizing structure for critical review.[44,45] Bradford Hill's original nine "criteria" have been further refined with modern usage and, as indicated in Table 1, four are considered of major importance.[46]

Consistency

There are now perhaps 300 peer-reviewed studies of the relation between income inequality and measures of health or homicide. They include both ecological and multilevel studies using cross-sectional, cohort, and time-series designs in many time periods. Looking at bivariate correlations before the use of control variables, the most recent full review found only 6 percent of studies (8 of 128) did not find at least one significant association between greater inequality and worse health.[13] After the use of many different kinds of control variables, many of which might be on the causal pathway, 70 percent of the studies reporting either positive or negative (but not "mixed") results found only significant associations between higher inequality and worse outcomes.

These relationships have been found in a wide variety of settings. Much research has focused on the rich, developed, market economies (United Kingdom, United States, Western Europe, Japan, Singapore, Australia, and New Zealand) and analyses of the 50 U.S. States.[15,28] Some studies have included developing, emerging, and developed nations, while others have focused on developing countries in particular. Infant mortality is the health outcome most often shown to be positively correlated with income inequality in less developed nations,[47-49] although there are also associations with lower life expectancy, higher HIV prevalence, and higher homicide rates.[47,50,51] Some studies have focused on particular world regions. For example, Marmot and Bobak found larger declines in life expectancy in the more unequal countries of Eastern Europe following the dissolution of the Soviet Union.[52] Biggs and colleagues studied 22 Latin American countries from 1960-2007 and found a substantial relationship between income inequality, life expectancy, infant mortality, and tuberculosis mortality rates; they also reported that when inequality was rising, economic growth was related to only a modest improvement in health, whereas during periods of decreasing inequality, there was a very strong effect of rising gross domestic product (GDP).[53] Other studies have shown an association between income inequality and health across states/regions within nations, including, for example, in Argentina,[54] Canada,[55] Brazil,[56] Chile,[57] China,[33] Ecuador,[58] India,[59] Italy,[60] Japan,[61] and Russia.[62]

The geographical scale at which income inequality is measured is, however, an important methodological issue because it points to a distinction between the large majority of supportive studies and the unsupportive minority. In one review, researchers found that after the use of controls, the proportion of analyses classified as wholly supportive of an income inequality effect on health was 83 percent among international studies, but it fell to 73 percent in large subnational areas, and to 45 percent in studies of small areas such as neighborhoods.[13] A similar pattern was

noted in an earlier meta-analysis of studies of income inequality and violent crime, including homicide[18] and a later meta-analysis of multilevel studies.[63] Hsieh and Pugh[18] concluded that "homogenous estimates of association between income inequality and homicide were reported by studies using states and nations as their sampling units but not by studies using smaller sampling units."

Table 1. Epidemiological framework for causal inference

Criteria	Importance
• **Consistency**	**The association has been replicated in different methodological, geographical, and time settings**
• **Temporality**	**The putative cause must precede the effect, this is an indisputable criterion for causality**
• Strength of association	The stronger an association is the less likely it is that there is some alternative unknown explanation
• Specificity	There is a high probability that an exposure is causally linked to some outcomes more than to others (many epidemiologists believe this criterion should be dropped)
• Dose response relationship	Increased exposure is related to increased outcomes
• Cessation of exposure	If exposure changes, positively or negatively, the incidence of the outcome will rise or fall
• **Consideration of alternative explanations**	**The association is not confounded by one or more other factors**
• **Biological plausibility**	**The association fits with existing biological knowledge**
• Coherence	The association is supported by other scientific knowledge

Source: Adapted from: Gordis I. Epidemiology. 5th ed. Philadelphia: Elsevier Saunders; 2013.
Note: Major criteria appear in bold font.

We have previously suggested that studies of income inequality are more supportive in large areas because in that context income inequality serves as a measure and determinant of the scale of social stratification or how hierarchical a society is.[13] Income inequality in small areas is affected by the degree of residential segregation of rich and poor, and the health of people in deprived neighborhoods is likely to be poor, not because of the inequality within each of those small areas, but because they are deprived in relation to the wider society. Studies from the United States and Sweden, which have compared the strength of association at different levels of aggregation, support this interpretation and the need to think carefully about scale before conducting studies.[64-66] Another

factor that might contribute to the same picture is the possibility that more unequal societies may give rise to greater residential segregation between rich and poor and thereby increase the inequality between areas and diminish the inequality within them.

Together, the studies provide overwhelming evidence that greater inequality is linked to worse health and more violence. Factors such as the size of area[64] and the use of conceptually inappropriate controls may provide plausible explanations of the minority of unsupportive studies.

Temporality

The large number of cross-sectional studies, undertaken over several decades, which link income inequality to health and violence, imply that there are relationships over time. As neither income distribution nor health are invariant over time, the fact that cross-sectional associations between them have been reported so many times is in itself an indication that they move together.

The preponderance of cross-sectional studies is partly a reflection of the limited availability of time series estimates of income inequality and, for some countries, for health outcomes. But there are now a growing number of studies of these relationships over time.

A meta-analysis of multi-level studies by Kondo and colleagues included data covering 59,509,857 individuals from nine cohort studies from Denmark, Finland, Norway, New Zealand, Sweden, and the United States, with followup periods ranging from 1-28 years.[67] The overall cohort relative risk (95 percent confidence interval) per 0.05 unit increase in the Gini coefficient, a measure of income inequality, was 1.08 (1.06 to 1.10). Studies with baseline data collection after 1990 and a length of followup greater than 7 years had a marginally higher relative risk; these interactions were not modified by the size of the area in which inequality was measured.[63]

In an international panel study of 21 developed countries over 30 years, controlling for serial correlation and stratifying by age and sex, Torre and Myrskylä, found that high inequality was associated with increased mortality of males and females aged 1–49 and older women but not older men.[68]

High inequality was associated with increased mortality of males and females aged 1–49 and older women but not older men.

Zheng[69] reviewed 79 studies of income inequality in relation to mortality: four aggregate and seven multilevel studies examined lagged effects up to 10 years with mixed results. However, all of these studies tested the lagged effect of income inequality in a particular year, treating it as a time-invariant variable and failing to control for a series of previous, subsequent, and contemporaneous income inequalities. Zheng also reviewed eight similar studies of self-rated health; seven of these found a significant effect, with two suggesting the strongest effect persisting through 15 years. Again, however, none of these studies looked at serial measures of inequality. Zheng's reviews included studies published through 2008 and contained in four previously published reviews.[11-13,67]

We have conducted a further primary systematic search for time-series and panel studies of income inequality and health and identified an additional nine studies, containing 53 analyses, including studies of mortality, life expectancy, infant mortality, under-5 survival rate, and self-rated health.[a] Of these, 55 percent support a longitudinal effect of income inequality on health (60 percent of within-country studies in the United States, the United Kingdom, and Norway), 37.5 percent do not, and 7.5 percent had mixed results, but all of these studies also suffer from the same methodological problem of not considering time-variant income inequality.

Studies also varied in the inclusion of control variables in the analyses of income inequality and health, including measures of aggregate or individual income or education; ethnic mix; unemployment; alcohol or tobacco consumption; birth, fertility, and divorce rates; benefit payments; health expenditures; and other variables. As some of these may be mediating or moderating factors in a causal pathway leading from income inequality to health, the inclusion of some is questionable, and the estimates of the effect of inequality would be underestimated.

Zheng[69] went on to conduct a discrete time-hazard analysis of U.S. national-level income inequality on mortality, controlling for individual income, in 701,179 individuals with a 21-year followup. A detrimental effect of rising inequality began to affect mortality after 3-5 years, and the effect size increased until mortality plateaued at a higher level after 12 years. This finding was robust to different model specifications and different measures of inequality. This study probably provides the best estimates of the average lag time between changes in inequality and mortality. It seems to accord well with other studies, though of course lag times will vary between age groups and causes of death.

A review of time-series and panel studies of income inequality and crime, that included seven studies examining homicide rates (five conducted in the United States or Canada or by international cross-country comparisons), found a significant increase in the murder rate with rising income inequality.[34] A study in West Germany found no association between income inequality and homicide rates,[70] as did a single international study,[71] although the statistical methods of this study were criticized.

Clarkwest analyzed State-level data within the United States from 1970 to 2000 to examine the effects of initial levels and change in income inequality on 10-year changes in life expectancy, finding that States with higher levels of inequality experienced less subsequent improvement in life expectancy.[72]

It is important to note that, for reasons which are not well understood, health continues to improve over time in most developed countries, with life expectancy rising by approximately 2-3 years with each decade. Against this background rate of improvement, the effect of changes in income inequality is to speed up or slow this background rate. Only when there are catastrophic rises in inequality, as in Russia and Eastern Europe during the transition from communism, does life expectancy actually fall.[52]

[a] See Appendix: Summary of Studies of Income Inequality and Health.
Available at http://www.york.ac.uk/healthsciences/our-staff/kate-pickett/#publications

A number of studies have suggested that death rates among the elderly show little or no relation to current inequality, see for example, Torre & Myrskyla.[68] However, analyses have shown that the health of older people is independently influenced by socioeconomic status at three different points in the life course, by fetal health and perhaps also by social security and welfare provisions in childhood.[73-75] If health in later life is similarly affected by lifetime exposure to inequality, we will need lifelong measures of exposure before we know whether the health of the elderly really is insensitive to their lifetime experience of inequality.

Strength of Association

In general, the stronger an association between putative cause and effect, the less likely it is that the relationship can be explained by other factors. But with health outcomes that are multifactorial, not all causes will have strong effects if they are necessary but not sufficient to cause the outcome alone. Nevertheless, at a population level, even moderate effects can have large impacts.

In international, cross-sectional, unadjusted studies of income inequality in relation to health and social problems in rich countries, the strength of the statistically significant associations vary.[15,28,37,76] Correlations with income inequality are higher for mental illness and teenage birth rates (both, $r=0.73$) and drug use and child well-being (both, $r=0.63$) than for life expectancy, infant mortality, obesity, and homicide (all, $r<0.5$). However, when researchers treated income inequality as a common cause of many health and social problems and combined them in one index, which tends to emphasize their common variance, the correlation with an index of problems was so high ($r=0.87$) that any alternative explanations would need to have extraordinarily strong effects (Figure 1).

When estimated in multilevel models, the size of the effect of inequality usually looks much smaller, as described earlier in this chapter in the section on Consistency. The difference is a matter of what is included as an effect of inequality. Some of the early multilevel studies of the effects of inequality were based on the assumption that the relationship between individual income and health was a reflection of the direct effects of what people's material circumstances did for their health regardless of anyone else in society. The desire was to separate out such effects before looking at the broader contextual effects of inequality, which were assumed to work through quite different pathways involving psychosocial processes hinging on relativities and social comparisons. However, a great deal of research attests to the likelihood that individual income is related to health because it is a marker of individual social status,[77,78] and that subjective social status may be more important than objective measures.[79] There is also evidence that greater inequality worsens outcomes such as math and literacy scores, social mobility, dropping out of high school, teenage birth rates, and mental illness, all of which might create feedback from higher inequality to increased numbers of people on low incomes.[37,39,80] If so, this would mean that multilevel models controlling out the effects of individual income risk seriously underestimate the effects of inequality.

Specificity

In some contexts, specificity is an outmoded causal criterion that dates from when the main health focus was on infectious diseases, which could only be caused by exposure to a specific pathogen. It is less relevant in a context where most health and social problems have multiple, interacting causes, and many outcomes share causes. However, there is an aspect of specificity in the relationship between income inequality and health that is helpful when considering causality and the pathways from one to the other. As we outlined in the History section of this chapter, the adverse effects of income inequality seem to be specific to outcomes that have an inverse social gradient.[26] For example, there was no social gradient for breast or prostate cancer mortality and no effect of income inequality, whereas there was a steep social gradient in working age all-cause mortality, and there was a strong association with income inequality. This would explain why the social outcomes included in Figure 1 are more common in unequal societies.

Broadly similar results have been found for child health.[81] In 29 Organization for Economic Co-operation and Development (OECD) countries, income inequality was positively related to post-neonatal mortality and teenage overweight, both of which have steep social gradients, but there was no association for suicide, which did not have clear evidence of a social gradient in some countries. One interpretation of this specificity is that income inequality intensifies the health effects of social hierarchy and social comparisons, thus increasing socioeconomic disparities in health. However, there was no association between income inequality and child asthma or adolescent smoking, both of which have some evidence of social gradients.

In international comparative studies, there is also a degree of specificity with regard to income inequality being associated with objective measures of health, rather than subjective measures. This is because, internationally, there is no correlation between life expectancy and the proportion of the population with good self-rated health, although these measures are correlated within countries.[82,83]

Dose-Response Relationship

A very large number of studies demonstrate statistically significant linear relationships between income inequality and health. The effects on inequality increase step by step from the most unequal of the 50 U.S. States and the most unequal countries to the most equal. However, Kondo and colleagues find a threshold effect, with higher relative risk of mortality in cohort studies with higher levels of income inequality (Gini coefficient >0.30) at baseline.[63] In other analyses of health and social problems among developed countries, the relationships tend to appear linear, with no evidence of a step change above some threshold level of inequality (see Figure 1). According to the World Bank World Development Indicators database, very few nations have Gini coefficients below 0.30,[84] the threshold identified by Kondo and colleagues; among developing nations there is only Afghanistan. Several former Soviet republics (Belarus, Bulgaria, Czech Republic, Kazakhstan, Romania, Slovak Republic, Ukraine) have Gini coefficients between 0.25-0.29, as do

the Scandinavian countries (Denmark, Finland, Norway, Sweden), Austria, Germany and Japan. Among OECD countries, only the Netherlands has not experienced a rise in income inequality since the mid-1980s. Thus, most of the world's population is exposed to income inequality above the threshold suggested by Kondo and colleagues, and the proportion of those exposed continues to rise.[85]

Even within more equal countries, inequality seems to matter. A recent study from Norway[86] found an independent effect of regional income inequality on mortality, after adjustment for regional-level social and economic characteristics. A study from Finland[87] suggested that widening differences in income inequality account for almost half of the increase in health inequalities, and one from Sweden found a detrimental effect of municipal income inequality on self-rated health.[66]

Whether each additional increment of inequality above a Gini of 0.30 has a greater effect than it does below that level remains unclear, but there is substantial agreement that there is a dose-response relationship above that level.

Cessation of Exposure

There can be no examples of cessation of exposure to inequality – only of exposure to more or less inequality. Interesting evidence comes from a study by Hamilton and Kawachi[88] that assessed whether or not individuals who migrate to the United States from countries with greater income inequality than the United States have better health than those who migrate from countries with less income inequality. Among immigrants who lived in the United States between 6 and 20 years, those for whom moving to the United States was a move towards greater equality had better self-reported health than those for whom it was a move towards greater inequality. Similarly, Auger and colleagues[89] found that income inequality was associated with mortality among non-immigrant Canadians but not migrants, although for long-term immigrants the effects tended to approach those of the Canadian-born population.

Also relevant is the striking reversal in international rankings in income inequality and population health between the United States and Japan in the three or four decades following the Second World War.[90] In the post-war period, the United States had much lower inequality than it does today and ranked high in the international league table for life expectancy, whereas Japan was highly unequal, with lower life expectancy. But by the end of the 1980s, Japan had become one of the most equal countries and had the highest life expectancy in the world. In contrast, the United States became rapidly more unequal from the late 1960s and is now among the most unequal societies in the developed world. During that period, the U.S. position slipped in the international life expectancy league tables, and it now ranks 40th according to the United Nations.[91]

Consideration of Alternative Explanations

Given that the epidemiological criteria examined so far support a causal interpretation for the role of inequality, we should ask whether there are any other possible explanations.

A paper by Deaton in 2003[92] reported that the proportion of black residents in States and Metropolitan Statistical Areas of the United States explained the income inequality-health association. This paper continues to be cited as evidence that income inequality does not affect health, despite the fact that several more recent studies found that ethnic heterogeneity does not confound the income inequality-health association in the United States.[12,93-96] International comparisons also show that income inequality is significantly related to health even after adjustment for ethnic heterogeneity.[97] Nor does ethnicity explain the income inequality-homicide relationship in U.S. States. To clarify this, one analysis of homicides in the 50 States confined attention to white perpetrators of homicides and showed they were significantly related to income inequality measured only among the white populations of each State.[98] It seems likely that ethnic differences attract more attention and seem more important not only when they become markers of social status differences but also when greater inequality makes social status differentiation more powerful, increasing the importance of "downward" social prejudices whether by class or ethnicity.[99,100]

Another proposed alternative explanation suggests not only that the relationship between individual income and health is curvilinear, such that a rise in income for the poor has a greater impact on health than an equivalent rise in income for the rich, but also that this effect reflects only the direct influence of material living standards on health – not inequality as such. The suggestion is that greater equality would improve average health but only for reasons related to what individual material circumstances do to health, regardless of other income and position in the income hierarchy. The assumption is that someone's health is affected only by their own income and is unaffected by where they are in the income hierarchy. Studies within the United States[101] and the United Kingdom,[102] as well as international comparisons,[47] disprove this explanation, as do the many multilevel studies of income inequality and health reviewed by Kondo et al, which control for individual income and socioeconomic status.[67] Such studies show a contextual effect of inequality over and above the effects of individual income.

Income inequality is, and can only be, an ecological variable describing the scale of income differences across a population. Because inferences are usually made only to other ecological variables, such as rates of health or social problems across the same population, the possibility of an ecological fallacy does not arise: inferences are not made from ecological variables to individual risk. However, studies that ask "Whose health is affected by inequality" suggest that although effects are probably strongest among the least well off, they extend to the majority of the population.[37,103]

It also has been suggested that the income inequality-health association reflects reverse causality – in other words, income inequality is a result of a larger proportion of the population being unhealthy, rather than a cause. The time series studies described above, which show that there are substantial lag periods between changes in inequality and changes in health, disprove this interpretation, as do the findings of cohort studies. In addition, income inequality has been related to many infant and child outcomes, including infant mortality, low birth weight, child well-being, and child mental health problems, which would not be expected to affect inequality.[27]

The suggestion of reverse causality faces two other difficulties. The first is that more unequal countries appear to do poorly on a wide range of health and social outcomes, while more equal

countries do well. If income inequality were a result of worsening outcomes, then it would be necessary to find an alternative explanation for why so many disparate problems – ranging from health to homicides, child well-being, mental illness, and drug abuse – all tend to be worse in some countries than in others. As they are such different outcomes, yet all with similar social gradients, it would be necessary to posit another, very deep-seated explanation closely related to social status differentiation. Lastly, a good deal is known of the economic policies that came in from the late 1970s and led directly to wider income differences.

In terms of pathways from income inequality to health, it has been suggested that more generous welfare regimes, public spending (e.g., on health or transport), more comprehensive social security, and increased investment in human capital development (e.g., education) are all characteristic of more equal societies, and that the relation between income inequality and health may therefore be mediated by these "neo-material" factors. From a neo-material perspective, the association between income inequality and health reflects people's lack of resources, as well as societal underinvestment in such things as "education, health services, transportation, environmental controls, availability of food, quality of housing, [and] occupational health regulations."[104]

Nevertheless, explicit tests of "neo-material" vs. psychosocial pathways from income inequality to healthy life expectancy, mortality, mental health, and homicide rates conclude that psychosocial factors, such as social capital and trust, mediate the relationship, whereas neo-material factors, such as public expenditure on health or social services, have little or no explanatory role.[105-107] Of course, insofar as welfare regimes, social security, and other programs redistribute income, it is difficult to disentangle the independent effects of income inequality and welfare regimes.

Lastly, the tendency is often to imagine that cultural differences lie behind and are the real reason for associations between income inequality and a poorer performance on a wide range of health and social outcomes. Any such hypotheses about the role of culture would of course have to be compatible with the evidence that shows that health changes follow changes in income distribution after a lag of some years. It would also have to be compatible with the evidence of associations in different parts of the world, at different levels of development and with different cultures. So, for instance, Mexico, Russia, and South Africa all have very high levels of income inequality and very high levels of violence, but their cultural identities are very different. Similarly, societies like Japan and the Scandinavian countries have low levels of inequality and low levels of violence despite obvious cultural differences among them. Also interesting is the cultural similarity between Portugal and Spain. Both countries were dictatorships until the mid-1970s, and they share a border. However, their performance on the Index of Health and Social Problems reflects (as shown in Figure 1) their substantial differences in income distribution.

Biological Plausibility

A psychosocial explanation of the effect of income inequality on health and behavioral outcomes is consistent with the biology of chronic stress, new studies of the neuroscience of social sensitivity,

and concepts from evolutionary biology. Income inequality is linked to lower levels of social cohesion and generalized trust, suggesting that inequality acts as a social stressor.[108-110]

Chronic stress impairs memory and increases risk of depression, lowers immune responses, elevates blood pressure and risk of cardiovascular disease, and affects hormonal systems.[111] Research shows that the ways in which we relate to one another, such as friendship, social support, and social networks, are as protective for health as smoking is deleterious.[112] If we have friends, we are less likely to contract a common cold infection in randomized controlled trials.[113] Likewise, if we have a difficult relationship with our spouses or partners, we heal more slowly in trials of experimental wound healing.[114] A meta-analysis of 208 laboratory studies of acute psychological stressors and cortisol responses shows that stronger cortisol responses were elicited if tasks were uncontrollable or characterized by "social-evaluative threat" (threats to self-esteem or social status).[115] Even low levels of psychological distress were found to be related to mortality in a meta-analysis of 10 large prospective cohort studies.[116] Telomere length, a measure of cell aging, was found to be shorter by age 9 among African American boys who lived in highly disadvantaged environments compared to those who were raised in more affluent environments.[117]

Neuroscience studies also highlight the importance of psychosocial factors for human physiology. A neuroimaging study showed that social pain (exclusion) activated the brain in the same ways as physical pain. The anterior cingulate cortex (ACC) was more active during experiences of social exclusion and was positively correlated with self-reported distress.[118] In another study, baseline sensitivity to physical pain predicted sensitivity to social rejection, and social exclusion was associated with more sensitivity to physical pain.[119] In two experiments, participants received either acetaminophen (a pain suppressant) or a placebo for 3 weeks. Acetaminophen reduced daily reports of social pain, and functional magnetic resonance imaging showed that acetaminophen reduced neural responses to social pain in areas of the brain previously shown to be related to both social and physical pain.[120]

Evolutionary explanations of human sensitivity to social relationships and hierarchies stress the importance of belonging and people's need for positive relationships and connectedness. Social exclusion affects cognitive, emotional, and behavioral outcomes, and adaptations to low social rank in both animals and humans include altered levels of hormones and behaviors, such as withdrawal, apathy, or hypervigilance.[121] A theory linking submission and subordination to depression suggests that it results from an inability to stop, or escape from, a submissive defeat strategy, and the evidence reviewed by Johnson and colleagues supports this; in more than 20 research studies, people with depression were more likely to report feeling inferior or experiencing shame.[122]

Coherence

In rich countries, there is no association between average levels of income (e.g. gross national income per capita) and measures of health, such as life expectancy.[123-126] Yet within rich countries, there are strong associations between individual income and life expectancy. This pattern suggests

that it is relative income within societies that is important for health in rich countries, in turn suggesting that psychosocial mechanisms are relevant.

Recent studies of income inequality in relation to psychological states and traits and sociological outcomes lend coherence to a psychosocial explanation of the health and social effects of income inequality on health. International comparisons show that status anxiety is higher in more unequal countries, for all socioeconomic groups.[127] Status anxiety and trust were found to mediate the association between income inequality and subjective well-being.[128] In more unequal countries, people exhibit higher levels of self-enhancement, i.e., believing themselves to be better than average.[129] In both ecological and multi-level analyses, people in more unequal U.S. States scored lower on a measure of agreeableness, reflecting less concern for social harmony and getting along with others.[130] In more unequal European countries, people show less solidarity; they are less willing to help others.[131]

Discussion and Conclusions

The body of evidence on income inequality and health points strongly to a causal connection. The major criteria of temporality, biological plausibility, consistency, and lack of alternative explanations are well supported. Of the small minority of studies that find no association, most can be explained by income inequality being measured at an inappropriate scale, the inclusion of mediating variables as controls, the use of subjective rather than objective measures of health, or followup periods that are too short.

Suicides seem to stand as an important exception to the general pattern: they tend to be more common in more equal societies, despite the evidence that depression is more common in more unequal societies.[39,132] A possible explanation is that social gradients in suicides are not always consistent internationally.[133] Another possibility is that there may be some truth in the view that violence can be directed either outwards or inwards against oneself. If suicide is, like homicide, often a response to adversity, we think it likely that greater equality increases a tendency to blame oneself rather than others for what goes wrong.

Epidemiological causal criteria are not exhaustive. A good test of the validity of a scientific theory is its ability to make successful, testable predictions. The theory that more equal societies are healthier arose from one international study[17] and has now been tested in many different contexts. The search for a mechanism led to the discovery that social relationships (social cohesion, trust, involvement in community life, and low levels of violence) are better in more equal societies, suggesting that inequality and health are linked through psychosocial processes related to social differentiation and relative deprivation.[61] That inequality does have powerful psychosocial effects is now amply confirmed.

We suggest that the most parsimonious explanation for the effects of income inequality is that larger income differences increase social distances, accentuating social class or status differences. This would explain why income inequality is most closely related to health when measured across whole

societies coterminous with social class hierarchies.[13,134] Rather than income inequality being a new and independent determinant of health, it is likely to act by strengthening the many causal processes (known and unknown) through which social class imprints itself on people throughout life. This would suggest why, not only health, but a wide range of other outcomes with social gradients are also related to inequality. It also suggests that if class and status are to become a less powerful influence both on individual lives and on whole societies, it will be necessary to reduce the material differences that so often constitute the cultural markers of social differentiation.

As whole populations are exposed to societal income inequality, estimates of the population attributable risk will be high even if, for some outcomes, the causal effect on some outcomes is modest. Kondo and colleagues[67] estimated that upwards of 1.5 million deaths (9.6 percent of total adult mortality for the 15-60 age group) could be averted in 30 OECD countries if each country reduced its Gini coefficient below 0.30. If individual income is also related to health partly through psychosocial mechanisms involving relative deprivation, then multilevel models that control out its effects may substantially underestimate the effects of inequality.[135] It has been estimated that if the United Kingdom reduced its inequality to the average in other OECD countries, the expenditure savings on physical and mental illness, violence, and imprisonment alone would amount to £39 billion per year.[136]

Future research should move beyond mere replication of these findings in different samples towards more explicit attempts to clarify the causal relationships, including studies of (1) different measures of income inequality (top- and bottom-sensitive measures, for example) in relation to different health and social outcomes, (2) time lags for different outcomes, (3) further modeling and testing of specific causal pathways, and (4) whether inequalities in wealth are as much a part of the picture as inequalities in income. Comparable measures of wealth inequality are available for only a limited number of countries, but initial explorations of the relationships with life expectancy are interesting. Life expectancy in Denmark, which seems to be an outlier in relation to its more equal distribution of income, appears to fall into place in relation to its large inequalities in wealth.[137]

The evidence that large income differences have damaging health and social consequences is already far stronger than the evidence supporting policy initiatives in many other areas of social and economic policy, and the message is beginning to reach politicians. The world leaders we mentioned at the start of this chapter have all referred to inequality as a cause of social and economic harm. But to recognize the problem is not the same as tackling it effectively. The gap between the richest and poorest 20 percent of households in countries like the United States and United Kingdom is not only very much wider than it used to be in the 1970s, but it is still twice as large as in some other successful market democracies. The reason why politicians do not do more is almost certainly a reflection of the undemocratic power of money in politics and the media.[138] Narrowing the gap will require not only redistributive tax policies but also a reduction in income differences before tax. The halving of top tax rates since the 1970s has led not only to a widening of income differences after tax but, more surprisingly, to an acceleration in pre-tax income differences particularly in the private sector where pay for top executives seems unrelated to company performance.[139-141]

Schrecker[142] has written about the risks of policymakers requiring unachievable standards of proof in social epidemiology before they are willing to act, and Popper[30] emphasized that scientific theories are never finally proven true. Adopting too high a standard of evidence may mean that it is never considered strong enough. Schrecker quotes Michael Marmot[143] as saying "While we should not formulate policies in the absence of evidence to support them, we must not be paralyzed into inaction while we wait for the evidence to be absolutely unimpeachable."

Acknowledgments

We are grateful to Hector Rufrancos for collating the time-series literature. This paper was published previously by Social Science & Medicine (2014 Dec 30; epub); the abstract is available at http://www.ncbi.nlm.nih.gov/pubmed/25577953. The opinions presented herein are those of the authors and may not necessarily represent the official position of the Agency for Healthcare Research and Quality, the National Institutes of Health, or the U.S. Department of Health and Human Services.

Authors' Affiliations

Kate E. Pickett, Department of Health Sciences, University of York, York, UK; Richard G. Wilkinson, Division of Epidemiology and Public Health, University of Nottingham, Nottingham, UK.

Address correspondence to: Kate E. Pickett, Department of Health Sciences, Seebohm Rowntree Building, Area 3, University of York, Heslington, York, YO10 5DD, UK;
email kate.pickett@york.ac.uk

References

1. Obama B. State of the Union address, 2014. Available at http://www.whitehouse.gov/the-press-office/2014/01/28/president-barack-obamas-state-union-address; Accessed December 5, 2014.
2. Cameron D. Hugo Young lecture, November 20, 2009. Available at http://www.theguardian.com/commentisfree/2009/nov/10/big-society-government-poverty-inequality. Accessed December 5, 2014.
3. Pope Francis. Evangelii Gaudium. Vatican City: Vatican Press; 2013.
4. Lagarde C. Speech at World Economic Forum, January 23, 2013. Davos, Switzerland. Available at https://www.imf.org/external/np/speeches/2013/012313 htm. Accessed December 7, 2014.
5. Moon B-K. Remarks at informal General Assembly Thematic Debate on Inequality, July 8, 2013. Available at http://www.un.org/apps/news/infocus/sgspeeches/statments_full.asp?statID=1918#.UfDb943EPO4. Accessed January 27, 2015.
6. Elliott L. Climate change will 'lead to battles for food', says head of World Bank. The Guardian; April 3. 2014. Available at http://www.theguardian.com/environment/2014/apr/03/climate-change-battle-food-head-world-bank. Accessed January 27, 2015.
7. Global risks 2013. Geneva: World Economic Forum; 2014. Available at http://www.weforum.org/reports/global-risks-2013-eighth-edition. Accessed January 27, 2015.
8. Ortiz I, Cummins M. Global inequality: beyond the bottom billion – a rapid review of income distribution in 141 countries. New York, NY: UNICEF; 2011.
9. Divided we stand: why inequality keeps rising. Available at http://tinyurl.com/pz28olv. Accessed January 27, 2015. Paris: OECD Publishing; 2011.
10. Wagstaff A, van Doorslaer E. Income inequality and health: what does the literature tell us? Annu Rev Public Health 2000;21:543-67.
11. Lynch J, Smith GD, Harper S, et al. Is income inequality a determinant of population health? Part 1. A systematic review. Milbank Q 2004;82(1):5-99.
12. Subramanian SV, Kawachi I. Income inequality and health: what have we learned so far? Epidemiol Rev 2004;26:78-91.
13. Wilkinson RG, Pickett KE. Income inequality and population health: a review and explanation of the evidence. Soc Sci Med 2006;62(7):1768-84.
14. Macinko JA, Shi L, Starfield B, et al. Income inequality and health: a critical review of the literature. Med Care Res Rev 2003;60(4):407-52.
15. Wilkinson RG, Pickett KE. Income inequality and social dysfunction. Annu Rev Sociol 2009;35:493-512.
16. Loftin C, Hill RH. Regional subculture and homicide: an examination of the Gastil-Hackney thesis. Am Sociol Rev 1974;39:714-24.
17. Rodgers GB. Income and inequality as determinants of mortality: an international cross-section analysis. Popul Stud 1979;33:343-51.
18. Hsieh C-C, Pugh MD. Poverty, income inequality, and violent crime: a meta-analysis of recent aggregate data studies. Criminal Justice Rev 1993;18:182-202.
19. Wilkinson RG. National mortality rates: the impact of inequality? Am J Public Health 1992;82(8):1082-4.
20. Wilkinson RG. Income distribution and life expectancy. Br Med J 1992;304(6820):165-8.
21. Wilkinson RG, Kawachi I, Kennedy B. Mortality, the social environment, crime and violence. Soc Health Illness 1998;20(5):578-97.
22. Wilson M, Daly M. Life expectancy, economic inequality, homicide, and reproductive timing in Chicago neighbourhoods. Br Med J 1997;314(7089):1271-4.
23. Pickett KE, Kelly S, Brunner E, et al. Wider income gaps, wider waistbands? An ecological study of obesity and income inequality. J Epidemiol Community Health 2005;59(8):670-4.
24. Pickett KE, Mookherjee J, Wilkinson RG. Adolescent birth rates, total homicides, and income inequality in rich countries. Am J Public Health. 2005;95(7):1181-3.
25. Pickett KE, James OW, Wilkinson RG. Income inequality and the prevalence of mental illness: a preliminary international analysis. J Epidemiol Community Health 2006;60(7):646-7.
26. Wilkinson RG, Pickett KE. Income inequality and socioeconomic gradients in mortality. Am J Public Health 2008;98(4):699-704.
27. Pickett KE, Wilkinson RG. Child wellbeing and income inequality in rich societies: ecological cross sectional study. Br Med J 2007;335:1080.
28. Wilkinson RG, Pickett KE. The problems of relative deprivation: why some societies do better than others. Soc Sci Med 2007;65(9):1965-78.
29. Popper K. The logic of scientific discovery. Abingdon: Routledge; 2002.
30. Popper K. Conjectures and refutations: the growth of scientific knowledge. Abingdon: Routledge; 2014.
31. Kaplan GA, Pamuk ER, Lynch JW, et al. Inequality in income and mortality in the United States: analysis of mortality and potential pathways. Br Med J 1996;312(7037):999-1003.

32. Kennedy BP, Kawachi I, Prothrow-Stith D. Income distribution and mortality: cross sectional ecological study of the Robin Hood index in the United States. Br Med J 1996;312(7037):1004-7.
33. Pei X, Rodriguez E. Provincial income inequality and self-reported health status in China during 1991-7. J Epidemiol Community Health 2006;60:1065-9.
34. Rufrancos H, Power M, Pickett KE, et al. Income inequality and crime: a review and explanation of the time–series evidence. Sociol Criminol 2013;1:103.
35. Wilkinson RG. Unhealthy societies: the afflictions of inequality. London: Routledge; 1996.
36. Kawachi I, Kennedy BP, Lochner K, et al. Social capital, income inequality, and mortality. Am J Public Health 1997;87(9):1491-8.
37. Wilkinson RG, Pickett K. The spirit level: why equality is better for everyone. London: Penguin; 2010.
38. Pickett KE, Wilkinson RG. Inequality: an underacknowledged source of mental illness and distress. Br J Psychiatry 2010;197:426-8.
39. Messias E, Eaton WW, Grooms AN. Economic grand rounds: income inequality and depression prevalence across the United States: an ecological study. Psychiatr Serv 2011;62(7):710-2.
40. Burns JK, Tomita A, Kapadia AS. Income inequality and schizophrenia: increased schizophrenia incidence in countries with high levels of income inequality. Int J Soc Psychiatry 2014;60(2):185-96.
41. Johnson SL, Wibbels E, Wilkinson R. Economic inequality is related to cross national prevalence of psychotic symptoms. Forthcoming.
42. Smoking and health: report of the Advisory Committee to the Surgeon General of the United States. Washington, DC: Department of Health, Education, and Welfare, 1964.
43. Hill AB. The environment and disease: association or causation? Proc Royal Soc Med 1965;58(5):295-300.
44. Bhopal RS. Concepts of epidemiology: an integrated introduction to the ideas, theories, principles, and methods of epidemiology. Oxford: Oxford University Press; 2002.
45. Rothman KJ, Greenland S, Lash TL. Modern epidemiology. Baltimore, MD: Lippincott Williams & Wilkins; 2008.
46. Gordis L. Epidemiology. 5th ed. Philadelphia: Elsevier Saunders; 2013.
47. Babones SJ. Income inequality and population health: correlation and causality. Soc Sci Med 2008;66(7):1614-26.
48. Flegg AT. Inequality of income, illiteracy and medical-care as determinants of infant-mortality in underdeveloped-countries. Popul Stud 1982;36(3):441-58.
49. Hales S, Howden-Chapman P, Salmond C, et al. National infant mortality rates in relation to gross national product and distribution of income. Lancet 1999;354(9195):2047.
50. Drain PK, Smith JS, Hughes JP, et al. Correlates of national HIV seroprevalence: an ecologic analysis of 122 developing countries. J Acquir Immune Defic Syndr 2004;35(4):407-20.
51. Fajnzylber P, Lederman D, Loayza N. Inequality and violent crime. J Law Econ 2002;45:1-40.
52. Marmot M, Bobak M. International comparators and poverty and health in Europe. Br Med J 2000;321(7269):1124-8.
53. Biggs B, King L, Basu S, et al. Is wealthier always healthier? The impact of national income level, inequality, and poverty on public health in Latin America. Soc Sci Med 2013;71(2):266-73.
54. De Maio FG, Linetzky B, Ferrante D, et al. Extending the income inequality hypothesis: ecological results from the 2005 and 2009 Argentine National Risk Factor Surveys. Glob Public Health 2012;7(6):635-47.
55. Daly M, Wilson M, Vasdev S. Income inequality and homicide rates in Canada and the United States. Can J Public Health 2001;43(2):219-36.
56. Rasella D, Aquino R, Barreto ML. Impact of income inequality on life expectancy in a highly unequal developing country: the case of Brazil. J Epidemiol Community Health 2013;67(8):661-6.
57. Subramanian SV, Delgado I, Jadue L, et al. Income inequality and health: multilevel analysis of Chilean communities. J Epidemiol Community Health 2003;57(11):844-8.
58. Larrea C, Kawachi I. Does economic inequality affect child malnutrition? The case of Ecuador. Soc Sci Med 2005;60(1):165-78.
59. Rajan K, Kennedy J, King L. Is wealthier always healthier in poor countries? The health implications of income, inequality, poverty, and literacy in India. Soc Sci Med 2013;88:98-107.
60. De Vogli R, Mistry R, Gnesotto R, et al. Has the relation between income inequality and life expectancy disappeared? Evidence from Italy and top industrialised countries. J Epidemiol Community Health 2005;59(2):158-62.
61. Kondo N, Kawachi I, Subramanian SV, et al. Do social comparisons explain the association between income inequality and health?: Relative deprivation and perceived health among male and female Japanese individuals. Soc Sci Med 2008;67(6):982-7.
62. Walberg P, McKee M, Shkolnikov V, Chenet L, Leon DA. Economic change, crime, and mortality crisis in Russia: regional analysis. Br Med J 1998;317(7154):312-8.

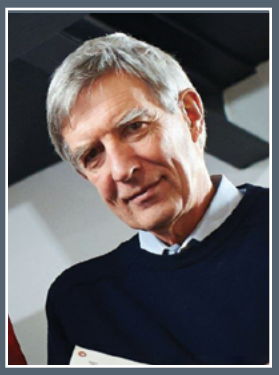
63. Kondo N, van Dam RM, Sembajwe G, et al. Income inequality and health: the role of population size, inequality threshold, period effects and lag effects. J Epidemiol Community Health 2012;66(6):e11.
64. Chen Z, Gotway Crawford CA. The role of geographic scale in testing the income inequality hypothesis as an explanation of health disparities. Soc Sci Med. 2012;75(6):1022-31.
65. Franzini L, Ribble J, Spears W. The effects of income inequality and income level on mortality vary by population size in Texas counties. J Health Soc Behav 2001;42(4):373-87.
66. Rostila M, Kolegard ML, Fritzell J. Income inequality and self-rated health in Stockholm, Sweden: a test of the 'income inequality hypothesis' on two levels of aggregation. Soc Sci Med 2012;74(7):1091-8.
67. Kondo N, Sembajwe G, Kawachi I, et al. Income inequality, mortality, and self rated health: meta-analysis of multilevel studies. Br Med J 2009;339:b4471.
68. Torre R, Myrskylä M. Income inequality and population health: an analysis of panel data for 21 developed countries, 1975-2006. Popul Stud 2014;68(1):1-13.
69. Zheng H. Do people die from income inequality of a decade ago? Soc Sci Med 2012;75(1):36-45.
70. Entorf H, Spengler H. Socioeconomic and demographic factors of crime in German: evidence frompanel data of the German states. Int Rev Law Econ 2000;20(1):75-106.
71. Messner SF, Raffalovich LE, Chrock P. Reassessing the cross-national relationship between income inequality and homicide rates: implications of data quality control in the measurement of income distribution. J Quant Criminol 2002;18(4):377-95.
72. Clarkwest A. Neo-materialist theory and the temporal relationship between income inequality and longevity change. Soc Sci Med 2008;66(9):1871-81.
73. Smith GD, Hart C, Blane D, et al. Lifetime socioeconomic position and mortality: prospective observational study. Br Med J 1997;314(7080):547-52.
74. Hoynes HW, Schanzenbach DW, Almond D. Long run impacts of childhood access to the safety net. Working paper 18535. Cambridge, MA: National Bureau of Economic Research; 2012. Available at http://www.nber.org/papers/w18535. Accessed January 27, 2015.
75. Johnson RC, Schoeni RF. Early-life origins of adult disease: national longitudinal population-based study of the United States. Am J Public Health 2011;101(12):2317-24.
76. Pickett KE, Wilkinson RG. Child wellbeing and income inequality in rich societies: ecological cross sectional study. Br Med J 2007;335(7629):1080.
77. Adler NE, Epel ES, Castellazzo G, et al. Relationship of subjective and objective social status with psychological and physiological functioning: preliminary data in healthy white women. Health Psychol 2000;19(6):586-92.
78. Marmot M. Status syndrome: how your social standing directly affects your health and life expectancy. London: Bloomsbury; 2004.
79. Singh-Manoux A, Martikainen P, Ferrie J, et al. What does self-rated health measure? Results from the British Whitehall II and French Gazel cohort studies. J Epidemiol Community Health 2006;60(4):364-72.
80. Doing better for children. Paris: OECD Publishing; 2009.
81. Bird PK. Social gradients in child health and development in relation to income inequality. Who benefits from greater income equality? York: University of York; 2014.
82. Barford A, Dorling D, Pickett K. Re-evaluating self-evaluation. A commentary on Jen, Jones, and Johnston (68:4, 2009). Soc Sci Med 2010;70(4):496-7; discussion 8-500.
83. Dorling D, Barford A. The inequality hypothesis: thesis, antithesis, and a synthesis? Health Place 2009;15(4):1166-9; discussion 3-5.
84. World development indicators: distribution of income or consumption. Washington, DC: World Bank; 2014. Available at http://wdi.worldbank.org/table/2.9. Accessed January 27, 2015.
85. Smeeding T, Morelli S, Thompson J. Recent trends in income inequality in the developed countries. Progressive Economy 2014;March(2):24-33.
86. Elstad J. Does the socioeconomic context explain both mortality and income inequality? Prospective register-based study of Norwegian regions. Int J Equity Health 2011;10(1):1-11.
87. Aittomaki A, Martikainen P, Rahkonen O, et al. Household income and health problems during a period of labour-market change and widening income inequalities - a study among the Finnish population between 1987 and 2007. Soc Sci Med 2014;100:84-92.
88. Hamilton TG, Kawachi I. Changes in income inequality and the health of immigrants. Soc Sci Med 2013;80:57-66.
89. Auger N, Hamel D, Martinez J, et al. Mitigating effect of immigration on the relation between income inequality and mortality: a prospective study of 2 million Canadians. J Epidemiol Community Health 2012;66(6):e5.
90. Bezruchka S, Namekata T, Sistrom MG. Improving economic equality and health: the case of postwar Japan. Am J Public Health 2008;98:216-21.
91. United Nations Department of Economic and Social Affairs (UN DESA) Population Division. World population prospects, the 2010 revision. New York, NY: United Nations, 2011.

92. Deaton A, Lubotsky D. Mortality, inequality and race in American cities and states. Soc Sci Med 2003;56(6):1139-53.
93. Ash M, Robinson DE. Inequality, race, and mortality in U.S. cities: a political and econometric review of Deaton and Lubotsky (56:6, 1139-1153, 2003). Soc Sci Med 2009;68(11):1909-13.
94. Ram R. Income inequality, poverty, and population health: evidence from recent data for the United States. Soc Sci Med 2005;61(12):2568-76.
95. Subramanian SV, Kawachi I. The association between state income inequality and worse health is not confounded by race. Int J Epidemiol 2003;32(6):1022-8.
96. Subramanian S, Kawachi I. Response: in defence of the income inequality hypothesis. Int J Epidemiol 2003;32(6):1037-40.
97. Ram R. Further examination of the cross-country association between income inequality and population health. Soc Sci Med 2006;62(3):779-91.
98. Daly M, Wilson M. Cultural inertia, economic incentives, and the persistence of "southern violence". In: Schaller M, Norenzayan A, Heine S, et al (Eds): Evolution, culture, and the human mind. New York, NY: Psychology Press; 2010.
99. Kennedy BP, Kawachi I, Lochner K, et al. (Dis)respect and black mortality. Ethn Dis 1997;7(3):207-14.
100. Andersen R, Curtis J. The polarizing effect of economic inequality on class identification: evidence from 44 countries. Res Soc Stratif Mobil 2012;30(1):129-41.
101. Wolfson M, Kaplan G, Lynch J, et al. Relation between income inequality and mortality: empirical demonstration. Br Med J 1999;319(7215):953-5.
102. Wood AM, Boyce CJ, Moore SC, et al. An evolutionary based social rank explanation of why low income predicts mental distress: a 17 year cohort study of 30,000 people. J Affect Disord 2012;136(3):882-8.
103. Subramanian SV, Kawachi I. Whose health is affected by income inequality? A multilevel interaction analysis of contemporaneous and lagged effects of state income inequality on individual self-rated health in the United States. Health Place 2006;12(2):141-56.
104. Lynch JW, Smith GD, Kaplan GA, et al. Income inequality and mortality: importance to health of individual income, psychosocial environment, or material conditions. Br Med J 2000;320(7243):1200-4.
105. Elgar FJ. Income inequality, trust, and population health in 33 countries. Am J Public Health 2010;100(11):2311-5.
106. Elgar FJ, Aitken N. Income inequality, trust and homicide in 33 countries. Eur J Public Health 2011;21(2):241-6.
107. Layte R. The association between income inequality and mental health: testing status anxiety, social capital, and neo-materialist explanations. Eur Soc Rev 2012;28(4):498-511.
108. Delhey J, Dragolov G. Why inequality makes Europeans less happy: the role of distrust, status anxiety, and perceived conflict. Eur Sociol Rev 2014;30(2):151-65.
109. Lancee B, Van de Werfhorst HG. Income inequality and participation: a comparison of 24 European countries. Soc Sci Res 2012;41:1166-78.
110. Uslaner EM, Brown M. Inequality, trust, and civic engagement. Am Politics Res 2005;33(6):868-94.
111. Sapolsky R. Sick of poverty. Sci Am 2005;293(6):92-9.
112. Holt-Lunstad J, Smith TB, Layton JB. Social relationships and mortality risk: a meta-analytic review. PLoS Med 2010;7(7):e1000316.
113. Cohen S, Doyle WJ, Skoner DP, et al. Social ties and susceptibility to the common cold. JAMA 1997;277:1940-44.
114. Kiecolt-Glaser JK, Loving TJ, Stowell JR, et al. Hostile marital interactions, proinflammatory cytokine production, and wound healing. Arch Gen Psychiatry 2005;62(12):1377-84.
115. Dickerson SS, Kemeny ME. Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. Psychol Bull 2004;130(3):355-91.
116. Russ TC, Stamatakis E, Hamer M, et al. Association between psychological distress and mortality: individual participant pooled analysis of 10 prospective cohort studies. Br Med J 2012;345.
117. Mitchell C, Hobcraft J, McLanahan SS, et al. Social disadvantage, genetic sensitivity, and children's telomere length. Proc Natl Acad Sci U S A 2014;111(16):5944-9.
118. Eisenberger NI, Lieberman MD, Williams KD. Does rejection hurt? An FMRI study of social exclusion. Science 2003;302(5643):290-2.
119. Eisenberger NI, Jarcho JM, Lieberman MD, et al. An experimental study of shared sensitivity to physical pain and social rejection. Pain 2006;126(1-3):132-8.
120. DeWall CN, MacDonald G, Webster GD, et al. Acetaminophen reduces social pain behavioral and neural evidence. Psychol Sci 2010;21(7):931-7.
121. DeWall CN, Deckman T, Pond RS, Jr., et al. Belongingness as a core personality trait: how social exclusion influences social functioning and personality expression. J Pers 2011;79(6):1281-314.
122. Johnson SL, Leedom LJ, Muhtadie L. The dominance behavioral system and psychopathology: evidence from self-report, observational, and biological studies. Psychol Bull 2012;138(4):692-743.
123. Preston SH. The changing relation between mortality and level of economic development. Popul

Stud 1975;29(2). Reprinted in: Int J Epidemiol 2007;36(3):484-90.
124. Cutler D, Deaton A, Lleras-Muney A. The determinants of mortality. J Econ Perspect 2006;20(3):97-120.
125. Kenny C. Getting better: why global development is succeeding - and how we can improve the world even more. New York: Basic Books; 2011.
126. Wilkinson R, Pickett K. The spirit level: why more equal societies almost always do better. London: Penguin; 2010.
127. Layte R, Whelan C. GINI DP 78: Who feels inferior? A test of the status anxiety hypothesis of social inequalities in health. Amsterdam: Amsterdam Institute for Advanced Labour Studies; 2013.
128. Delhey J, Dragolov G. Why inequality makes Europeans less happy: the role of distrust, status anxiety, and perceived conflict. Eur Sociol Rev 2014;30(2):151-65.
129. Loughnan S, Kuppens P, Allik J, et al. Economic inequality is linked to biased self-perception. Psychol Sci 2011;22(10):1254-8.
130. de Vries R, Gosling S, Potter J. Income inequality and personality: are less equal U.S. states less agreeable? Soc Sci Med 2011;72(12):1978-85.
131. Paskov M, Dewilde C. Income inequality and solidarity in Europe. Res Soc Strat Mob 2012;30(4):415-32.
132. Daly MC, Oswald AJ, Wilson D, et al. Dark contrasts: the paradox of high rates of suicide in happy places. J Econ Behav Org 2011;80(3):435-42.
133. Burrows S, Laflamme L. Socioeconomic disparities and attempted suicide: state of knowledge and implications for research and prevention. Int J Inj Contr Saf Promot 2010;17(1):23-40.
134. Kondo N, van Dam RM, Sembajwe G, et al. Income inequality and health: the role of population size, inequality threshold, period effects and lag effects. J Epidemiol Community Health 2011;66(6):e11.
135. Pickett KE, Wilkinson RG. Greater equality and better health. Br Med J 2009;339:b4320.
136. The cost of inequality. London: The Equality Trust; 2014.
137. Nowatzki NR. Wealth inequality and health: a political economy perspective. Int J Health Serv 2012;42(3):403-24.
138. Gilens M, Page BI. Testing theories of American politics: elites, interest groups, and average citizens. Perspect on Politics 2014;12(3):564-81.
139. Chemi EG, Giorgi A. The pay-for-performance myth. Bloomberg Business Week, July 222, 2014. Available at http://www.businessweek.com/articles/2014-07-22/for-ceos-correlation-between-pay-and-stock-performance-is-pretty-random. Accessed January 27, 2015.
140. Piketty T, Saez E, Stantcheva S. Optimal taxation of top labor incomes: a tale of three elasticities. Cambridge, MA: National Bureau of Economic Research; 2011.
141. Tosi HL, Werner S, Katz JP, et al. How much does performance matter? A meta-analysis of CEO pay studies. J Manage 2000;26(2):301-39.
142. Schrecker T. Can health equity survive epidemiology? Standards of proof and social determinants of health. Prev Med 2013;57(6):741-4.
143. Marmot M. Inequalities in health: causes and policy implications. In: Tarlov A, St. Peter R (Eds): The society and population health reader, vol 2. A state and community perspective. New York: New Press; 2000.

Kate Pickett, FRSA, FFPH, is Professor of Epidemiology in the Department of Health Sciences at the University of York and previously was a National Institute for Health Research Career Scientist. She co-authored (with Richard Wilkinson) The Spirit Level: Why More Equal Societies Almost Always Do Better and is a co-founder of The Equality Trust. She was awarded a 2013 Silver Rose Award from Solidar for championing equality and the 2014 Charles Cully Memorial Medal by the Irish Cancer Society.

Richard Wilkinson, FRSA is Emeritus Professor of Social Epidemiology at the University of Nottingham, and Honorary Professor at UCL and the University of York. He co-authored (with Kate Pickett) The Spirit Level: Why More Equal Societies Almost Always Do Better and is a co-founder of The Equality Trust. He was awarded a 2013 Silver Rose Award from Solidar for championing equality and the 2014 Charles Cully Memorial Medal by the Irish Cancer Society.



Labor Policy and Work, Family, and Health in the Twenty-First Century

Lisa F. Berkman

Abstract

Socioeconomic disparities in health can be identified with some precision, yet solutions to reduce inequalities and improve overall population health are not as easily specified. Solutions centered on helping poor and lower wage men and women stay in the paid labor market may have health benefits, reduce inequalities in health, and improve overall population health. The Earned Income Tax Credit, pro-family work policies and practices, parental leave, and retirement policies are examples of labor policies that impact the health and well-being of working families. Policies enabling adults to participate in the paid labor force while not risking the health and well-being of themselves or their family members show particular value. Metrics for evaluating social and economic policies do not currently include health metrics. The health outcomes, which are spillovers of such policies, would increase the benefits of those policies in any cost-benefit equations. Developing sound health metrics for the evaluation of seemingly "non-health-related" social and economic policies is essential to improving population health. Such metrics will, however, need to be very specific and tailored to the policies they are intended to evaluate. We want to ensure that Americans, particularly those living in poverty and the working poor, aren't robbed of healthy years of life.

Introduction

The United States has continued to experience modest increases in life expectancy (LE) over the past decades; however, LE in the United States also has declined dramatically in cross-country rankings. It is important therefore to consider what conditions shape patterns of LE in the United States and elsewhere.[1] While most attention has focused on changing behavioral risks (notably smoking), here we explore the interactions between working conditions, labor force participation, and family dynamics, especially for women, because these conditions have enormous potential to influence population health. Furthermore, women's LE has fared worse than men's in these international rankings. When the interactions between work and family are coupled with the major demographic transitions in the world today, they may be among the most powerful social determinants of health.

Fundamentally, the United States, along with most nations, faces a number of very serious demographic transitions, including the aging of their societies, increased numbers of older workers, increased labor force participation among women, and changing patterns of fertility. Re-imagining how work might be organized and constructing innovative approaches to work that acknowledge the shifting demographics and new constellations of families in the 21st century may enable the United States to become not only a healthier nation but also a more productive one.

Here in the United States, for the most part, work and many labor policies are organized as if we were still living as people did in the 1940s when many work policies were established. During that era, there usually was one breadwinner in each family, and people had a LE of 66. In the middle of the 20th century, just over half of Americans who reached the age of 21 could expect to reach age 65. Many workers paying into Social Security would never live long enough to receive benefits, especially African Americans whose life expectancy was just over 50 in 1935. Men retiring today can expect to live about 17 additional years, or about 4-5 years longer than they did in 1940.[2] Women's LE is even longer. These changes are so dramatic that it is no surprise that our expectations cannot fathom them. Furthermore, in the mid-20th century, few men had any caretaking responsibilities while working or in retirement. The contemporary scenario is dramatically different, with most women in the paid labor force and caregiving needs spread out across three or even four generations. Among men, home responsibilities have taken on increased significance over the last years, and work life strains have increased for both men and women.[3] Thus, both men and women may need accommodations in the workplace to incorporate caregiving responsibilities. Furthermore, families with young children, as well as older workers with other family caregiving needs, share the need for increased workplace flexibility.

Many social policies developed in the mid-20th century that appeared to be optimal at the time, may now be costing us in terms of both economic and health gains. Yet, we have rarely evaluated the health impact of these policies and practices. The premise in this chapter is that the limited evaluation of the health impacts of such social policies and the subsequent lack of implementation of new policies based on best evidence to enable both men and women to maintain active work and family lives have cost the United States a great deal in terms of poor health.

U.S. Life Expectancy: Low International Rankings and Rising Inequality

To briefly set the stage for this exploration, overall LE in the United States has lost ground compared to that of other nations in the last decades, especially for women. I was a member of a recent National Academy of Sciences panel on diverging trends in longevity. The panel found that the United States ranked at the bottom among 21 developed, industrialized nations,[1] and poor rankings were particularly striking for women. In the 1980s, U.S. rankings were in the middle among Organization for Economic Co-operation and Development (OECD) countries.[1] While it is true that LE improved during this time by 5.6 years for men and 3.6 years for women, other countries gained substantially more in terms of LE, leaving us behind. Furthermore, almost all of those gains were concentrated among the most socioeconomically advantaged segments of the U.S. population, and they were more substantial for men than for women. The poorest Americans experienced the greatest health disadvantage compared to those in other countries.[4,5] At a recent National Institutes of Health (NIH) conference, the discussion was focused on the steps required for the United States to reach the OECD average—not even the top—in the next 20 years. It seems we have given up on achieving better than average health.

The panel found that the United States ranked at the bottom among 21 developed, industrialized nations and poor rankings were particularly striking for women.

More concerning is the widening gap in mortality rates between those at the bottom and those at the top of the economic ladder in the United States. This gap has widened over the last 25 years. These patterns are evident whether we look at education, income, or wealth differentials. Because the evidence is clearest that education itself is causally linked to health and functioning,[6,7] in this chapter I will focus on these associations. For instance, the mortality rate for men with less than a high school education in 2007 was about 7 per 100. For those with 16 years or more of education, the rate was less than 2 per 100. This corresponds to a 3.5-fold risk of dying in 2007, compared to 2.5 times the risk in 1993. For less educated women, their mortality risk actually increased absolutely during this time, giving rise to an increased risk from 1.9 to 3 in 2007.[8] This pattern holds even if we confine our analyses to white women.[9] While it is true that fewer adults are in the less educated pool in later years, giving rise to questions about selection issues, it is also true that adults in the highest educational categories have grown over this same time period, suggesting increased compositional heterogeneity in these groups. Overall, while selection into education level occurs, it does not appear to completely explain this widening gap.

Although mortality gaps in socioeconomic status have existed for centuries, the magnitude of these differences has grown substantially over time in the United States. These widening disparities suggest that either disparities in the underlying determinants of illness and mortality have also been growing over time or that support to buffer these stressful conditions has changed. In either case, while we may not be able to eliminate health disparities, the fact that the size of the risks varies so much suggests that such large inequalities are not inevitable or innate, giving us hope that there are ways to reduce the burden of illness for our most vulnerable citizens.

Work and Family Strain: A New Model

What can account for such diverging trends? The National Academy of Sciences Panel on Explaining Divergent Levels of Longevity in High-Income Countries[1] was convened to try to reconsider what might account for those diverging trends, and the panel—of which I was a member—considered a very broad range of conditions. Several members identified important behavioral trends, most critically tobacco consumption. Several others examined the role of social networks and economic inequality. These factors seemed to account for either none or very small portions of these diverging trends, although they continue to be significant risk factors for mortality within countries. Smoking clearly accounts for some of the diverging trends, but the full explanation for the diverging trends seemed still to be a mystery. At this time, we must start to consider what else has happened, especially to women, between 1980 and today, and how experiences over the last few decades might account for such diverging trends. One of the factors that must be acknowledged is that the vast majority of women have joined the workforce during this period, including those with young children. In the United States, fertility patterns have changed slightly, but they continue to be relatively steady compared to many other countries, although there is also a substantial rise is single parenthood. This pattern led to the development of a model in which work and family systems combine to influence health. Most importantly, it seemed likely that social and economic policies and practices enacted in many European countries might protect or buffer women from the strain experienced by growing demands from both work and family.

The United States differs from virtually all European and most other industrialized countries in its very low levels of social protection in the face of demographic trends related to women's labor force participation, population aging, and rather steady trends in fertility. Except for the Family Medical Leave Act (FMLA) that extends unpaid leave to those in the labor force (under certain conditions), there are virtually no federally mandated parental or family leave policies in the United States. There are selected State- and city-level polices and many private corporate policies that may protect women and potentially men from the joint strains of work and family demands. This situation makes the United States relatively unique among industrialized countries.

With support from the National institute on Aging, I launched a project to better examine this issue. The project involved an international set of colleagues, including economists, social epidemiologists, and demographers on social protection and women's health. We have created a work-family strain model built closely on the model of job demands and job control developed by Karasek.[10] In this model (Figure 1), job demands are defined by both physical and psychological demand, whereas control is defined as the degree to which an employee can shape work hours or pace or generally have control over his or her schedule or workplace situations. In the Karasek model, often the combination of high job demands and low job control is associated with the worst health outcomes. A third dimension of job strain related to workplace social support was incorporated into the model after initial work. Here we envision a situation in which risks of poor health are particularly high among working women with dependent children or older family members who experience high demands with low control and minimal formal or informal support. We explicitly add to the job strain model a dimension related to the family.

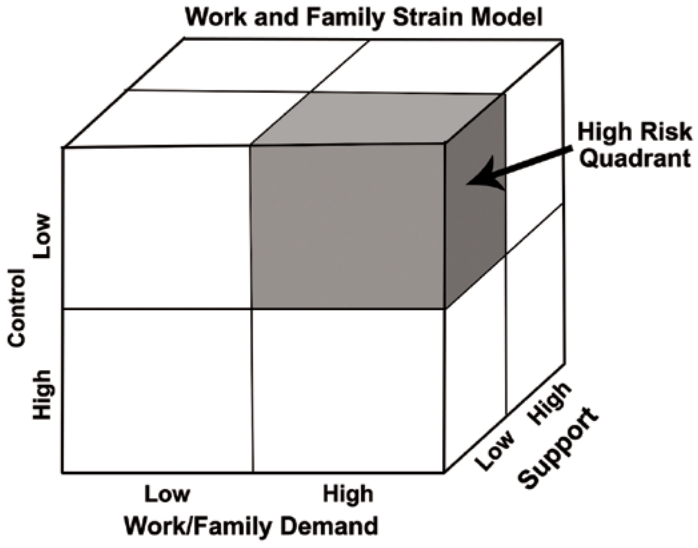


Figure 1. Work/family/support demand control model

As we started to look at women's experiences in other countries, we found that very few of the women experience all three of these conditions together, e.g. high work demand, high family demand, and low formal or informal support. For instance, while France is similar to the United States in terms of trends in fertility and work, the country has very strong social protection policies. In other countries,

women join the workforce, but when they do, their fertility drops. This general framework helped us to evaluate the specific health risks and protections that were important.

Our team of researchers reviewed some evidence on demographic trends and related family policy. There are clear and steady demographic trends regarding the growing number of women in the labor force with young children in United States. In the 1940s and 1950s, when most of our work policies were designed, there were very few women with young children in the workforce. Now, the vast majority of women are in the labor force. By 2008, 71 percent of mothers with children under 18 and 60 percent of mothers with children age 3 or younger were working.[11] At the same time, there was no formal U.S. Federal policy for granting weeks of paid parental leave between 1980 and 2000; most other countries had increased leave policies substantially during this period.

Evidence that Labor Policies and Practices Impact Health

In this section, recent work related to work and family policies is reviewed in terms of health impacts. For the last 10 years, I have been a member of the Work, Family, & Health Network, which was convened by NIH and the Centers for Disease Control and Prevention (CDC).[a] This network has produced some important findings about how work/family dynamics impact health based on both observational and experimental evidence. In addition, I will discuss evidence suggesting that several labor policies, specifically the Earned Income Tax Credit (EITC) and maternity leave policies, impact adult health in addition to bettering the lives of very young children. Finally, some of the recent findings on retirement and cognition shed further light on the range of ways in which work might influence health.

First, it is important to acknowledge that most evidence suggests that people who work and people who have children are healthier than those who do not work and who do not have children.[9] Both selection into these "states" and the causal effects of work and family life on health are likely operating. Understanding these dynamics is important. For these purposes however, we are most interested in the interface of work and family, particularly for women who may experience role strain if formal social protection or informal social support does not buffer the demands of work and family or who may experience the benefits of multiple roles if demands are not overwhelming.

The working hypothesis of the Work, Family, & Health Network is that both work and family can be health promoting in the presence of social protection or support enabling some flexibility during times when heavy work and family demands compete with each other and become stressful. We suspect this is true for all working families. We further hypothesize that in general, low and middle income working class families will benefit more from many social policies than better off families because they have less exposure to corporate social and economic policies that would be of benefit, and they have fewer financial resources to maintain work life balance. A more precise evaluation of these interactions between socioeconomic conditions and social policies is beyond the scope of this chapter but is of central importance in the long-running policy debates.

[a] For more information about the Work, Family, & Health Network, see http://projects.iq.harvard.edu/wfhn/home. Accessed January 7, 2015.

In an observational study of employees working in nursing homes (also called extended care or long term care facilities), we found that workers whose managers were attentive to work-family issues had half the cardiovascular risks as assessed by objective biomarkers from blood or clinical exam and healthier patterns of sleep compared to those who worked for less family-friendly managers.[12] Specifically, employees whose managers maintained family friendly practices were less likely to be overweight, had lower risk of diabetes, and lower blood pressure. Based on objective measures of sleep using actigraphy monitors, these same employees slept almost 30 minutes more per night than their counterparts. A critical aspect of this study was that we interviewed managers about their practices and coded their responses. Respondents themselves were not reporting on manager's practices, thus eliminating a selective reporting bias. In an ongoing randomized controlled trial that was built on these initial findings, we will be able to evaluate whether changing workplaces practices for managers and employees will alter cardiometabolic risks.

The Work, Family & Health Network is also concerned about whether such practices are actually good for the company that implements them. We hypothesize that schedule control, job flexibility, and work-family flexibility are good for the bottom line and good for workers. Here we are concerned specifically with work redesign to reduce work-family conflict. Lead investigators from the University of Minnesota, Moen and Kelly, worked with ROWE (Results Only Work Environment) as it began as an innovation developed and championed by insiders at Best Buy Co., Inc. Cali Ressler and Jody Thompson, both human resources (HR) employees at Best Buy's corporate headquarters, created ROWE. Within 5 years, ROWE had been implemented in the vast majority of departments in the Best Buy corporate headquarters and in other firms as well. In 2005, Moen and Kelly established a research partnership with Best Buy. From 2006 to 2008, Moen, Kelly, and their colleagues observed ROWE sessions and conducted surveys of employees before and after they began ROWE, as well as surveying employees in departments that continued with traditional work practices.[13-15]

Results of this workplace initiative indicate the workplace redesign had positive impacts for the organization and for employees' work, personal lives, and health. Workplace units that experienced ROWE compared to those that did not had reduced turnover, with 6 percent of employees in ROWE leaving the organization within the study period as compared to 11 percent of employees in traditional departments, as well as employees' plans to leave in the future.[12] Employees in the intervention group reported significantly increased schedule control and decreased work-family conflict,[13,14] as well as increased sleep, energy, and self-reported health.[16] There also were positive impacts on smoking, drinking, and exercise frequency.[17] Overall, members of the Work, Family and Health Network have seen that workplace practices that provide schedule control and opportunities to reduce work-family conflict turn out to be health-promoting for employees, and evidence indicates such policies are good for the organization as well.

Turning next to the health impacts of State and national policies, two policies are of special interest: the first related to the EITC in the United States and the second related to variations over time and place in maternal leave policies in Europe. Strully and colleagues[18] assessed whether treating the EITC primarily as a natural experiment for low-income mothers would reveal a causal effect of prenatal poverty on birth weight and selected maternal outcomes. Expansions of the EITC in

conjunction with the 1996 Welfare Reform that happened under President Clinton can be seen as part of a much broader trend of liberalizing U.S. welfare policy, making it so that cash assistance more frequently depends on labor market participation and wages. During this time, there was a huge rise in the EITC permitting an analysis of the EITC compared with the Temporary Assistance for Needy Families (TANF) program. TANF is a cash assistance program for low-income families that does not require employment by the recipient to receive the benefit.

The EITC for unmarried mothers with less than a high school education increased their market wages in significant ways. It also increased the birth weight of infants by almost 16 grams, which is a very remarkable birth weight increase. Even more interesting, it decreased smoking among mothers during this time. The comparison that Strully and colleagues draw in this study between EITC and TANF is related to the differential effects of the policies on smoking.[18] TANF actually increases smoking in mothers. That is, they use that extra income to buy cigarettes, whereas tobacco consumption decreases by 5 percent among unmarried women who receive EITC. Thus, Strully reports a beneficial effect of the EITC for disadvantaged women, allowing them to be in the market, keep earning money, have healthier babies, and improve their health behaviors. This is one example of a set of policies that has been evaluated for the unanticipated spillover effects that social policies may have on health.

Recent evidence from several studies of maternity leave policies in the United States and Europe suggests that, by protecting employment among mothers in the period around birth, maternity leave leads to better long-term labor market outcomes after maternity leave, including wage level and growth, career prospects, labor market attachment, and employability.[19-22] Thus, not only may maternity leave policies benefit children and mothers around the period of birth, they also may have long-term benefits for mothers that extend for decades in later adulthood. Recent work from our group[23] using data from the Study of Health and Retirement in Europe (SHARE), coupled with data from national policies, shows that maternal leave policies in Europe are associated with significant reductions in depression in later life for women who were working and had their first child while living in countries with generous maternity benefits.

Finally, turning to recent research on men and women at a different point in the life course, recent work on retirement policies shows their potential impact on cognition.[24] Retirement may well turn out to have heterogeneous outcomes, with retirement for people working in very physically demanding jobs versus those in less dangerous and/or physically arduous jobs having different health impacts. As we identify more consistent patterns here, policies may well need to identify two trajectories to retirement: one for those who are able to continue working and the other for those with disabilities limiting their capacity to work or those who work in occupations that are physically demanding. The evidence here is not conclusive but rather intriguing. Rohwedder and Willis[24] have looked at SHARE and the U.S. Health and Retirement Study (HRS) to explore the associations between retirement and what they call "mental retirement." They examine the associations between the percent of people not working for pay among 60- to 64-year-olds in a number of countries and cognitive decline in those countries. As countries increase the percent of men aged 60-64 not working for pay, cognitive function among men in this cohort declines. The authors report an association between cognitive function and not working for pay at a country level.[23]

There is a substantial literature on social engagement and cognition suggesting that as people stay socially engaged, their cognitive function is more likely to be maintained.[25,26] Thus, working may be one way that people stay socially engaged.

Directions for Future Research

Health impact evaluations for social and economic policies are rarely conducted, and few methods or metrics have been adapted for such evaluations. In the future, such evaluations will be essential for us to identify the exact health impacts and related costs and benefits so that we can incorporate health impacts into bottom-line evaluations related to economic benefits. Cross-country evaluations are particularly valuable in this regard as they permit us to assess the benefits of national policies across time and place. Health metrics will need to be developed and tailored to policies. For instance, we might anticipate that some policy changes, for example EITC and family leave benefits, will impact health outcomes across the life course and among multiple generations.

Other policies, for instance those related to older workers, might influence outcomes critical at older ages. The development of health impact and related health metrics in early stages and encouraging NIH, the National Science Foundation, and other funding entities to invest in the rigorous health evaluation of current and past policies would yield important information. Such policies are likely to have both spill-over and cross-over effects—that is, they may influence multiple outcomes for individuals, as well as cross over to impact others beyond those directly affected by the policy. Additionally, it would be enormously beneficial to invest in the development of workplace strategies and novel policies and practices that enable women and men to remain in the workforce while also attending to the roles and responsibilities related to caretaking in their families and communities. If these innovations rest on solid experimental or quasi-experimental designs, they will provide policymakers with strong evidence. We have only to look at the case of the woman employed by United Parcel Service (UPS) who was pregnant and needed to take unpaid leave to understand that our country needs to develop more family friendly work based policies.[27] We need much more work on the development of randomized or quasi-experimental studies to identify new, real world policies and practices to enable men and women across the life course with diverse demographic profiles to maintain their engagement in the labor force and care for their families.

Implications for Public Health, Knowledge, and Dissemination

Our international rankings and rising levels of inequality with regard to life expectancy are well known in public health but are not widely recognized or part of the public discourse. The public often assumes—because of our high costs for health care—that the Nation is protected from major morbidity and mortality risks. Public health is an intersectoral field, with the determinants of health ranging from the physical to the social environment. The field of public health would be enhanced by wider communication of the degree to which broad policy sectors, notably labor, education, welfare, and urban planning, shape overall patterns and distributions of population health. This chapter has focused on the impact of work policies in relation to work and family life and the degree to which they may be important determinants of public health.

Conclusion

Research suggests that labor policies and practices that support men and women in the labor force, and especially help those with caregiving obligations, are health promoting. These policies and practices have health effects that are not often "counted" as we think about their costs and the broader range of their benefits. For both young parents in the work force and older workers facing retirement, work appears to be health promoting for those without disability. If we as a society optimize opportunities for work across the life course, we will simultaneously need to develop options for those with disability or in situations where work in not possible. We have not focused on this second option in this work but need to acknowledge the necessity of building policy options in this area as well.

Men and women continue to need opportunities for flexibility and schedule control to enter and remain in the labor force, given the inevitability of having to care for children, parents, or partners at some point in time. Our goal for women and men should be to enable them to be successful, both in the work force and in their family roles if they choose. Currently in the United States, we make this very difficult. Our labor policies challenge working families to remain committed to work and also to their families. For example, over half (54 percent) of low wage earners lack sick leave or vacation time to take care of family members, and around 30 percent of middle income families lack such leave.[28] Even fewer have parental leave.

We have shown that we can identify the socioeconomic disparities in health with some precision. Solutions that help to maintain low income men and women and the working poor in the paid labor force have clear health benefits. The EITC, pro-family work policies and practices, parental leave, and retirement policies are examples of policies that impact the health of working families. Policies enabling adults to participate in the paid labor force while not risking the health and well-being of themselves or their family members show particular value. Metrics for evaluating social and economic policies do not currently include health metrics. The health spillovers of such policies would increase the benefits of the policies in any cost-benefit equations. We want to ensure that Americans, particularly those living in poverty and low and middle wage working families, aren't robbed of healthy years of life.

Acknowledgments

This research was conducted as part of the WFHN (www.WorkFamilyHealthNetwork.org), which is funded by a cooperative agreement through the National Institutes of Health and the Centers for Disease Control and Prevention: Eunice Kennedy Shriver National Institute of Child Health and Human Development (U01HD051217, U01HD051218, U01HD051256, and U01HD051276); National Institute on Aging (U01AG027669); Office of Behavioral and Social Sciences Research and National Institute for Occupational Safety and Health (U01OH008788 and U01HD059773). This work was also supported by a grant from the National Institute on Aging (5R01AG040248-03) and by the MacArthur Foundation Research Network on an Aging Society. The opinions expressed herein are those of the author and may not reflect the position of the Agency for Healthcare Research and Quality, the National Institutes of Health, or the U.S. Department of Health and Human Services.

Author's Affiliation

Director, Harvard Center for Population and Development Studies, Thomas D. Cabot Professor of Public Policy, Epidemiology, and Global Health and Population, Harvard School of Public Health.

Address correspondence to: Lisa F. Berkman, PhD, Harvard Center for Population and Development Studies, 9 Bow Street, Cambridge, MA 02138; email lberkman@hsph.harvard.edu

References

1. Crimmins E, Preston S, Cohen B, eds. Explaining divergent levels of longevity in high-income countries. Washington, DC: National Academies Press; 2011.
2. Social Security Administration. Actuarial life table, 2010. Available at http://www.ssa.gov/oact/STATS/table4c6.html. Accessed January 7, 2015.
3. Galinksky E, Aumann K, Bond J. Times are changing: gender and generation at work and at home. New York, NY: Families and Work Institute; 2008.
4. Avendano M, Glymour M, Banks J, et al. Health disadvantage in U.S. adults aged 50 to 74 years: a comparison of the health of rich and poor Americans with that of Europeans. Am J Public Health 2009;99(3):540-8.
5. Banks J, Marmot M, Oldfield Z, et al. Disease and disadvantage in the United States and in England. JAMA 2006;295(17):2037-45.
6. Lleras-Muney A. The relationships between education and adult mortality in the United States. Rev Econ Stud 2005;72:189-221.
7. Glymour M, Kawachi I, Jencks C, et al. Does childhood schooling affect old age memory or mental status? Using state schooling laws as natural experiments. J Epdemiol Community Health 2008;62(6):532-7.
8. Ma J, Anderson R, Jemal A. Widening educational disparities in premature death rates in twenty six states in the United States, 1993-2007. PLoS One 2012;7(7).
9. Montez J, Hummer R, Hayward M, et al. Trends in the educational gradient of U.S. adult mortality from 1986 to 2006 by race, gender, and age group. Res Aging 2011;33(2):145-71.
10. Karasek R, Theorell T.Healthy work: stress, productivity, and the reconstruction of working life. New York: Basic Books; 1990.
11. Women in the labor force: a data book. Report 1018. Washington, DC: U.S. Department of Labor, U.S. Bureau of Labor Statistics; 2009. Available at http://www.bls.gov/cps/wlf-databook-2009.pdf. Accessed January 7, 2015.
12. Berkman L, Buxton O, Ertel K, et al. Managers' practices related to work-family balance predict employee cardiovascular risk and sleep duration in extended care settings. J Occup Health Psychol 2010;15(3):316-29.
13. Moen P, Fan W, Kelly E. Team-level flexibility, work-home spillover, and health behavior. Soc Sci Med 2013;(84):69-79.
14. Moen P, Kelly E, Hill R. Does enhancing work-time control and flexibility reduce turnover? A naturally-occurring experiment. Social Problems 2011;1(58):69-98.
15. Moen P, Kelly E, Tranby E, et al. Changing work, changing health: can real work-time flexibility promote health behaviors and well-being? J Health Soc Behav 2011;4(52):404-29.
16. Hill R, Tranby E, Kelly E, et al. Relieving the time squeeze? Effects of a white-collar workplace change on parents. J Marriage Fam 2013(5):1014-29.
17. Kelly E, Moen P, Tranby E. Changing workplaces to reduce work-family conflict: schedule control in a white-collar organization. Am Sociol Rev 2011;2(76):265-90..
18. Strully K, Rehkopf D, Xuan Z. Effects of prenatal poverty on infant health: state earned income tax credits and birth weight. Am Sociol Rev 2010;75(4):534-62.
19. Brugiavini A, Pasini G, Trevisan E. The direct impact of maternity benefits on leave taking: evidence from complete fertility histories. Adv Life Course Res 2013;18:46-67.
20. Rossin M. The effects of maternity leave on children's birth and infant health outcomes in the United States. J Health Econ 2011;30(2):221-39.
21. Rossin-Slater M, Ruhm C, Waldfogel J. The effects of California's paid family leave program on mothers' leave-taking and subsequent labor market outcomes. J Policy Anal Manag 2013;32(2):224-45.
22. Ruhm C. Policies to assist parents with young children. Future Child 2011;21(2):37-68.

23. Avendano M, Berkman L, Brigiavin A, et al. Do maternity leave benefits improve mothers health at old age: evidence from 11 European countries during 1960-1010. Soc Sci; submitted for publication.
24. Rohwedder S, Willis R. Mental retirement. Econ Perspect 2010 Winder;24(1):119-38.
25. Ertel KA, Koenen KC, Berkman LF. Incorporating home demands into models of job strain: findings from the Work, Family, & Health Network. J Occup Environ Med 2008;50(11):1244-52.
26. Fratiglioni L, Paillard-Borg S, Winblad B. An active and socially integrated lifestyle in late life might protect against dementia. Lancet Neurol 2004;3(6):343-53.
27. Liptak A. Case seeking job protections for pregnant women heads to Supreme Court. The New York Times 2014 December 1; p A18.
28. Heymann S. The widening gap: why working families are in jeopardy and what can be done about it. New York: Basic Books; 2000.

Lisa Berkman, PhD, is Director, Harvard Center for Population and Development Studies, and Thomas D. Cabot Professor of Public Policy, Epidemiology, and Global Health and Population at the Harvard School of Public Health (HSPH). Previously, she was Chair of the Department of Society, Human Development, and Health at HSPH. She also is the former head of the Division of Chronic Disease Epidemiology at Yale University. She has devoted much of her work to understanding determinants of population health by comparing European countries with the United States. Her research focus has been on social and policy influences on health outcomes. Dr. Berkman has authored several books and more than 200 publications.

Social and Behavioral Interventions to Improve Health and Reduce Disparities in Health

David R. Williams and Valerie Purdie-Vaughns

Abstract

Large racial and socioeconomic status (SES) differences in health persist in the United States. In this chapter, we summarize empirical evidence that suggests that promising interventions exist to address the prominent features of these social inequalities in health in the United States. Research indicates that conditions of life linked to SES and geographic location are important drivers of social inequalities in health. We review research showing that policies and procedures that seek to enhance access to economic resources and improve neighborhood and housing conditions can have positive effects on the SES and health of disadvantaged populations. Also, effectively addressing health inequalities among adults requires a life course approach and efforts to address the accumulation of adversity over one's lifetime. Studies suggest that investing in early childhood interventions can lead to striking improvements in both socioeconomic and health indicators in adulthood. Given that there are racial differences in SES at every level of economic status, we also discuss the need for interventions to address the residual effects of race. We consider values affirmation interventions as an example of a race-targeted intervention that is seeking to identify aspects of racial disadvantage that may be missed by interventions that target an overall population. The challenges and opportunities of successfully reducing SES and racial/ethnic disparities are discussed.

Introduction

In virtually every society, socioeconomic status (SES), whether measured by income, education, occupational status, or wealth, is a strong predictor of variations in health.[1] Similarly, in race-conscious societies such as Australia, Brazil, Canada, New Zealand, South Africa, the United Kingdom, and the United States, racial groups characterized by legacies of social exclusion, economic disadvantage, and political marginalization have worse health than the dominant racial groups in their societies.[2] These large, pervasive, and persistent social inequalities are significant public health challenges of our time.

In this chapter, we review research that suggests we can improve health and reduce inequalities in health through interventions that target the underlying social and psychological conditions that drive health. We begin with salient examples of racial and SES variations in health in the United States. Next, we review research on interventions that have the potential to enhance income, improve

neighborhood and housing conditions, early childhood experiences, and psychological factors linked to stigmatized racial status. Such interventions have the potential to improve both SES and health and reduce social inequalities in health. This review is not exhaustive, but it does seek to showcase research from randomized control trials and other studies that used rigorous evaluation designs. We conclude that there is a pressing need to develop a scientific research agenda for future interventions to reduce social inequalities in health.

Social Inequities in Health

Coronary heart disease (CHD), the leading cause of death in the United States, accounts for one in every four deaths or about 600,000 deaths annually. Table 1 presents national age-specific CHD mortality rates for blacks and whites.[3] Both black men and black women have death rates from CHD that are substantially higher than those of whites from ages 25 through 84 and are especially elevated in early adulthood. A similar pattern is evident for multiple health outcomes in which disadvantaged racial groups have markedly earlier onset of disease, greater severity of disease, and poorer survival than their more advantaged counterparts.[4]

Table 1. Age-specific heart disease death rates for 2010 for whites and blacks and black/white ratios

Age	Males			Females		
	White (W) Rate*	Black (B) Rate*	B/W Ratio	White (W) Rate*	Black (B) Rate*	B/W Ratio
25-34	9.5	20.6	2.17	4.1	11.0	2.68
35-44	33.2	63.5	1.91	13.1	35.5	2.71
45-54	111.2	190.9	1.72	40.7	96.6	2.37
55-64	257.0	437.8	1.70	98.2	218.6	2.23
65-74	536.3	847.8	1.58	268.4	475.9	1.77
75-84	1,475.1	1,807.1	1.23	941.6	1,227.2	1.30
85+	4,943.1	4,202.7	0.85	4,086.7	3,783.8	0.93

Source: National Center for Health Statistics, 2012.[3]

* Rates per 100,000 population.

SES—whether measured by income, education, occupational status, or wealth—is patterned by race and ethnicity,[4] and racial differences in SES contribute to disparities in health. Table 2 considers the complex relationship between race, SES, and CHD.[5] For blacks and whites, both males and females, there is a graded association between education and CHD mortality, with each higher level of education associated with a lower rate of death. Blacks who have less than 12 years of formal education have death rates from CHD that are at least twice as high as their counterparts with a

college degree or more education. These differences are even larger among whites. Similar patterns exist for a broad range of outcomes in the United States and elsewhere in which health improves, generally in a stepwise manner as SES levels rise.[1,6] Thus, although the lowest SES groups face the largest shortfalls in health, income and education affect the health of all individuals within a society.

Table 2 also shows that at every level of education, blacks have higher death rates from CHD than whites, with the black/white ratios increasing with every level of education such that the residual effect of race is markedly larger among college graduates than among blacks and whites who have not finished high school. This suggests that individual-level income and education do not fully account for the multiple components of social and economic disadvantage that are linked to minority racial status. Research reveals that income and education are not equivalent across race, with blacks and Hispanics, compared to whites, having lower earnings at each level of education, less wealth at every level of income, and less purchasing power because of higher costs of goods and services in their communities.[4] Added factors linked to racial/ethnic status also adversely affect the health of disadvantaged minority populations. Minorities live in markedly more health-damaging residential environments than whites and have higher exposure to multiple types of acute and chronic stressors over their life course, including the health-damaging aspects of institutional and interpersonal racism.[7,8]

Table 2. Heart disease death rates, age-standardized, for blacks and whites aged 25-64, 2001

Education	Females			Males		
	Black (B) Rate*	White (W) Rate*	B/W Ratio	Black (B) Rate*	White (W) Rate*	B/W Ratio
All	106.1	37.9	2.80	194.9	100.7	1.94
<12 years	132.9	97.8	1.36	262.9	214.9	1.22
12 years	142.1	50.1	2.84	258.2	145.2	1.78
13-15 years	73.0	25.1	2.91	120.0	73.1	1.64
16+ years	62.8	16.9	3.72	99.2	51.1	1.94
Low/High Ratio	2.12	5.79		2.65	4.20	

Source: Adapted from Jemal, Ward, Anderson, et al., 2008.[5]

* Rates per 100,000 population.

Research reveals, for example, that the opportunities and resources that people have to be healthy are strongly patterned by place. There are multiple pathways by which neighborhood conditions can affect health.[9,10] For example, community variation in educational and economic resources, other social assets, and collective efficacy contribute to social inequalities in educational attainment, income, and employment. Neighborhoods also vary in access to health promoting goods and resources, ranging from the quality of the built environment, public services, commercial resources, and shopping that can promote and sustain health. Health-damaging exposures in homes and

neighborhoods tend to co-occur and are clustered in residential spaces where low SES individuals and racial/ethnic minorities comprise a disproportionate share of the residents. These include the density of fast food outlets and liquor stores, the concentration of tobacco advertising, higher levels of chemicals and pollutants in air, soil, and water, as well as greater exposure to social disorder, including crime and violence.

Interventions to Address Social Inequalities

Scientific evidence indicates that reducing economic and social disadvantage, providing infrastructures that promote economic opportunity, and enhancing income to achieve an adequate standard of living can improve the health of disadvantaged populations.

Increased Household Income and Health

Policy initiatives that provide households with additional income can lead to improved health. The Earned Income Tax Credit (EITC), the largest government cash transfer program to low-income working families in the United States, provides a cash award through the tax system. A study using variation in the Federal EITC over time and the presence of State EITCs found that the additional income reduced the rate of low birth weight and increased mean birth weight, with the effects being larger for blacks than for whites.[11] One study used changes in State EITC as a natural experiment and found that State EITCs increased birth weight and reduced maternal smoking,[12] while another study found that additional income from the EITC reduced hearing limitations among adults.[13] Additional income from the Supplemental Security Income (SSI) program also reduced disability in the elderly, with every $100 increase in maximum monthly SSI benefits found to be associated with a 4.6 percent reduction in disability.[14] Other U.S. research documents that Social Security payments to individuals aged 65 years and older that began in 1940 have been associated with a mortality decline from non-infectious disease among the elderly.[15] Moreover, subsequent legislatively mandated increases in benefit payments were also associated with steeper declines in mortality for the elderly compared to younger people. Similarly, data from 18 OECD countries revealed that pension payments for the elderly were inversely related to all-cause mortality.[16]

The Great Smoky Mountains Study in North Carolina is a natural experiment that has documented that income supplements can improve health and reduce disparities. This study assessed the impact of the extra income that American Indian households received, due to the opening of a casino, on the health of the youth. The study found declining rates of deviant and aggressive behavior among adolescents whose families received additional income.[17] The lower risk of psychiatric disorders in adolescence when the youth lived at home persisted into young adulthood.[18] Additional income was also associated with higher levels of education and lower incidence of minor criminal offenses in young adulthood—especially for the households that were poor at the time of the inception of income supplements—and the elimination of Native American-white disparities on both of these outcomes.[19]

Conditional cash transfer (CCT) programs are initiatives that provide cash payments to low income families contingent on regular health care visits, school attendance, or participation in educational programs. They have been widely used in middle and lower income countries. A review of 13 CCT programs in low and middle income countries, all using experimental or quasi-experimental

designs, found that the programs were successful in increasing the utilization of preventive health services and immunization rates, improving nutritional and health outcomes, and encouraging healthy behaviors.[20] Early large scale programs in Latin America showed striking effects of reduced illness rates, child stunting, and rural infant mortality.[21] Design differences exist across the various CCT studies, and we are uncertain of all the optimal conditions.[20] Evidence indicates that additional cash is responsible for observed effects, although conditionality may be important for political support of CCT programs.[22]

Other evidence of the positive health impact of additional income comes from historical data, which have shown that the widening and narrowing of racial economic inequality have been associated with parallel changes in racial health inequalities. The Civil Rights Movement and related social policies addressing poverty led to improvements in the household income of blacks and a narrowing of the black-white gap in income from the mid-1960s to the late 1970s. The economic gains were larger for black women than for black men, and the improvements in life expectancy for African American females were larger than those of black males and whites.[23] Another analysis of national data revealed that between 1968 and 1978, black men and women, aged 35 to 74 years, had a larger percentage decline in mortality from all causes than their white peers.[24] Increases in life expectancy at birth during this period for black men (4.6 years) and women (5.7 years) were larger on both a relative and an absolute basis than those for whites (2.7 years and 2.8 years, respectively).

Another study documented that black women who were born during the time of improved economic well-being for blacks (1967-1969) had better health status as adults and were less likely to have infants with low birth weight and low APGAR scores than those born earlier (1961-1963).[25] In contrast, the health gains between these two cohorts for white women were negligible. The desegregation of Southern hospitals was also associated with a reduction in the black-white gap in infant mortality in Southern States between 1965 and 1975 and enabled an additional 5,000 to 7,000 black babies to survive infancy during this period.[26] However, the racial gap in health worsened when the economic gap widened. In 1978, black households received 58 cents of income for every dollar whites earned, but the income of blacks fell relative to that of whites during the 1980s. In tandem with the widening economic gap, racial disparities in health widened for multiple indicators of health status during that decade.[8] For example, for 5 years in a row, the life expectancy of blacks declined from the 1984 level, while there were small but consistent increases in life expectancy for whites during the same period.

A meta-analysis of financial incentives in high-income countries found that they were successful, at least in the short-term, in increasing healthy behaviors such as reducing cigarette smoking and drug misuse, with the size of the of the financial payment positively associated with the magnitude of the effect.[27] Similarly, a meta-analysis of nine weight loss trials found no benefit of additional income on long-term weight loss (12 months or longer).[28] On the other hand, a randomized controlled trial focused on smoking cessation found that financial incentives can have long-term effects on health behavior change when the incentive is large ($750) and provided incrementally, with the largest payment ($400) at the 12-month benchmark.[29] In sum, there is a dose-response relationship between financial incentives and behavior change, and health behaviors that are complex need a frequent and incremental schedule of incentives and reinforcement payments, with the timing of the financial reward in close proximity to the behavior to both initiate and sustain changes.[30]

Improving Neighborhood and Housing Conditions

A review of housing interventions that have been evaluated for their health impact concluded that although the assessment of health outcomes has been limited in these studies, improving neighborhood conditions has been associated with better self-reported measures of health.[31] In the Yonkers housing intervention, for example, public housing residents randomized to move to newly constructed public housing with better conditions reported better health, less substance abuse, and less neighborhood disorder and violence 2 years later compared to families that had not moved.[32] People in the intervention group also had higher rates of employment and lower levels of welfare use compared to those in the control group. The strongest evidence comes from the Moving to Opportunity (MTO) project in which low-income public housing residents in five American cities were randomized to a treatment group and two comparison groups. The treatment group received housing vouchers that allowed them to secure housing. After 3 years, criminal victimization was lower, and the health of parents and sons was better in the treatment group.[31] After 10 to 15 years, the treatment group had lower rates of severe obesity and diabetes risk[33] and higher levels of mental health and subjective well-being.[34]

Interventions that Address Early Childhood Conditions

The foundations of SES and health in adulthood are laid in childhood, and preschool interventions have had striking findings. Since 1977, three large randomized control trials with low income, white, African American, and Hispanic adolescent or unmarried pregnant women have documented that the Nurse-Family Partnership (NFP) program positively affects a number of child and parent health and socioeconomic outcomes.[35] Controls in these trials received standard prenatal and postnatal health care services and transportation vouchers to get to their medical care site. Intervention group participants received home visits by nurses during their pregnancy and the first 2 years of their infants' lives. These visits focused not only on improving maternal health practices and facilitating access to any needed treatment but also sought to assist participants with parenting skills, links to needed human services, planning for subsequent pregnancies, and facilitating maternal educational and employment opportunities. The NFP led to improvements in prenatal health-related behaviors, pregnancy outcomes, intervals between the birth of the first and second child, the stability of relationships with current partners, and maternal employment. It also reduced rates of child abuse and neglect, subsequent pregnancies, and use of welfare programs and food stamps. The positive effects of the program were stronger among mothers at high risk (low income, unmarried, or a teenager). Cost-benefit analysis documented an $18,054 per family return to society (due to reduced crime, substance abuse, teen pregnancy, child abuse and neglect, and domestic violence).[36]

In the Michigan-based Perry Preschool Program, African American 3 to 4 year olds from a public housing project were randomized to a 2-year school-based early childhood intervention that included sessions at school and home visits by the teacher.[37] At age 10, children who attended the preschool did not have higher IQ scores than the controls, but they had higher test scores.[38] At age 40, the intervention group had higher income, education, health insurance coverage, and home ownership and lower rates of crime, out-of-wedlock births, and welfare assistance compared to the controls.[37] At age 40, the intervention group also had better overall health and engaged in fewer

risky behaviors (driving without a seat belt, smoking, illicit use of sedatives, marijuana, LSD, cocaine, heroin), although there were no differences in reported medical conditions.[39]

The North Carolina Abecedarian Project (ABC) is a randomized long-term study in which economically disadvantaged infants, born between 1972 and 1977, were randomly assigned to a high-quality early childhood program.[40] The early childhood interventions, which enrolled the children from birth to age 5, consisted of cognitive and social stimulation, caregiving, and supervised play for 8 hours per day. At age 21, the intervention group had fewer symptoms of depression, lower marijuana use, a more active lifestyle, and significant educational and vocational benefits compared to the controls.[41,42] By their mid-30s, children who received the preschool intervention had lower levels of multiple risk factors (elevated blood pressure, metabolic syndrome, and excess weight) than controls.[40] Thus, access to pediatric care, good nutrition, and a safe and nurturing environment in the preschool years translated into better health in early adulthood. Economic analyses reveal that early childhood programs have a net return to society of $3 to $17 for each dollar invested.[43]

Psychological Interventions to Address Racism

As noted, the residual effects of race when SES is controlled reminds us of the importance of intervening on race-related aspects of social experience that can affect SES and health. Research reveals that racial discrimination is ubiquitous in the lives of racial minorities and can lead to elevated risk of a broad range of negative health outcomes and explain some of the residual effect of race when SES is controlled.[8] Given historical and contemporary institutional bias, stereotypes, and discrimination, one need not encounter overt discrimination for the effect of race-related stress to undermine well-being. For instance, an African American patient visiting a white physician could have concerns that he or she will be perceived as unintelligent and noncompliant. Experiencing such concerns, termed "stereotype threat," is stressful and can trigger a coordinated set of neuro-physiological responses involving the hypothalamic-pituitary-adrenocortical (HPA) axis, the cardiovascular system, and the immune system.[44] Psychological interventions can reduce at least some of the negative health effects of race-related stressors and thus contribute to a reduction of health disparities. While they do not eliminate exposure to race-related stressors, psychological interventions can buffer people from their pernicious effects by altering how they perceive, appraise, and respond to psychologically threatening situations. Such interventions can be easy, cost effective, and have enduring benefits, particularly when the efforts are well-timed.[44]

Research shows that the effects of stereotype-threat on academic performance can be mitigated by short, structured writing exercises, called values-affirmations.[44] With this exercise, individuals select and write about a personally relevant value (e.g., humor, honesty, relationships with others, religion). Affirmations buffer people's sense of self-integrity—their sense of being competent, effective, and able to control important outcomes—in environments that are threatening the negative effects of racial stigmatization on academic performance. Thus, people are better able to tolerate and cope with a threat (e.g., poor health) if they affirm their global sense of personal worth in a different domain (e.g., strong family). Affirmed individuals retain their awareness of environmental threats, such as racial bias or awareness of negative stereotypes implicating

their group, but such threats lose some of their power to undermine well-being. Research using randomized, double-blind field experiments shows that values-affirmations raised minority students' academic performance during a 2-year study, reducing the racial achievement gap in this sample.[44]

One recent values-affirmation experiment tested whether the intervention could reduce racial disparities in weight over the course of 2 years.[45] Latino and white college students were randomly assigned to either a values-affirmation or control writing exercise. Results showed that Latinos' perception of social identity threat in their freshmen year was associated with a greater likelihood of being clinically overweight 2 years later. The intervention severed this relationship for affirmed Latino students, leading to a 35 percent drop in the likelihood of being overweight 2 years later relative to controls.[45] Another double-blind randomized psychological intervention, using a variation of values-affirmation termed "social belonging intervention" examined overall health and physician visits.[46] The intervention was a 2-hour procedure implemented once during college students' freshman year and sought to neutralize the psychological perception of threat that many minority college students have. It did so by providing information to students that social adversity on campus was initially common to all students, but temporary, and sought to help them internalize this message. This study documented that the intervention reduced physician visits and improved self-reported health over a 3-year observation period among African American but not white college students.[46] The intervention also improved African Americans' academic performance, reducing the black-white achievement gap in this sample by one half.

Another randomized controlled trial found that a values-affirmation intervention may have contributed to behavioral change in hypertensive patients.[47] In this experiment, all patients received an educational intervention. The intervention group additionally received both positive-affect induction (small gifts throughout the year) and values-affirmation. The experiment found that compared to those in the control group, African Americans in the intervention group had higher medication adherence (as assessed by electronic pill monitors) over 12 months, but there was no effect on blood pressure levels.

A values-affirmation intervention improved patient-physician communication among low SES African American patients with hypertension in a randomized controlled trial.[48] Patients were randomly assigned to complete a values-affirmation task (treatment) or a control task at their health care clinic immediately before their appointment with their primary care provider. The study found that affirmed patients gave and requested more information about their medical condition and their interaction style (based on audio-recordings) and were characterized as being more interested, friendly, responsive, interactive, and respectful and less distressed. Interestingly, no differences were found on self-report measures of satisfaction, trust, stress, and mood.

Research Implications

Our review suggests that a broad range of interventions have the potential to improve SES and health and to possibly reduce or eliminate disparities. Notably, these interventions address factors

outside of the health care system. Medical care is a determinant of health, and there are well-documented disparities by race and SES in access to care and the quality of care.[49] At the same time, contemporary medical care delivery places little emphasis on prevention and on the underlying determinants of illness, and not surprisingly, although the United States leads the world in health care spending, it ranks at or near the bottom of industrialized nations on indicators of health status.[49] Thus, eliminating disparities in health requires investment in the social and economic factors like income, education, and housing that are the fundamental causes of disease and the behaviors that drive the onset of illness. Investing in addressing the social determinants of health has the potential to improve health, reduce social inequalities in health, enhance the quality of life, and even slow down the growth of medical expenditures.[49] Research reveals that addressing social determinants to eliminate disparities would save more lives than expenditures on medical care. For example, one study calculated that 176,000 deaths were averted in the United States during the decade of the 1990s due to declines in overall mortality.[50] It showed that even if we credited all of this decline to advances in medical care, five deaths (a total of 886,000 over the decade) would be averted by eliminating the black-white gap in mortality for every life saved by medical advances.

Yet, there is also much that we do not understand. For example, although research suggests that improving housing and neighborhood quality can enhance health, studies like the Moving to Opportunity (MTO) program[a] have been too small to affect racial segregation and the concentration of poverty in the participating cities (key factors that initiate and sustain poor neighborhood and housing conditions). Importantly, the observed effect sizes are small, and we do not know the specific aspects of the programs that lead to variation in outcome and the number, type, and size of the key components of the multifaceted interventions that are needed to observe maximal impact. Moreover, these studies did not address or dismantle either the institutional or individual mechanisms of discrimination that contribute to residential segregation.[31] It is not feasible to think that the problem of disadvantaged neighborhoods will be solved by moving people, and families should not have to move out of their neighborhoods to live in a better neighborhood. Research is needed to identify the conditions under which improvements in housing and neighborhood conditions can translate into health improvement and to identify the specific underlying mechanisms. We also need to better understand the potential health consequences of mixed-income developments and gentrification processes.

We also currently lack a firm empirical base to determine which strategies to improve health and reduce disparities are likely to have the greatest impact, and we do not know which domains should be tackled first. Research is needed that would provide data on the relative costs and benefits of the full range of promising social and behavioral interventions. This is of critical importance, since social disadvantages tend to cluster in people and places, and cost constraints often limit the ability to implement multiple interventions simultaneously. The Great Smoky Mountains Study[18] highlighted issues of timing, sensitive periods, and sequencing that may matter for the impact of additional income. The reduction in adolescent risk behaviors was seen only in the youngest cohort (age 12 when the income supplements began) who had experienced the longest exposure to the additional income with no effect evident in the two older cohorts (age 14 and age 16) at initial

[a] See Moving to Opportunity for Fair Housing at http://portal hud.gov/hudportal/HUD?src=/programdescription/mto.

supplement. Research is needed that would provide a clearer blueprint of the maximally effective timing and sequencing of specific interventions and the needed clustering and combinations of interventions that are likely to generate the greatest impact for specific health behaviors and indicators of health status. Research is also needed to identify the thresholds of additional income that would be needed to trigger health benefits for various target populations in a range of contexts.

Relatedly, interventions need to be evaluated for the extent to which they are differentially effective across social groups.[51] For example, when CCTs have not targeted the poorest segment of society, groups that are better off economically have received greater benefit.[20] In the future, researchers should give greater attention to the potential of the differential impact of interventions for various subgroups of the population. Higher SES populations are more likely to be aware of, receptive to, and maximally utilize new information regarding minimizing health risks. Thus, the policies most likely to benefit disadvantaged social groups are the upstream interventions that address the underlying social determinants of health.

It is instructive that the studies reviewed here that found clear evidence of reducing disparities in health were interventions that explicitly targeted supplemental income to blacks and American Indians. The Civil Rights initiatives that sought to enhance the economic well-being of disenfranchised and economically disadvantaged racial groups were associated with improved health and a reduction in health disparities. Similarly, the Great Smoky Mountains study,[17] in which supplemental income was provided to American Indians, also documented reductions in disparities. Reductions in disparities in academic performance were also evident for values-affirmation interventions. These interventions were offered to both blacks and whites but reduced the racial gap in academic performance because these theoretically driven interventions were designed to address a race-related risk factor. Future research needs to give greater attention to the effects of intervention for various population subgroups, and interventions should collect the data that would facilitate the identification of any differential effects. Relatedly, the cost-effectiveness of interventions needs to be assessed for population subgroups.

Another subgroup that is important is gender. Reanalysis of data from the Perry Preschool and the Abecedarian and Early Training projects reveals that females benefit more than males from these early childhood interventions.[52] For some outcomes, significant effects are evident only for males. At the same time, the effects for males are larger in the Abecedarian project (an intensive program that began at birth) than in the Perry and Early Training programs,[52] and the health benefits in the mid-30s of the Abecedarian project, as noted earlier, are larger for males than females. These data highlight the importance of explicitly testing for group differences and attending to features of interventions (such as intensity and timing) that may account for variation.

We considered values-affirmation interventions as an example of interventions to address one aspect of racism. Racism adversely affects health through multiple pathways,[8] and future research must deploy a comprehensive set of strategies to minimize the negative effects of the multiple pathogenic effects of racism. Racism's most pernicious effects are likely to operate through institutional mechanisms such as residential segregation that lead to dramatic racial differences in living conditions and economic resources.[53] The racial gap in wealth is an example of a distinctive social disadvantage created by institutional racism. Racial differences in wealth are much larger

than those for income. For example, in 2009, for every dollar of income that white households earned in the United States, black households earned 63 cents, and Hispanics earned 73 cents.[54] More strikingly, blacks have 6 cents and Latinos have 7 cents for every dollar of wealth owned by whites.[55] Data on economic hardship illustrate how blacks are more economically distressed than whites. One national study found that, even after adjustment for multiple indicators of SES (income, education, transfer payments, home ownership, employment status, disability, and health insurance) and demographic characteristics (age, gender, marital status, children, and residential mobility), blacks were more likely than whites to experience economic hardships such as being unable to meet essential expenses; unable to pay full rent or mortgage; unable to pay full utility bills, had utilities shut off; had telephone service shut off; or being evicted from their apartment or home.[56] Thus, the African American and Latino poor are poorer than the white poor, and initiatives targeted at low SES groups need to take the full depth of this racial economic inequality into account.

An important challenge moving forward is to identify how we can best design interventions to ensure maximal benefits to populations in greatest need. A related issue is the need to identify when global interventions can reduce racial disparities and when race-specific interventions are indispensable. Some of the values-affirmation interventions reviewed earlier were used alone, while some were used with other psychological strategies to change health behaviors. More research is needed to identify specific conditions, populations, and health outcomes that modulate the effect of psychological interventions on reducing health disparities. More generally, we need to understand the mechanisms and pathways responsible for observed effects. In the Great Smoky Mountains Study, improved parenting appeared to be responsible for the health enhancing effects that came from additional household income.[19]

Understanding the pathways might help us to identify why the health effects of some interventions are evident for some health outcomes but not others. For example, positive effects of additional income that were evident for alcohol and cannabis abuse and dependence in the Great Smoky Mountains Study did not exist for nicotine dependence.[57] Similarly, additional income from the EITC reduced hearing limitations but was unrelated to self-reported health and functional limitations.[13] We also need to better understand the conditions under which interventions targeted at multiple levels are more impactful. For example, we do not know if we would observe greater health effects in an intervention focused on additional income if it were combined with a values-affirmation intervention.

There is also an urgent need to identify, better understand, and minimize negative unintended effects that were evident for some of the interventions reviewed. In the Great Smoky Mountains Study, negative effects on health linked to additional income were documented. There was an increase in accidental deaths during the specific months that households actually received the cash payments, probably due to increases in vehicular travel and increased substance use.[58] Additional income was also associated with increased adolescent obesity among Indian families whose incomes were low before the receipt of the casino income with no effect for those families whose income had been high.[59] The gains in obesity in adolescence persisted into young adulthood. These findings highlight the importance of being attentive to unintended consequences of interventions that seek to improve health.

Future research and intervention also need to consider that some race-related aspects of life may have broader application to other stigmatized social groups. For example, interventions focused specifically on a race-related aspect of life such as discrimination, may have wider currency. A recent study of white adolescents found that perceived discrimination based on social class accounted for 13 percent of the association between poverty and allostatic load (a summary measure of biological dysregulation), suggesting that discrimination based on social class may be an important but neglected contributor to SES differences in health.[60] Similarly, research on the effects of discrimination on health also finds that although racial minorities report higher levels of interpersonal discrimination than whites, the negative effects of discrimination on health are very similar across races.[61]

Practice Implications

Our review reveals that improvements in living conditions can have decisive health consequences. It implies that a seismic shift is needed in health policy. Health initiatives are needed that pay greater attention to the social determinants of health. They also need to have a dual focus: improving the health of the overall population and reducing gaps in health. While it is true that even the highest SES Americans, in almost every State, are not experiencing a level of good health that is attainable now,[62] there are large shortfalls in health by race and SES, and global strategies that may have the largest impact on improving population health will likely widen disparities.[63] Thus, the implementation and evaluation of health initiatives need to go beyond demonstrating that they improve health for low SES and vulnerable racial/ethnic populations. In order to reduce social inequalities in health, interventions need to be designed to evaluate their impact on health equity. Reducing social inequalities in health will require targeted interventions that improve the health of disadvantaged groups more rapidly than that of the rest of the population.[64] This issue has received inadequate attention in the world of health policy and intervention.

The research reviewed indicates that while there is much we need to learn, there is a substantial and growing body of scientific evidence suggesting that social and behavioral interventions, far removed from traditional health policy, can improve health and potentially reduce shortfalls in health faced by low SES and economically disadvantaged racial and ethnic populations. Braveman and colleagues[65] have persuasively argued that policymakers should not wait for the optimal evidence but should take action based on the best available science. They articulated the importance of using evidence from multiple sources, evaluating the quality of evidence based on multiple criteria with study design being only one factor informing the strength of the evidence, and taking responsible actions in the face of less-than-certain knowledge. Accordingly, decisionmakers should give greater attention to applying our current knowledge to improving the Nation's health. Equally important, efforts are needed to explicitly allocate resources to capitalize on opportunities such as natural experiments that can expand our knowledge base and help us to make strategic investments in social and behavioral research that will strengthen the science of improving health and reducing and ultimately eliminating gaps in health that are avoidable and unfair.

Acknowledgments

Preparation of this chapter was supported by a grant (P50 CA 148596) from the National Cancer Institute. We thank Maria Simoneau and Liying Shen for assistance with preparing the manuscript. The opinions presented herein are those of the authors and may not necessarily represent the position of the Agency for Healthcare Research and Quality, the National Institutes of Health, or the U.S. Department of Health and Human Services.

Authors' Affiliations

David R. Williams, Department of Social and Behavioral Sciences, Harvard TH Chan School of Public Health; Department of African and African American Studies and of Sociology, Harvard University; Department of Psychiatry and Mental Health, University of Cape Town, South Africa. Valerie Purdie-Vaughns, Department of Psychology, Columbia University, New York, NY.

Address correspondence to: David R. Williams, Harvard TH Chan School of Public Health, 677 Huntington Avenue, 6th floor, Boston, MA 02115; email dwilliam@hsph.harvard.edu

References

1. Closing the gap in a generation: health equity through action on the social determinants of health. Final Report of the CSDH. Geneva, Switzerland: World Health Organization; 2008.
2. Williams DR. Miles to go before we sleep: racial inequities in health. J Health Soc Behav 2012;53(3):279-95.
3. National Center for Health Statistics. Health, United States 2012. Hyattsville, MD: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention; 2013.
4. Williams DR, Mohammed SA, Leavell J, et al. Race, socioeconomic status, and health: complexities, ongoing challenges, and research opportunities. Ann N Y Acad Sci 2010;1186(1):69-101.
5. Jemal A, Ward E, Anderson RN, et al. Widening of socioeconomic inequalities in U.S. death rates, 1993–2001. PLoS ONE 2008;3(5):e2181.
6. Braveman PA, Cubbin C, Egerter S, et al. Socioeconomic disparities in health in the United States: what the patterns tell us. Am J Public Health 2010;100:S186-96.
7. Sternthal MJ, Slopen N, Williams DR. Racial disparities in health: how much does stress really matter? Du Bois Rev 2011;8(1):95-113.
8. Williams DR, Mohammed SA. Racism and health I: pathways and scientific evidence. Am Behav Sci 2013;57(8):1152-73.
9. Diez Roux AV, Mair C. Neighborhoods and health. Ann N Y Acad Sci 2010;1186(2010):125-45.
10. Miller WD, Pollack CE, Williams DR. Healthy homes and communities: putting the pieces together. Am J Prev Med 2011;40(1, Suppl 1):S48-57.
11. Hoynes HW, Miller DL, Simon D. Income, the Earned Income Tax Credit, and infant health. Working paper 18206. Washington, DC: National Bureau of Economic Research; 2012.
12. Strully KW, Rehkopf DH, Xuan Z. Effects of Prenatal poverty on infant health: state Earned Income Tax Credits and birth weight. Am Sociol Rev 2010;75(4):534-62.
13. Larrimore J. Does a higher income have positive health effects? using the Earned Income Tax Credit to explore the income-health gradient. Milbank Q 2011;89(4):694-727.
14. Herd P, Schoeni RF, House JS. Upstream solutions: does the supplemental security program reduce disability in the elderly? Milbank Q 2008;86(1):5-45.
15. Arno PS, House JS, Viola D, et al. Social security and mortality: the role of income support policies and population health in the United States. J Public Health Policy 2011;32(2):234-50.
16. Norström T, Palme J. Public pension institutions and old-age mortality in a comparative perspective. Int J Soc Welf 2010;19:S121-30.

17. Costello EJ, Compton SN, Keeler G, et al. Relationships between poverty and psychopathology: a natural experiment. JAMA 2003;290(15):2023-9.
18. Costello EJ, Erkanli A, Copeland W, et al. (2010). Association of family income supplements in adolescence with development of psychiatric and substance use disorders in adulthood among an American Indian population. JAMA 2010;303(19):1954-60.
19. Akee RKQ, Copeland WE, Keeler G, et al. Parents' incomes and children's outcomes: a quasi-experiment using transfer payments from casino profits. Am Econ J Appl Econ 2010;2(1):86-115.
20. Ranganathan M, Lagarde M. Promoting healthy behaviours and improving health outcomes in low and middle income countries: a review of the impact of conditional cash transfer programmes. Prev Med 2012;55(suppl 1):S95-105.
21. Barham T. A healthier start: the effect of conditional cash transfers on neonatal and infant mortality in rural Mexico. J Dev Econ 2011;94(1):74-85.
22. Fernald LCH, Gertler PJ, Neufeld LM. Role of cash in conditional cash transfer programmes for child health, growth, and development: an analysis of Mexico's Oportunidades. Lancet 2008;371(9615):828-37.
23. Kaplan GA, Ranjit N, Burgard S. Lifting gates—lengthening lives: did civil rights policies improve the health of African-American women in the 1960s and 1970s? In: Schoeni RF, House JS, Kaplan GA, et al (eds), Making Americans healthier: social and economic policy as health policy. New York: Russell Sage Foundation Publications; 2008.
24. Cooper RS, Steinhauer M, Schatzkin A, et al. Improved mortality among U.S. blacks, 1968-1978: the role of antiracist struggle. Int J Health Serv 1981;11:511-22.
25. Almond D, Chay KY. The long-run and intergenerational impact of poor infant health: evidence from cohorts born during the civil rights era. Washington, DC: National Bureau of Economic Research; 2006. Available at http://users nber.org/~almond/chay_npc_paper.pdf. Accessed March 3, 2015.
26. Almond D, Chay KY, Greenstone M. Civil rights, the war on poverty, and black-white convergence in infant mortality in the rural South and Mississippi. MIT Dept of Economics Working Paper Series, Report 07-04. Cambridge, MA: Massachussetts Institute of Technology; 2006. Available at http://dspace.mit.edu/bitstream/handle/1721.1/63330/civilrightswaron00almo.pdf?sequence=1. Accessed March 3, 2015.
27. Marteau TM, Ashcroft RE, Oliver A. Using financial incentives to achieve healthy behaviour. Br Med J 2009;338.
28. Paul-Ebhohimhen V, Avenell A. Systematic review of the use of financial incentives in treatments for obesity and overweight. Obes Rev 2008;9(4):355-67.
29. Volpp KG, Troxel AB, Pauly MV, et al. A Randomized, controlled trial of financial incentives for smoking cessation. N Engl J Med 2009;360(7):699-709.
30. Lynagh M, Sanson-Fisher R, Bonevski B. What's good for the goose is good for the gander. Guiding principles for the use of financial incentives in health behaviour change. Int J Behav Med 2011:1-7.
31. Acevedo-Garcia D, Osypuk TL, Werbel RE, et al. Does housing mobility policy improve health? Housing policy debate. 2004;15:49-98.
32. Fauth RC, Levanthal T, Brooks-Gunn J. Short-term effects of moving from public housing in poor- to middle-class neighborhoods on low-income, minority adults' outcomes. Soc Sci Med 2004;59(11):2271-84.
33. Ludwig J, Sanbonmatsu L, Gennetian L, et al. Neighborhoods, obesity, and diabetes — a randomized social experiment. N Engl J Med 2011;365(16):1509-19.
34. Ludwig J, Duncan GJ, Gennetian LA, et al. Neighborhood effects on the long-term well-being of low-income adults. Science 2012;337(6101):1505-10.
35. Olds DL. The nurse–family partnership: an evidence-based preventive intervention. Infant Ment Health J 2006;27(1):5-25.
36. Lee S, Aos S, Miller M. Evidence-based programs to prevent children from entering and remaining in the child welfare system: benefits and costs for Washington. Olympia, WA: Washington State Institute for Public Policy; 2008.
37. Schweinhart LJ, Montie J, Xiang Z, et al. Lifetime effects: The High/Scope Perry Preschool study through age 40. Ypsilanti, MI: High/Scope Press; 2005.
38. Heckman J. Skill formation and the economics of investing in disadvantaged children. Science 2006;312(5782):1900-2.
39. Muennig P, Schweinhart L, Montie J, et al. Effects of a prekindergarten educational intervention on adult health: 37-year follow-up results of a randomized controlled trial. Am J Public Health 2009;99(8):1431-7.
40. Campbell F, Conti G, Heckman JJ, et al. Early childhood investments substantially boost adult health. Science 2014;343(6178):1478-85.
41. Campbell FA, Wasik BH, Pungello E, et al. Young adult outcomes of the Abecedarian and CARE early childhood educational interventions. Early Child Res Q 2008;23(4):452-66.
42. McLaughlin AE, Campbell FA, Pungello EP, et al. Depressive symptoms in young adults: the influences of the early home environment and early educational child care. Child Dev 2007;78(3):746-56.

43. Karoly L, Kilburn R, Cannon J. Early childhood interventions: proven results, future promise. Santa Monica, CA: RAND Corporation; 2005.
44. Cohen GL, Sherman DK. The psychology of change: self-affirmation and social psychological intervention. Annu Rev Psychol 2014;65:333-71.
45. Cook JE, Purdie-Vaughns V, Garcia J, et al. A values-affirmation intervention reduces body mass due to social identity threat. Submitted.
46. Walton GM, Cohen GL. A brief social-belonging intervention improves academic and health outcomes of minority students. Science 2011;331(6023):1447-51.
47. Ogedegbe GO, Boutin-Foster C, Wells MT, et al. A randomized controlled trial of positive-affect intervention and medication adherence in hypertensive African Americans. Arch Intern Med 2012;172(4):322-6.
48. Havranek EP, Hanratty R, Tate C, et al. The effect of values affirmation on race-discordant patient-provider communication. Arch Intern Med 2012;172(21):1662-7.
49. Braveman PA, Egerter SA, Mockenhaupt RE. Broadening the focus. Am J Prev Med 2011;40(1):S4-18.
50. Woolf S, Johnson R, Fryer G, et al. The health impact of resolving racial disparities: an analysis of U.S. mortality data. Am J Public Health 2004;94(12):2078-81.
51. Whitehead M. A typology of actions to tackle social inequalities in health. J Epidemiol Community Health 2007;61(6):473-8.
52. Anderson ML. Multiple inference and gender differences in the effects of early intervention: a reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. J Am Stat Assoc 2008;103(484):1481-95.
53. Williams DR, Collins C. Racial residential segregation: a fundamental cause of racial disparities in health. Public Health Rep 2001;116(5):404-16.
54. Statistical abstract of the United States: 2012 (131st Edition) Washington, DC: U.S. Census Bureau; 2011. Available at http://www.census.gov/compendia/statab/.
55. Detailed tables on wealth and asset ownership. Washington, DC: U.S. Census Bureau; 2014. Available at http://www.census.gov/people/wealth/data/dtables.html. Accessed March 3, 2015.
56. Bauman K. Direct measures of poverty as indicators of economic need: evidence from the survey of income and program participation. Technical Working Paper No. 30. Washignton DC: U.S. Census Bureau; 1998.
57. Costello EJ, Erkanli A, Copeland W, et al. Association of family income supplements in adolescence with development of psychiatric and substance use disorders in adulthood among an American Indian population. JAMA 2010;303(19):1954-60.
58. Bruckner TA, Brown RA, Margerison-Zilko C. Positive income shocks and accidental deaths among Cherokee Indians: a natural experiment. Int J Epidemiol 2011;40(4):1083-90.
59. Akee RKQ, Simeonova E, Copeland W, et al. Does more money make you fat? The effects of quasi-experimental income transfers on adolescent and young adult obesity. IZA Discussion Paper No. 5135. Bonn, Germany; Institute for the Study of Labor; 2010. Available at http://ftp.iza.org/dp5135.pdf. Accessed March 3, 2015.
60. Fuller-Rowell TE, Evans GW, Ong AD. Poverty and health: the mediating role of perceived discrimination. Psychol Sci 2012;23(7):734-9.
61. Williams DR, Mohammed SA. Discrimination and racial disparities in health: evidence and needed research. J Behav Med 2009;32(1):20-47.
62. Egerter S, Braveman P, Cubbin C, et al. Reaching America's health potential: a state-by-state look at adult health. Washington, DC: Robert Wood Johnson Foundation Commission to Build a Healthier America; 2009. Available at http://www.commissiononhealth.org/Report.aspx?Publication=72672. Accessed March 3, 2015.
63. Mechanic D. Disadvantage, inequality, and social policy. Health Aff 2002;21(2):48-59.
64. Lorenc T, Petticrew M, Welch V, et al. What types of interventions generate inequalities? Evidence from systematic reviews. J Epidemiol Community Health 2013;67(2):190-3.
65. Braveman PA, Egerter SA, Woolf SH, et al. When do we know enough to recommend action on the social determinants of health? Am J Prev Med 2011;40(1):S58-66.

David R. Williams, PhD, MPH, is the Florence and Laura Norman Professor of Public Health at the Harvard T.H. Chan School of Public Health and Professor of African and African American Studies and of Sociology at Harvard University. His prior academic appointments were at Yale University and the University of Michigan. He is an elected member of the Institute of Medicine and the American Academy of Arts and Sciences. He was ranked as the Most Cited Black Scholar in the Social Sciences in 2008 and as one of the World's Most Influential Scientific Minds in 2014. Dr. Williams has served as the staff director of the Robert Wood Johnson Foundation Commission to Build a Healthier America and as a scientific advisor to the award-winning PBS film series, Unnatural Causes: Is Inequality Making Us Sick?

Valerie Purdie-Vaughns, PhD, is associate professor, Department of Psychology, Columbia University, core faculty for the Robert Wood Johnson Health & Society Scholars Program. She is the founder and Director for the Laboratory of Intergroup Relations and the Social Mind. Previously, Dr. Purdie-Vaughns served on the faculty at Yale University. Her research focuses on the interplay between social contexts, marginalization (i.e., being an outsider) and human behavior. In particular, she is interested in the interplay between institutions, social situations, and biology and how they affect the development and persistence of marginalization in adolescents and adults.

The Burden of Non-Communicable Diseases in the Developing World: A Role for Social and Behavioral Research

Wendy Baldwin

Abstract

Developing countries are undergoing an epidemiological transition, with a shift from mortality predominately driven by infectious diseases to mortality driven largely by non-communicable diseases (NCDs). NCDs—primarily cardiovascular disease, diabetes, chronic respiratory disease, and most cancers—are strongly influenced by social and behavioral factors. These behaviors, which have multiple drivers, are tobacco use, excessive alcohol use, poor nutrition, and sedentary lifestyle. The increasing impact of these behaviors on mortality highlights the importance of a lifecycle approach to lowering risk of disease. While there is much to do at different life cycle stages, there is a window of opportunity to focus research, policy, and programmatic attention on adolescence. Adolescence is the time when both tobacco and alcohol use typically begin, and it is a time of increasing independence. It is also a time to ensure that healthy eating and regular exercise are solidified as habits for a lifetime. At a time of rising concern about the current and future burden of NCDs in the developing world, there are a number of opportunities involving data, research, programmatic, and policy initiatives centered on the role of adolescents that should be pursued.

Introduction

The predominance of death from non-communicable diseases (NCDs), as opposed to infectious diseases, is the hallmark of the epidemiological transition.[1] In high-income countries, this transition occurred generations ago, and now NCDs generally represent diseases and causes of death of the elderly. However, NCDs are not just diseases of older, wealthier populations; they are now poised to be the dominant cause of death in developing countries. The earlier onset of NCDs in low-income countries means that deaths and disability occur during economically productive ages. Since the behavioral risk factors that drive most of these diseases can be addressed during adolescence, there is an opportunity for primary prevention. While primary prevention does not eliminate the need for secondary prevention and treatment of disease, it does offer an opportunity to lower the burden on individuals, families, health systems, communities, and nations.

The burden of NCDs on the developing world has been recognized by the World Health Organization. In 2011, the United Nations General Assembly held a special meeting on the topic, only the second such meeting focused on a health issue.[2] Notably, the first was 10 years earlier and addressed HIV and AIDS. There has been continuing attention from WHO on the challenges that NCDs present for low- and middle-income countries, as well as for more wealthy nations, as can be seen in the NCD action plan which includes specific targets for improved use of medications as well as reductions in risk factors.[3] The WHO action plan provides a wide-ranging blueprint of actions, targets, and potential actors that, together, could be a formidable global response to the crisis NCDs represent. Clearly, much needs to be done to strengthen screening, diagnosis, and treatment. Also, for a comprehensive response, many sectors outside of the health care arena need to be involved. One way to heighten attention to primary prevention and articulate what other sectors could do is through attention to adolescence as a critical life transition point and an opportunity for intervention to lower lifelong risks for NCDs.

While the focus in this chapter is on the potential for intervention during adolescence, other life stages also represent important intervention points. A broad life-cycle approach has many potential benefits, since there is a growing literature linking nutrition in early years—even prenatally—to diet and the ability to maintain a healthy weight in later life. However, two of the main risk factors—tobacco and alcohol use—typically begin during adolescence, and as young people take more responsibility for their lifestyle choices, adolescence may be an ideal time to reinforce healthy habits. Establishing and/or reinforcing healthy habits during adolescence is not without challenges, but changing well-ingrained habits later in life is likely more difficult. Adolescence may not be the best time to instill positive behaviors, nor is it the last time, but it presents a valuable window of opportunity at a critical life cycle stage.

Non-Communicable Diseases in Low- and Middle-Income Countries

Developing countries—that is, the low- and middle-income countries (LMICs)—now face an increase in the proportion of deaths that result from NCDs (Figure 1). Of course, this would be expected with the decline in infectious diseases. But, the proportion of deaths attributable to NCDs is only part of the story. The mortality rate—deaths per 100,000 population of all ages—is 705 for males and 520 for females for the world as a whole. However, in Sub-Saharan Africa the comparable figures are 869 and 746.[4] It might be tempting to think that these are diseases of the elderly. Death is still a lamentable (but unavoidable) outcome for all of us, but age of death does vary and can be influenced. In LMICs, the age of death from NCDs is far earlier and is more likely to occur during the economically productive years than in higher income countries. Premature death is an additional trauma for families that may rely on that individual for their economic well-being, happiness, and even survival. In high-income countries, 13 percent of deaths due to NCDs were premature, i.e. before the age of 60. However, in middle-income countries, the proportion of premature NCD mortality is 28 percent, and in low-income countries it is 41 percent.[5] Furthermore, the morbidity that often precedes a death caused by an NCD can include serious, life-altering consequences, such as amputations from complications of diabetes or disability from chronic cardiac or respiratory diseases.

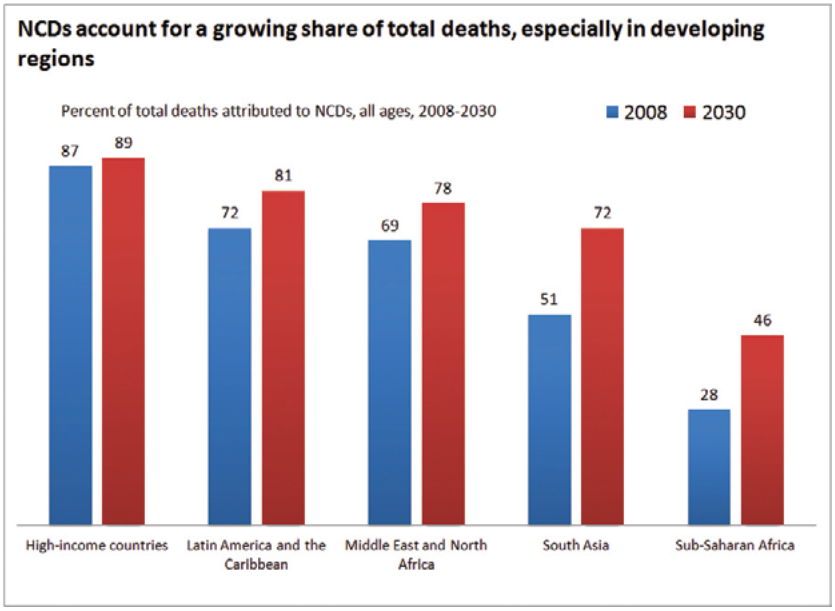


Figure 1. Projected burden of deaths from chronic diseases

Source: Baldwin W, Kaneda T, Amato L, et al. Noncommunicable diseases and youth: a critical window of opportunity for Latin America and the Caribbean. Washington, DC: Population Reference Bureau; 2013. Used with permission.

The economic impact of NCDs has been assessed by the World Economic Forum, which projects the cumulative impact by 2030 of $47 trillion or 5 percent of the global GDP.[6] In fact, the costs of addressing NCDs could actually undermine the economic growth that has been shown to date. The World Bank has stated that "the developing countries cannot afford to treat their way out of this,"[7] and the United Nations Development Program has declared that the rise in NCDs threatens those economic gains made in the developing world.[8]

It is appropriate to ask about the ability of health systems to meet this challenge, and clearly, they will have to be strengthened. NCDs are typically, but not always, chronic diseases that may have an onset years before a death occurs. In addition to the impact on the individual and the family, this also means that there is a prolonged impact on the health care delivery system. Much of health care around the world is provided by the family, but the advances in drugs to help manage diabetes, hypertension, cardiovascular disease, chronic respiratory diseases, and many cancers bring expectations for the health care delivery system as well. Discovery of elevated blood glucose, blood pressure, or cholesterol usually comes with access to a robust primary care, including specific screening programs. The potential for management of risk (secondary prevention) through medications is quite impressive, but there needs to be health education, availability of essential medicines, and access to a health system that is able to sustain potentially decades of need for those affected.

Lessons learned from dealing with diseases such as HIV infection and AIDS can be helpful, but the potential costs of strengthening health systems to meet the kinds of demands NCDs will present is daunting. Also, the low- and middle-income countries have not conquered the challenges of

infectious diseases, and thus, many will face a "double burden" of disease. The outbreak of Ebola in West Africa in 2014 is a reminder of the potential of infectious diseases and the limitations of health systems in many countries.[9] A growing awareness of the burden imposed by NCDs is leading to increased surveillance of mortality and risk factors. The Institute for Health Metrics and Evaluation at the University of Washington has been a leader in evaluating the current and future burden of disease. In their report "Global Burden of Disease 2010," they concluded there were several big challenges, all of which are relevant for the present analysis.[10] First, there is the demographic transition and projected longer lifespans that make healthy aging critically important. Second, there is an epidemiologic transition that will result in the move toward NCDs as the overwhelming cause of death and a shift to concern for disability and not just mortality. The growth in disability will be influenced by the extent to which healthy lifestyles are established and maintained. Also, there is a change in the risk factors from those that reflected environmental factors (e.g., clean water) to those that are behavioral (e.g., healthy lifestyle). It is not surprising that the WHO has focused on the four behavioral risk factors that form the basis for the rise in NCDs. Finally, while health systems are facing enormous challenges, attention to primary prevention of NCD risks involves many other sectors besides the health system that can lower future burdens on the health system itself.[10]

The World Bank has projected that there will be a slight rise in deaths attributed to NCDs in high-income countries.[11] They expect the rise to be slight because the proportion is already so high (see Figure 1); in other parts of the world, however, the increase will be substantial. In Latin America and the Caribbean, the proportion of deaths due to NCDs is fairly high now at 72 percent, but this figure is projected to rise to 81 percent by 2030. In South Asia the current level is about half (51 percent), but it is projected to rise to 72 percent by 2030. Africa is still heavily influenced by infectious diseases, so the proportion of deaths due to NCDs is only 28 percent, but the increase will be stunning over the next 15 years or so, to 46 percent by 2030.

The continued rise in deaths attributable to NCDs is foreshadowed by the rise in risk behaviors, such as tobacco use, obesity, and sedentary lifestyle. These risk behaviors are typically higher in urban areas than in rural areas, so the trend toward greater urbanization portends rising risk factors. Also, unlike in industrialized countries where these risks are often higher in lower income groups, in developing countries the levels are often higher in higher income groups.[12,13] The National Institute on Aging has supported considerable research on non-communicable diseases in the elderly in low- and middle-income countries and observed the pattern that as economic conditions improve in a country, more people may have discretionary income to spend on tobacco, alcohol, and higher calorie food.[14] They may be able to alter the conditions that kept them physically active (car ownership rises with rising income) and so their risks rise. Urban areas that are magnets for growth may provide fewer opportunities for physical activity and healthy diet. Many developing countries have growing economies, and as people are able to move into a better standard of living (frequently in an urban area), we can expect to see the risk factors for NCDs rise as well.

The World Health Organization focused on the four NCDs that are driving the epidemiological transition—diabetes, cardiovascular disease, chronic respiratory disease, and most cancers.[15] These four NCDs are responsible for 80 percent of the NCD mortality rate, and they also tie to four shared underlying risk factors: tobacco use, excessive alcohol consumption, unhealthy diet/obesity, and sedentary lifestyle.

Tobacco use is the single most important modifiable risk factor; cigarette smoking is responsible for 71 percent of lung cancer deaths, 42 percent of chronic respiratory disease, and 10 percent of cardiovascular disease (mortality). Excessive use of alcohol has many unhealthy sequelae, including heart disease and some cancers. Unhealthy diet—i.e., a diet that is high is sugars, saturated fats, trans-fatty acids, and salt and low in fruits and vegetables—and resulting obesity are linked to high blood pressure, type 2 diabetes, and heart disease. Poor diet and low physical activity often result in obesity, which is now viewed as a global epidemic.[15]

Obesity is not just a health risk, it is also a reflection of reduced ability to function productively. Finally, sedentary lifestyle is estimated to result in a 20-30 percent increase in all-cause mortality.[5] Two of these risk factors—tobacco and alcohol use—typically begin during adolescence, and adolescence provides an opportunity to establish and solidify good eating and activity levels. This is not without challenges, but perhaps it is not as daunting as waiting and trying to change deeply entrenched behaviors later in life.

In addition to their role in the four NCDs described above, these risk behaviors affect other health issues as well. Tobacco use is a risk factor for poor pregnancy outcome, for example. Excessive alcohol consumption is often a trigger for interpersonal violence and for accidents and injuries, especially those that are traffic-related. Road traffic accidents and interpersonal violence are the leading causes of death among young people worldwide.[16] The link to productivity can perhaps best be seen in the fact that in the United States, the major health reason for young men and women being refused entry into the military is overweight or obesity.[a,17]

Risk Behaviors Among Adolescents

A first step is to understand the risk behaviors of youth. There are several key data sources that provide insight into the level and nature of compromising health behaviors among adolescents in the developing world. The Global Youth Tobacco Survey (GYTS) is a school-based survey that collects data on students aged 13-15 using standardized methods across the multiple countries implementing the survey.[18] That ensures that the method for establishing the sampling frame, selecting schools and classes, and processing of data are comparable. The structure of the sample allows for national estimates with regional level stratification possible. While there is a common protocol, countries may add questions. The questionnaire is self-administered and address multiple measures of consumption along with other knowledge and attitude measures. To date, 145 countries have participated. Funding is provided by the Canadian Public Health Association, the National Cancer Institute (NCI), the United Nations Children's Fund (UNICEF), and the World Health Organization (WHO).[19]

The Global School-Based Student Health Survey (GSHS) was developed by WHO in collaboration with UNICEF, UNESCO (United Nations Educational, Scientific, and Cultural Organization), and UNAIDS (the Joint United Nations Programme on HIV/AIDS) and with technical assistance from the Centers for Disease Control and Prevention (CDC).[20] This, too, is a school-based survey

[a] Jack Dilbeck, Research Analyst, United States Accessions Command, Fort Knox, KY, as cited in Ready, Willing, Unable to Serve, report by Mission Readiness: Military Leaders for Kids; cdn missionreadiness.org/MR-Ready-Willing-Unable.pdf. Accessed January 13, 2015.

conducted primarily among students aged 13-17. The purpose is to collect data on health behaviors and protective factors that can be used for international comparisons and the establishment of trends and serve as a basis for countries to develop policies, establish programs, and address youth health needs. This survey also uses a standardized sampling process and core questionnaire modules. In some locations, countries have added specific questions to supplement the core questions. Data are collected with a self-administered questionnaire that can be completed in one regular class period. The core modules are alcohol use, dietary behaviors, drug use, hygiene, mental health, physical activity, protective factors, sexual behaviors, tobacco use, and violence/unintentional injury. More than 100 countries have had people trained in the methodology, and 73 have completed at least one GSHS. More than 420,000 students have participated in the surveys.[20]

A third survey provides international comparisons with data gathered from a household survey instead of a school-based survey. The Demographic and Health Surveys (DHS) have been conducted in more than 90 countries and, for many, in multiple waves.[21] Funded by USAID (the U.S. Agency for International Development), the DHS are a valued resource for international comparisons of health and reproductive behavior. The core DHS survey is of women in the reproductive years, with data collection generally beginning at age 15. There are some country variations, and some countries also include men in their surveys. Sharing a common core protocol, there are country-country variations within the standardized format and training for implementing the DHS. Since the DHS typically includes adolescent women, it provides an additional data point for measures of obesity, smoking, and alcohol consumption.

Each of these surveys provides a view into the health behavior of youth. The GSHS and GYTS are school-based surveys, which could of course limit their coverage. However, school enrollment/attendance is increasing around the globe, helped by the focus given with the United Nations' millennium development goals and activities.[22] But, there could be some gaps in coverage among population groups that have low school involvement. Also, the risk of gaps in coverage is greater when older teens are being considered, but even at young adolescent ages, this can be problematic if specific groups (girls or rural or ethnic groups) have low school enrollment. The DHS is primarily a survey of women in their reproductive years, although some sites include men in the survey. All three of these surveys generate data that are publicly available and built on a shared platform so that the survey design and questionnaires are comparable.

Individual countries, or groups of countries, may conduct other surveys of youth risk behaviors covering an entire country or specific regions, cities, or schools. These surveys provide an excellent complement to the broad national data and, depending on the local circumstances, may go into more details on specific risk factors. The STEPS (STEPwise approach to surveillance) provides a tool to gather data about NCD risks. Supported by WHO, it enables countries to gather data about youth risks, such as tobacco, alcohol, physical activity, obesity, and diet. STEPS is a valuable resource, but individual country surveys are not necessarily focused on youth.[b,23]

We might conclude that basic surveillance of youth risk factors does exist for many countries, although not for all countries and not always for identical time periods and ages. There is also a

[b] Note that in Colombia, the age range for the STEPwise Survey was 15-64, but in some countries it was 25 or older. The survey from Colombia (Spanish language only) can be found at http://www.who.int/chp/steps/2010_STEPS_Survey_Colombia.pdf.

need for more surveys and repeat surveys to assess changes over time in the levels of behaviors that place youth at risk for later disease. The desire for more and better data is universal, but it should not detract from the basic picture of risks that is before us.

Extent of Risk to Youth in Developing Countries

The Population Reference Bureau (PRB) has arrayed data for youth risk behaviors in LAC (Latin America and the Caribbean) to provide an overview of risks in that region (see Figure 2). The measures were mostly from school-based surveys and involved 13-15 year olds.[24] There were 27 countries with recent (2008 or later) data on at least three risk categories. For each risk behavior, cut offs were determined to identify countries as being at high, medium, or low risk (see technical notes to the PRB data sheet for a full discussion).

Tobacco

For tobacco use, countries were considered at high risk if the proportion of youth aged 13-15 smoking in the past 30 days was 16 percent or more and at medium risk if the proportion was 7-15.9 percent. For the region, eight countries were at high risk for tobacco use for male youth and five for females. In no case was the risk level for females greater than the risk level for males. Fourteen countries were at medium risk for tobacco for males and 13 for females. Canada, Puerto Rico, Antigua and Barbuda, St. Kitts-Nevis, and Brazil were at low risk for both males and females. Discussion of the definitions and cut offs for each risk factor can be found on the PRB Web site.[24]

Research generally shows that early onset of tobacco use is associated with high levels of use in later life and greater difficulty in quitting. Also, it also signals a likely longer period of exposure to tobacco.[25,26] The ongoing accumulation of evidence about the risks of smoking and early onset of smoking has led to policy and programmatic strategies to impede initiation and facilitate cessation of smoking. There has been concern that the tobacco companies specifically target youth with advertising and marketing approaches.[27] One might conclude that there is no health benefit of tobacco, youth are smoking at significant and often increasing rates, and prevention efforts need to accommodate to the potential for directed marketing toward young people.

While there is less information about changes over time in risk factors, there are indications that tobacco use may be declining in some of the most developed countries. In the United States, for example, youth smoking declined between 2000 and 2011, but the decline has not continued.[28] In some countries where overall levels of smoking have declined, the gender differential has narrowed. Generally, males are more likely to smoke than females, but the rates for females seem to be rising. In some countries, such as Argentina and Uruguay, the rates of smoking for girls exceeds that for boys, with 27 percent of girls and 21 percent of boys in Argentina having smoked in the past 30 days; in Uruguay, the rates of smoking are 23 percent for girls and 16 percent for boys.[29]

Overall, of those surveyed in the GYTS who were aged 13-15 at the time, the median percentage for having ever smoked cigarettes was 33 percent. One benefit of these data is that they also ask about age of initiation, which for most is before age 18 and for a significant number of respondents, is before age 10. For all countries surveyed, the median percentage who had their first cigarette before age 10 was 23.9 percent, although it is above 80 percent in parts of India.[30]

	Population & Youth							NCD Mortality & Prevalence							NCD Risk Factors Among Youth							
	Total Population (millions) 2012	Percent of Population Ages 10-24	Secondary School Net Enrollment Rate, 2007/2010		Adolescent Fertility Rate, 2010-2015	Percent Living in Urban Areas	GNI PPP per Capita (int'l $) 2011	Age-standardized Death Rate for All NCDs (per 100,000) 2008		% of All Deaths due to NCDs	Hypertension Prevalence among Adults (%)		Type II Diabetes Prevalence among Adults (%)		Cigarette Use		Alcohol Use		Physical Inactivity		Overweight or Obese
			Male	Female				Male	Female		Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Female
NORTHERN AMERICA																						
Canada	34.9	19			11	80	39,730	387	265	89	16	17	8	7								
United States	313.9	20	89	90	27	79	48,890	458	326	87	32	30	9	8								
Puerto Rico	3.7	23			51	99					33	35	13	13								
CENTRAL AMERICA																						
Belize	0.3	32	64	65	71	44	6,070	507	455	62	29	24	8	18								
Costa Rica	4.5	27			62	62	11,950	431	333	81	26	25	8	8								
El Salvador	6.3	33	57	59	89	63	6,690	539	449	67	21	19	9	7								
Guatemala	15.0	33	43	44	98	50	4,800	503	421	47	13	14	9	8								
Mexico	116.1	27	70	73	66	77	15,060	543	412	78	32	31										
CARIBBEAN																						
Antigua and Barbuda	0.1		85	85		30	15,670	544	511	80	47	38										
British Virgin Islands	0.03		79	89		40					41	31	12	12								
Cayman Islands	0.05		98	94		100					25	26	6	9								
Dominica	0.1		84	88		67	12,460	682	519	85	32	32	22	12								
Grenada	0.1	30	95	86	35	40	10,530	722	442	81												
Jamaica	2.7	29	80	87	72	52	7,770	498	479	68	34	37	10	15								
St. Kitts-Nevis	0.1		89	88		32	14,490	621	553	83	38	32	5	10								
Saint Lucia	0.2	27	85	85	56	28	9,080	597	405	78	46	38										
St. Vincent and the Grenadines	0.1	27	84	86	54	40	10,560	649	509	76												
Trinidad and Tobago	1.3	23	65	70	32	13	24,940	896	506	78	35	28										
SOUTH AMERICA																						
Argentina	40.8	25	78	87	54	91	17,250	613	366	80	32	33	8	10								
Brazil	194.3	26			76	84	11,500	614	428	74	22	25	5	6								
Chile	17.4	24	83	86	56	87	16,330	501	313	83	29	25	8	10								
Colombia	47.4	27	72	77	84	76	9,640	438	351	66	28	19	3	2								
Ecuador	14.9	28	73	74	81	66	8,310	434	336	65	43	35										
Guyana	0.8	32	78	83	101	29	3,460	735	602	66	43	38										
Peru	30.1	29	77	78	63	74	10,160	408	339	60	38	31										
Suriname	0.5	27	49	55	35	67	7,710	696	450	71	34	27		20								
Uruguay	3.4	23	68	76	59	94	14,740	651	378	87	33	31	6	5								

Figure 2. Youth risk data sheet

Source: Baldwin W, Kaneda T, Amato L, et al. Noncommunicable diseases and youth: a critical window of opportunity for Latin America and the Caribbean. Washington, DC: Population Reference Bureau; 2013. Used with permission.

Alcohol

For alcohol use among adolescents, countries were identified as at high risk if the proportion of youth aged 13-15 who had consumed alcohol in the past 30 days was 40 percent or more and at medium risk if the proportion was 20-39.9. For alcohol, as for tobacco, the measure chosen was use in the past 30 days.[24]

Alcohol consumption is a more difficult risk behavior to assess than tobacco use, since there can be positive or at least neutral effects of moderate drinking. The "risk factor" is usually described as excessive or harmful alcohol consumption. While it might be good to have more nuanced measures of risk, we presumed that reporting alcohol use at ages 13-15 constituted some level of risk behavior.

In the Latin American/Caribbean countries included in the study,[24] 10 countries were at high risk for youth alcohol use for both males and females and an additional three countries for males only. Eleven countries were at medium risk for both males and females. Only El Salvador and Guatemala were at low risk for both males and females in the country as a whole.

Excessive alcohol use is clearly a risk factor for later disease as well as a trigger for interpersonal violence, including gender-based violence and road traffic accidents (Figure 3).[31] Earlier onset of alcohol use, with or without a family history of alcoholism, is a risk for alcohol abuse or alcoholism.[32,33] This provides further reason for concern for drinking among adolescents, especially when alcohol use among 13-15 year olds is over 50 percent in a number of countries.[24]

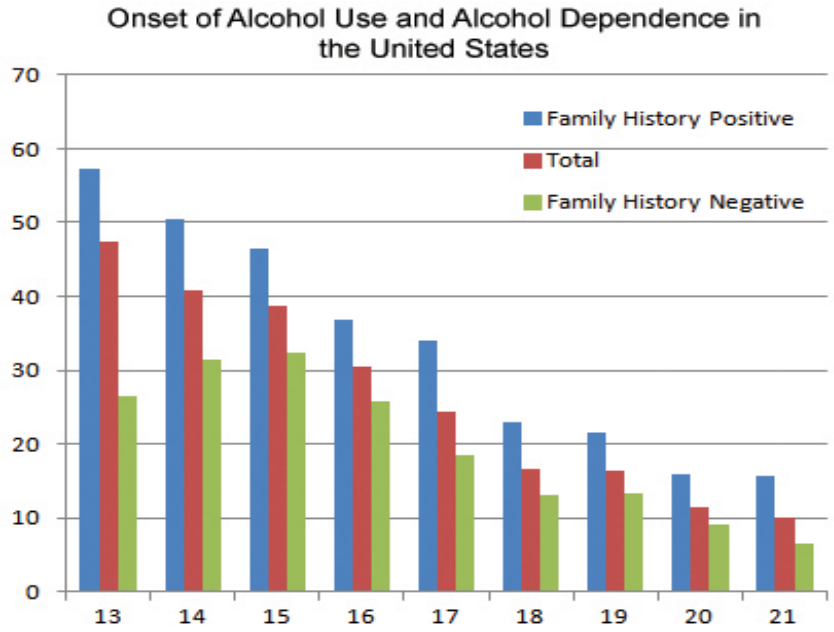


Figure 3. Onset of alcohol use and alcohol dependence in the United States

Source: Baldwin W, Kaneda T, Amato L, et al. Noncommunicable diseases and youth: a critical window of opportunity for Latin America and the Caribbean. Washington, DC: Population Reference Bureau; 2013. Used with permission.

Reports from the Global School-Based Student Health Survey[20] show the proportion of school age youth who report drinking in the previous 30 days. Some countries with largely Muslim populations—where alcohol use rates are very low—do not collect data about adolescent experiences with alcohol.[34] Although some countries have similar rates for males and females, when they differ it is where males are more likely to drink than females. Perhaps more alarming is the percentage reporting binge drinking, i.e. five or more drinks at one time. For example, in Guatemala City, over 11 percent of school-going youth aged 13-15 reported drinking so much that they were "really drunk" one or more times in their life, a figure that was about the same for boys as for girls.[35]

Diet, Physical Activity, and Body Mass Index

While the many aspects of diet are more difficult to measure and track across countries, it is possible to look at the levels of obesity or overweight, which is the typical outcome of a diet that is low in fruits and vegetables and high in sugar and processed carbohydrates combined with a sedentary lifestyle. Data from the Demographic and Health Surveys (DHS) show worrying levels and rises in body weight among women across many countries.[36] This may seem out of place for countries still battling malnutrition and childhood stunting, but India has as many obese in their population as they do stunted, a situation mirrored in many countries in the Middle East.[37]

In the 22 countries in Latin America and the Caribbean where there were data for female obesity/overweight, none were below 10 percent of the population being overweight or obese; five were at moderate risk (10-19.9 percent), and 17 were at high risk, with 20 percent or more of the adolescent female population being overweight or obese.[24] Recent press attention has focused on Mexico, which now exceeds the United States in obesity,[38] but Mexico is taking innovative steps toward prevention, as well as steps to ban television ads promoting high calorie foods and soft drinks.[39]

Data for physical activity among males and females in Latin America and the Caribbean show that 14 countries are at high risk for insufficient physical activity, with 70 percent or more youth reporting that they did not meet the WHO guidelines of 60 minutes or more a day for 5 out of the last 7 days.[24,40] Where there is a gender difference, it favors males—with males being more physically active. Levels of physical activity appear to fall with urban living and with many youth failing to meet the minimum levels of activity of 1 hour a day, at least 5 days a week. NCDs are a burden in the developing world now, and yet the underlying risk behaviors are increasing. This would lead us to expect an even greater burden of NCDs in the future if there is no improvement in the risk behaviors discussed here.

Searching for Successful Interventions

Although a comprehensive review of all interventions on these four risk factors—tobacco use, excessive alcohol consumption, unhealthy diet/obesity, and sedentary lifestyle—is beyond the scope of this paper, there are indications that where there is appreciation of the risks, there have been attempts to address them. These actions may be at the level of individual behavior change up to broad policy changes. There have been few rigorous evaluations of such interventions in the developing world, but there are some promising approaches. Interventions may be aimed at

individual behavior change or focused on changing the social and cultural environment. Some broad initiatives may not have a specific focus on youth and, if aimed at the total population, may have a differential impact on youth. For example, analyses in different countries show that an increase in taxes on tobacco is associated with declines in use (Figure 4).[41] Also, the impact on use appears greater for younger smokers (or potential smokers). In the United States, as well as in other countries, a growing awareness of the risks of tobacco have restricted where one can smoke, making it both more inconvenient and less acceptable to smoke. While there may be concern that adolescents would still seek to smoke as a sign of adult behavior or defiance, even in most high-prevalence countries, the majority of youth do not smoke, and rates of smoking have fallen in many countries that have launched broad anti-smoking campaigns.[42-44]

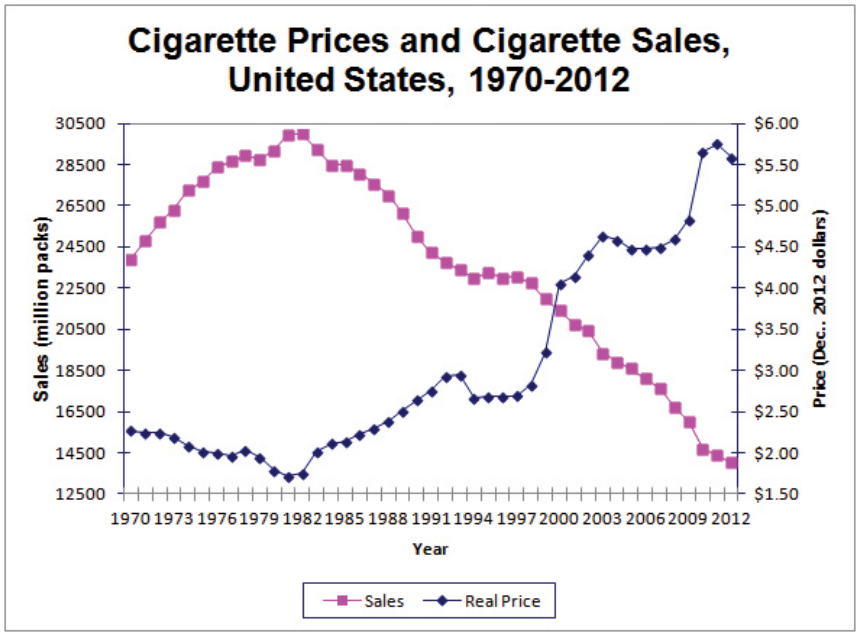


Figure 4. U.S. cigarette prices vs. consumption, 1970-2012

Source: Chaloupka F, Hugan J. Cigarette prices and cigarette sales, United States, 1970-2012. Chicago: University of Illinois at Chicago, Health Policy Center; 2014. Used with permission.

In Mexico, an increase in the tax on tobacco is estimated to have reduced consumption by 5 percent, and the relationship of taxes to consumption looks very similar for the United States, Mexico, and South Africa. Several countries have taken at least some of the proceeds from increased taxes to support public health programs or smoking cessation programs. These are not specifically focused on youth, but youth are more price-sensitive and so the impact is greater for them. Approaches that make it less acceptable or convenient to smoke can have the same overall effect. There are some useful approaches to explore, and this is an ideal time to do so when there has been a shift in marketing by tobacco companies away from the wealthier countries and to the developing world.[45]

Primary prevention is often located in the broader social network, rather than just focused on individual behaviors. In the case of levels of physical activity, communities (which includes schools

and religious organizations, as well as families) can ensure support for activities such as sports, dance, safe transportation, bike paths, and other opportunities for youth to be active. Mexico and other countries periodically close off major roads to car traffic to encourage recreational biking.[46,47] Several African countries, such as Ethiopia, have supported running clubs for youth. The success of Kenyan and Ethiopian runners in the New York and Boston marathons from 2000 to the present speaks to the feasibility and acceptability of running as an activity for youth.

While it is important to build positive food choices at the individual level, there are also activities at the community level that can be helpful. Clearly, schools have an opportunity to increase exposure to and support for healthy food choices. Broader campaigns, such as the "Less Salt, More Life" campaign in Argentina,[48] is an agreement between the government and 25 companies to reduce the salt content of packaged foods.[49] One study that looked at the impact of taxing sugar content had some encouraging findings.[50] John Jemmot and colleagues have tested a rigorous study of a cognitive-behavioral approach to influencing diet and physical activity in South Africa showing that such an intervention was effective.[51]

What We Know and What We Need to Know

There are useful data already available that allow us to anchor our understanding of the issues of youth and risks for later non-communicable diseases. These are a product of far-sighted surveys that include a variety of measures of youth health, health risks, and related social factors. Agencies with a global view of these issues, such as WHO, CDC, and USAID, are providing assistance for youth-oriented data collection activities that provide generally comparable data for a number of countries and where the resulting data are made available for researchers and policymakers in local, national, and global settings.

A first step to advancing our global understanding of these issues would be for more countries to participate in the surveys and, for those who do participate, to repeat the measures at regular intervals. This would facilitate temporal study as well as enhance the ability to do cross-national studies. Of course, the drawbacks of such global studies is that they may miss issues of local relevance. However, it is possible to supplement core data collection with measures that are especially meaningful in specific settings.

In this chapter, I have used a single indicator for each risk factor, each one representing judgment about the value of that single measure. However, it is clear that there are different ways to measure each risk factor. Researchers could develop and test other measures. For example, would a measure of binge drinking be more useful in understanding the risks for NCDs? Overall, one would hope that future research would add to our understanding of the benefits of specific measures, the development of new measures, and greater understanding of how to get the best information from questions about risk factors. A recent study pointed out that people overestimate how much they are exercising. Self-report measures could probably be enhanced by taking advantage of such research. In this case, providing a definition of vigorous exercise (unable to carry on a conversation) could improve reporting.[52] Youth surveys could ask about the intensity of activity, not just the time spent doing the activity.

The issues around measuring physical activity go further into just what counts as physical activity. Most surveys ask about a simple, overall measure, such as "over the past 7 days, in how many were you physically active for at least 60 minutes a day"?[53] However, some surveys break up such questions, most often into three parts. In the STEPS survey in Colombia, the question has three parts and asks about physical activity related to work, transportation (e.g. walking or biking to school or work), and planned physical activity (e.g. exercise).[54] Results reveal gender differences that are not so surprising, since boys and men typically are involved in more vigorous work. But, are all forms of physical activity the same? One view is that only activity that is being structured as "exercise" should be considered. However, other research in the United States found that people who defined their activity as "exercise" were prone to reward themselves for doing it, and that rewards could undermine the benefits of the exercise if the rewards include foods high in sugar or fat.[55]

Measures of smoking and alcohol consumption are already quite varied and may include age of first experience, frequency, intensity, and consistency. While clearly one would like to avoid labeling as a "user" someone who had a single experience with alcohol or tobacco, but what would be the most meaningful measure that is plausible to include in broad-based youth surveys? Clearly there are different types of tobacco use (smoked, chewed, etc) that can be distinguished in data collection, but how much detail adds to our understanding of whether youth are beginning a potentially dangerous habit? Do youth distinguish between alcohol that is home-brewed from that commercially produced? In terms of alcohol use, can we extend research on how the circumstances of use—at home or with friends—affect future use?

Body mass index (BMI) seems to be a straightforward measure. However, measuring BMI among youth means that the measures are taken at different stages of pubertal development. Does that affect the meaning of those measures? Higher fat in early adolescence may fuel puberty among girls at least,[56] and so perhaps in thinking about the linkages of youth risks for NCDs and sexual and reproductive health, the assessments should start earlier than usual for adolescents. Surveys typically include a direct measure of weight and height from which to calculate BMI, are there other simpler techniques that would still be robust?

Improvements in data collection activities need to be based on knowledge about a behavior to obtain the most meaningful measures and important correlates of the behavior. Also, there is a need to develop measures that are valid and useful in survey settings. Much can be done in highly controlled or laboratory settings, but there is a value to measures that can be widely adopted in data collection efforts to allow for studies of differentials within and among countries and over time. In addition to data collected specifically on NCD risk behaviors, there is also a benefit from broader data initiatives that help set the context for young lives. Surveys of education, labor force participation, and family and household functioning could also be useful.

At the most basic level, vital statistics systems are a valuable adjunct to the work of understanding youth and NCDs. Birth registration is a key step to ensuring that young people know how old they are and can document it if needed for participation in school, work, or programs. Death registration, including cause of death reporting, is important for documenting the ultimate toll of NCDs, and for youth, it can also help to document the impact of causes of death, such as accidents, injuries, and suicide among youth.

Research Opportunities

Closely linked to the data issues is the research potential of using those data. Currently available data allow for many analyses within countries and comparisons across countries. A valuable addition to this dialogue would be research to model the impact of the present risk behaviors, the apparent trends, and the impact of changes in those anticipated trends. This modeling could then be used to estimate the likely burden of either not taking any action or the benefits of different interventions to lower risks. Since the cost of treating resulting diseases may vary country-by-country, it would be valuable to have this research conducted at the national level. National level estimates would likely be more useful to policymakers and those who are developing interventions.

The National Institutes of Health (NIH) has already supported many individual projects that deal with elements of youth risks for NCDs and called attention to the need to address the underlying factors that are involved in establishing and maintaining positive health behaviors in children and adolescents. See Table 1 for some examples of NIH program announcements calling for research on youth risks for NCDs. Initiatives by the National Institute on Alcohol Abuse and Alcoholism (NIAAA) address not only individual behaviors but the role of advertising in influencing drinking by young adolescents. Research on obesity pursues the ties to specific diseases and assesses the impact of broad social initiatives in influencing behavior. An innovative announcement specifically sought to improve the understanding of policy changes by facilitating research on the natural experiments that emerge through policy or program changes. There is a wealth of research on tobacco use and growing interest in how the built environment influences activity levels. The landmark Add Health Study[57] has shown the power of a large, representative sample of youth surveyed over time, as it has added to our understanding of how adolescent behaviors influence later health.

Table 1. Examples of behavioral health research funded by the National Institutes of Health

Program Announcement (PA) and Funding Unit	Topic Area/Focus	Example/Brief Description
PA 13-098, 99, 100 NICHD, NHLBI, OBSSR http://grants.nih.gov/grants/guide/pa-files/PA-13-098.html	School nutrition and physical activity policies, obseogenic behaviors, and weight behaviors	Effects of school physical activity and nutrition policies on youths' obseogenic behaviors
PAR 12-198 NIDDK, NCI, NHLBI, NIAAA, NICHD, NINR, ODS http://grants.nih.gov/grants/guide/pa-files/PAR-12-198.html	Improving diet and physical activity assessment	Refining and testing of methods of diet or physical activity assessments; developing or refining innovative methods to improve respondent self-report

Program Announcement (PA) and Funding Unit	Topic Area/Focus	Example/Brief Description
PA 14-038 NIDA, NIAAA http://grants.nih.gov/grants/guide/pa-files/PA-14-038.html	Sex and gender differences in drug and alcohol abuse and dependence	Assessing the effectiveness of gender-based prevention services and the factors that affect their availability, adoption, adaptation, implementation, sustainability, cost-benefit, and cost-effectiveness; addressing and comparing stages of the life cycle from preconception, through adolescence, into adulthood
PA 13-262, 26 NIAAA http://grants.nih.gov/grants/guide/pa-files/PA-13-262.html	Implications of new digital media use for underage drinking and drinking-related behaviors and prevention research	Understanding social media-related underage drinking, drinking-related norms and expectancies, and drinking-related problems
PA 13-191, 192, 242 NIAAA	Structural interventions among alcohol use and risk of HIV/AIDS	Intermediate-level factors that directly affect the drinking environment, retail prices, restrictions placed on alcohol
PAR 11-314 OBSSR, NCI, NHLBI, NIA, NIAAA, NIBIB, NIDCR, NICHD, NIEHA, NIGMS, NIMH, NINR http://grants.nih.gov/grants/guide/pa-files/PAR-11-314.html	Systems science and health in the behavioral sciences	Problems related to health behaviors that seem to cluster together—e.g. tobacco use and other risk behaviors, poor diet and nutrition with physical inactivity—and how these associations undermine effective interventions, programs, or policies
PA 11-087 NIAAA, OBSSR http://grants.nih.gov/grants/guide/pa-files/PA-11-087.html	Research on alcohol-related public policies, such as those detailed in the alcohol policies information system	Effects and effectiveness of alcohol-related public policies on underage drinking, taxation

Key: NCI = National Cancer Institute; NHLBI = National Heart, Lung, and Blood Institute; NIAAA = National Institute on Alcohol Abuse and Alcoholism; NIBIB = National Institute of Biomedical Imaging and Bioengineering; NIDCR = National Institute of Dental and Craniofacial Research; NICHD = National Institute of Child Health and Human Development; NIDDK = National Institute of Diabetes and Digestive and Kidney Diseases; National Institute on Drug Abuse; NIEHS = National Institute of Environmental Health Sciences; NIGMS = National Institute of General Medical Sciences; NIMH = National Institute of Mental Health; National Institute of Nursing Research; OBSSR = Office of Behavior and Social Science Research; ODS = Office of Dietary Supplements.

There is a body of work on youth risks for NCDs that can be a platform for relevant research and program or policy development in low- and middle-income countries, but such work also needs to be developed and adapted to fit the circumstances faced by youth in those settings. In combination, such work could provide a guide for how a society could foster the development of a broad set of positive health behaviors among youth and lower the future risk and burden of NCDs. Youth in developing countries are often facing rapid urbanization that may be disruptive to family and social supports. Urban areas may make an active lifestyle and healthy eating more difficult, but the density

found there may present opportunities for interventions. Urban living may simply be a current example of how researchers could take context into account in studying these health behaviors.

There is an opportunity for research to consider the four risk behaviors together to further our understanding of the collective impact that they have on future NCDs. Modeling the impact of changes in risk behavior in relation to later expected disease burden could advance our understanding of the likely impact of interventions or the cost of neglect.

In the 50 years since the U.S. Surgeon General's report[26] on the health risks of tobacco, there have been many interventions to address tobacco use and many tools to help spread awareness of those tools (e.g. Campaign for Tobacco Free Kids).[58] There is an opportunity now to ensure that those tools are widely available and culturally appropriate for the millions of young people in developing countries. One intervention that has spread is the increase in the cost of tobacco products through taxation.[41] Increased costs of tobacco, lower social tolerance of smoking, and a broader awareness of the health risks associated with smoking are all programmatic approaches that can be both broadly applied (tax increases apply to all) but have greater impact on youth. The large and growing portfolio of interventions could allow the developing world to move more quickly on addressing youth smoking than was the case in higher income countries.

Interventions to address youth alcohol consumption can include taxation, since increasing the price of alcohol tends to lower youth consumption and also provides a funding stream for other health interventions. Laws to prohibit the provision of alcohol to youth can help address underage drinking. Of course, the enforcement of such laws is critical. Developing parental, family, and community understanding of the risks of underage and excessive drinking may be a way to support enforcement. There are growing examples of campaigns to address problematic drinking, such as those focused on drinking and driving, that provide concrete tools (e.g. designated drivers, taxi services after partying) to help mitigate the risks of excessive alcohol consumption and provide a reminder of the inherent risks.

A great deal of attention has been given to the rise in overweight and obesity, and there have been a plethora of approaches to deal with it after the fact. This is not the place to review the myriad tools to help people lose weight; instead we need to ask how we can support youth to establish healthy eating patterns and maintain them into adulthood. Some social interventions, such as taxes on sodas or required labeling of foods and restaurant meals, have been tried with mixed results in terms of support for them and evidence of their effectiveness. A cross-national study[59] found a significant relationship between soft drink consumption and overweight and obesity, as well as a relationship to the prevalence of diabetes. Further research could sharpen our understanding of how soft drinks come to replace healthier options in different settings. Such research would have to confront the role of commercial interests in the transition in food choices in developing countries. Schools offer one setting to provide both nutritious food and education about the importance of healthy eating.

Low income urban areas, typically slums or informal settlements, may become food deserts where there is limited access to a wide range of nutritious, fresh foods. Policies that limit unhealthy fats, salt, and sugar in prepared foods or in restaurant foods are gaining traction and can be tested in a variety of settings with different forms of implementation.

It is unusual to see small children who are not physically active, but levels of activity seem to decline during adolescence. As young people begin to live lives filled with education, work, and family responsibilities, physical activity may suffer. That follows with certain types of employment and typically with urban living. Recent research has shed light on the linkage of physical activity and diet and shown that when physical activity is labeled as being "exercise" it is less effective than physical activity that just appears to be enjoyable, even when it is the same activity.[55] Individuals who saw the activity as "exercise" were more likely to reward themselves with treats, which were often ones that undermined healthy eating. Findings such as these could provide guidance in the ways interventions are developed and marketed for adolescents.

Policy Implications

In addition to numerous data and research opportunities, there are also opportunities at the policy level to address youth risks for later NCDs. Countries facing this looming shift in burden of disease could make an "adolescent health report card" to guide their actions. This could ask basic questions about the steps that were being taken to address each of the risk factors. For example, What steps are being taken to reduce youth smoking—taxation, health education, enforcement of laws about selling tobacco to minors? Are there laws about selling alcohol to minors, and are they enforced? Are there health education programs about the risks of early and extreme drinking? Do they include tools for youth to use to offset peer pressure? Are parents provided tools to help them identify problems and manage alcohol and tobacco use among their children? What steps are schools taking to support healthy eating and physical activity? Is there an infrastructure to support physical activity? In urban areas, are there ways to access healthy foods? Does the tax structure support unhealthy foods? Are economic development tools used to advance healthy eating and physical activity?

While the specific approaches will vary with different settings, a policy approach can be one that is implemented at different levels. Action at the national level would be needed for taxation approaches, but the enforcement of laws typically happens at the community level. Where there is widespread understanding of the risks of underage drinking and smoking, enforcement may be more rigorous. Community, cultural, and religious institutions can sponsor activities that support physical activity, such as sports. In most developing countries the health sector is already burdened, but there are NCD primary prevention activities that lie outside the health sector. That creates an opportunity to have more players who are challenged to contribute to primary prevention. The World Health Organization has articulated an essential package of "Best Buys" for population-based interventions.[60]

NIH has long supported research on health behaviors and is well-positioned to contribute to understanding the threats faced by the developing world in terms of NCDs. A review of the initiatives NIH is promoting shows that many could enlighten us about how youth form, establish, and maintain heathy behaviors. There is an opportunity to specifically focus on the policy issues that have been raised about tobacco, alcohol, nutrition, and physical activity, and there are structural interventions that may address alcohol use and HIV/AIDS and, at the same time, enlighten us about structural interventions more generally. There is even an expressed interest in improving the measures used for diet and physical activity. Other initiatives call up the role of schools, social, and other media and highlight gender/ethnic issues, setting the stage for this work to focus on adolescents and include the considerations of these behaviors in developing countries. In other words, NIH has already

established a framework of interest in important key issues related to the risk behaviors among youth in developing countries. Signaling to the investigator and reviewer communities that work in such settings is desired would help mobilize current initiatives to advance work in this area (see Table 1 for examples of these NIH-supported initiatives). This existing background work and interest could be directed to include relevant studies that address adolescent risk behaviors in developing countries. Research in non-U.S. settings often provides an opportunity to not only advance understanding of important health issues among our global partners, but it also may allow for the testing of hypotheses, developing novel measures, and testing of different types of interventions.

Conclusion

In sum, the developing world is facing an epidemiologic transition that makes non-communicable diseases the most prominent cause of death. Despite ongoing needs to address infectious disease and malnutrition, we also must pay attention now to the morbidity and mortality associated with heart disease, diabetes, cancer, and chronic respiratory disease. In addition to the need to screen, diagnose, and treat those with disease, we also need to articulate a primary prevention approach that can lower the future burden of NCDs on individuals, families, communities, and nations. Two of the primary behavioral risk factors—use of tobacco and alcohol—begin during adolescence. To promote a healthy diet and active lifestyle in later years, adolescence is an important time to solidify positive health behaviors that will, among other things, lower the risk from overweight and obesity.

Basic data are available about youth risk behaviors in the developing world, but vigilance is needed to enhance those data systems. Research on these risk factors is more heavily concentrated in settings where the burden of NCDs has been felt for longer; there is a need to extend this research attention to youth in developing countries. The policy community has an opportunity to take the available data, research findings, and experiences with interventions to make youth prevention of NCD risks a key, measurable part of their overall planning. Benefits would be felt not only in the health sector, but also more broadly, since rising levels of NCDs threaten economic productivity.

Acknowledgments

Special thanks go to Toshiko Kaneda of the Population Reference Bureau who has participated in the development of data on youth in Latin America and the Caribbean. The opinions presented herein are those of the author and do not necessarily represent the position of the Agency for Healthcare Research and Quality, the National Institutes of Health, or the U.S. Department of Health and Human Services.

Author's Affiliation

Wendy Baldwin, PhD, Former President and CEO, Population Reference Bureau, Washington, DC.

Address correspondence to: Wendy Baldwin, PhD, 1104 Mourning Dove Drive, Blacksburg, VA 24060; email wendybburg@gmail.com.

References

1. Omram AR. The epidemiologic transition: a theory of the epidemiology of population change. Milbank Q 1971;40:509-38.
2. United Nations high-level meeting on noncommunicable disease prevention and control: NCD summit to shape the international agenda. September 19-20, 2011, New York City. Available at www.who.int/nmh/events/un_ncd_summit2011/en/. Accessed January 12, 2015.
3. Global action plan for the prevention and control of NCDs, 2013-2020. Geneva: World Health Organization; 2013. Available at http://www.who.int/nmh/publications/ncd-action-plan/en/. Accessed January 12, 2015.
4. 2012 World population data sheet. Washington, DC: Population Reference Bureau; 2012. Available at http://www.prb.org/Publications/Datasheets/2012/2012-world-population-data-sheet.aspx. Accessed January 12, 2015.
5. Noncommunicable diseases country profiles 2011: WHO global report. Geneva: World Health Organization; 2011. Available at http://www.who.int/nmh/publications/ncd_profiles2011/en/. Accessed January 12, 2015.
6. Bloom DE, Cafiero E, Jané-Llopis E, et al. The global economic burden of noncommunicable diseases. PGDA working paper No. 87. Boston, MA: Harvard Institute for Global Health, Program on the Global Demography of Aging; 2012. Available at http://www.hsph harvard.edu/program-on-the-global-demography-of-aging/WorkingPapers/2012/PGDA_WP_87.pdf. Accessed January 12, 2015.
7. Scaling up action against NCDs: How much will it cost? Geneva: World Health Organization; 2011. Available at http://www.who.int/nmh/publications/cost_of_inaction/en/. Accessed January 20, 2015.
8. Malik K. Sustaining human progress: reducing vulnerabilities and building resilience. UNDP human development report, 2014. New York: United Nations Development Program; 2014. Available at http://www.undp.org/content/undp/en/home/presscenter/events/2014/july/HDR2014 html. Accessed January 12, 2015.
9. Chan M. Report by the Director-General to the special session of the Executive Board on Ebola. Geneva: World Health Organization; 2015. Available at http://www.who.int/dg/speeches/2015/executive-board-ebola/en/. Accessed March 12, 2015.
10. Lozano R, Naghavi M, Foreman K, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012;380(9859):2095-128.
11. Nikolic IA, Stanciole AE, Zaydman M. Chronic emergency: why NCDs matter. Washington, DC: World Bank; 2011. Available at http://tinyurl.com/knvnbph. Accessed March 12, 2015.
12. Gupta DK, Shah P, Misra A, et al. Secular trends in prevalence of overweight and obesity from 2006 to 2009 in urban Asian Indian adolescents aged 14-17 years. PLoS One 2011;6(2):e17221.
13. Reddy SP, Resnicow K, James S,. Rapid increases in overweight and obesity among South African adolescents: comparison of data from the South African National Youth Risk Behaviour Survey in 2002 and 2008. Am J Public Health 2012;102(2):262-8.
14. Noncommunicable diseases among older adults in low- and middle-income countries. Population Reference Bureau, Today's Research on Aging, Issue 26; 2012, August. Available at www.prb.org/pdf12/TodaysResearchAging26.pdf. Accessed January 13, 2015.
15. Obesity: preventing and managing the global epidemic. Report of a WHO consultation, WHO technical report series 894. Geneva: World Health Organization; 2000.
16. Viner RM, Coffey C, Mathers C, et al. 50-year mortality trends in children and young people: a study of 50 low-income, middle-income, and high-income countries. Lancet 2011;377(9772):1162-74.
17. Cawley J, MacLean JC. Unfit for service: the implications of rising obesity for the U.S. military. Health Econ 2012;21(11):1348-66.
18. Global Youth Tobacco Survey. Atlanta, GA: Centers for Disease Control and Prevention. Available at http://www.cdc.gov/tobacco/global/gtss/tobacco_atlas/pdfs/part3.pdf. Accessed January 13, 2015.
19. Global tobacco control. Atlanta, GA: Centers for Disease Prevention and Control, Office on Smoking and Health, National Center for Chronic Disease Prevention and Health Promotion; 2013. Available at http://www.cdc.gov/tobacco/global/index htm. Accessed January 13, 2015.
20. Global School-Based Student Health Survey. World Health Organization and Centers for Disease Control and Prevention. Available at http://www.cdc.gov/gshs/. Accessed January 13, 2015.
21. The DHS Program: Demographic and Health Surveys. Rockville, MD: USAID. Available at www.dhsprogram.com. Accessed January 13, 2015.

22. Millennium Development Goals and Beyond 2015. Goal 2: Achieve universal primary education. United Nations. Available at http://www.un.org/millenniumgoals/education.shtml. Accessed January 13, 2015.
23. STEPwise approach to chronic disease risk factor surveillance. 2010 STEPS survey report. Geneva: World Health Organization; 2010. Available (Spanish only) at http://www.who.int/chp/steps/2010_STEPS_Survey_Colombia.pdf. Accessed March 12, 2015.
24. Baldwin W, Kaneda T, Amato L. Noncommunicable diseases in Latin America and the Caribbean: Youth are key to prevention. Washington, DC: Population Reference Bureau; June 2013. Available at http://www.prb.org/pdf13/TechnicalNotes.pdf. Accessed January 13, 2015.
25. Taioli E, Wynder EL. Effect of the age at which smoking begins on frequency of smoking in adulthood. N Engl J Med 1991;325(13):968-9.
26. Reducing the health consequences of smoking: 25 years of progress. A report of the Surgeon General. Rockville, MD: U.S. Department of Health and Human Services; 1989. Available at http://profiles nlm nih.gov/ps/access/NNBBXS.pdf. Accessed January 14, 2015.
27. You're the target: global Marlboro campaign found to target teens. Washington, DC: Campaign for Tobacco-Free Kids; 2014. Available at http://global.tobaccofreekids.org/en/industry_watch/marketing/youre_the_target. Accessed January 14, 2015.
28. Tobacco product use among middle and high school students-United States, 2011 and 2012. MMWR 2013 Nov 15;62(45):893-7.
29. Global Tobacco Surveillance System Data (GTSSData). Atlanta GA: Centers for Disease Control and Prevention; 2010. Available at http://nccd.cdc.gov/GTSSData/default/default.aspx. Accessed January 14, 2015.
30. Global Youth Tobacco Survey Collaborative Group. Tobacco use among youth: a cross country comparison. Tob Control 2002;11(3):252-70.
31. Alcohol, violence, and aggression. Alcohol Alert from NIAAA, No. 38, October 1997. Rockville, MD: National Institute on Alcohol Abuse and Alcoholism. Available at http://pubs niaaa nih.gov/publications/aa38 htm. Accessed January 14, 2015.
32. Grant BF. Impact of a family history of alcoholism on the relationship between age at onset of alcohol use and DSM-IV alcohol dependence. NIAAA Epidemiological Bulletin 1998;22(2):144-8.
33. Grant BF, Dawson DA. Age at onset of alcohol use and its association with DSM-IV alcohol abuse and dependence: results from the National Longitudinal Alcohol Epidemiologic Survey. J Subst Abuse 1997;9:103-10.
34. Kaneda T, Naik R, Baldwin W. Noncommunicable diseases—risk factors among young people in Africa. Washington, DC: Population Reference Bureau; 2014. Available at http://www.prb.org/Publications/Reports/2014/ncd-risk-youth-africa.aspx. Accessed January 14, 2015.
35. Guatemala (Capital City Oficial): 2009 Fact Sheet. Global School-based Student Health Survey. Available at http://www.who.int/chp/gshs/Guatemala_capital_city_oficial_2009_FS.pdf. Accessed January 14, 2015.
36. Global Health Observatory data: obesity among women. Geneva: World Health Organization; 2011. Available at http://www.who.int/gho/urban_health/risk_factors/women_obesity/en/. Accessed March 12, 2015.
37. Kolcic I. Double burden of malnutrition: a silent driver of double burden of disease in low- and middle-income countries. J Glob Health 2012;2(2):020303.
38. The state of food and agriculture. Rome: Food and Agriculture Organization of the United Nations; 2013. Available at http://www fao.org/docrep/018/i3300e/i3300e.pdf. Accessed March 12, 2015.
39. OECD obesity update, 2014. Victoria, BC, Canada: Jesse Ferreras; 2014.
40. Global recommendations on physical activity for health. Geneva: World Health Organization; 2010. Available at http://whqlibdoc.who.int/publications/2010/9789241599979_eng.pdf?ua=1. Accessed January 14, 2015.
41. Raising cigarette taxes reduces smoking, especially among kids (and the cigarette companies know it). Fact sheet. Washington, DC: Campaign for Tobacco-Free Kids; 2012.
42. Holford TR, Mexa R, Warner KE, et al. Tobacco control and the reduction in smoking-related premature deaths in the United States, 1964-2012. JAMA 2014;311(2):164-71.
43. Ng M, Freeman MK, Fleming TD, et al. Smoking prevalence and cigarette consumption in 187 countries, 1980-2012. JAMA 2014;311(2):183-92.
44. Youth and tobacco use. Fact sheets. Atlanta, GA: Centers for Disease Control and Prevention; 2014. Available at http://www.cdc.gov/tobacco/data_statistics/fact_sheets/youth_data/tobacco_use/index.htm. Accessed March 12, 2015.
45. Bornhaeuser A, Bloom J. Smokefree air law enforcement: lessons from the field. Washington, DC: Global Smokefree Partnership; 2009.
46. Ciclovias – the 20 longest cycling streets in the world. City Clock Magazine (online); 2014. Available at http://www.cityclock.org/ciclovias-20-longest-cycling-streets-world/#.VM2YO9LF90Y. Accessed March 12, 2015.
47. Ciclovis recreativisde las Americas (Web site). Available at http://www.cicloviasrecreativas.org/en/map. Accessed March 12, 2015.

48. Argentine initiative to reduce salt consumption. "Less salt, more life." Washington, DC: Pan American Health Organization. Available at http://www.paho.org/panamericanforum/wp-content/uploads/2012/08/less-salt-more-life_PAHO-consortium_ARG.pdf. Accessed January 15, 2015.
49. Legetic B. America's initiative for cardiovascular disease prevention through dietary salt reduction. Washington, DC: Pan American Health Organization; 2012. Available at http://www.iom.edu/~/media/Files/Activity%20Files/Nutrition/ConsequencesSodiumReduction/2012-DEC-04/Presentations/06_Branka%20Legetic.pdf. Accessed January 15, 2015.
50. Cabrera Escobar M, Veerman JL, Tollman SM, et al. Evidence that a tax on sugar sweetened beverages reduces the obesity rate: a meta-analysis. BMC Public Health 2013;13(1):1072.
51. Jemmott JB, Jemmott LS, O'Leary A, et al. Cognitive-behavioural health-promotion intervention increases fruit and vegetable consumption and physical activity among South African adolescents: a cluster randomized controlled trial. Psychol Health 2011;26(2):167-85.
52. Canning KL, Brown RE, Jamnik VK, et al. Individuals underestimate moderate and vigorous intensity physical activity. PLoS One 2014;9(5):e97927.
53. Global School-based Student Health Survey: 2009 Guatemala GSHS questionnaire. Atlanta, GA: Centers for Disease Control and Prevention; 2009. Available at http://www.who.int/chp/gshs/2009_Guatemala_GSHS_Questionnaire.pdf. Accessed January 15, 2015.
54. STEPwise approach to chronic disease risk factor surveillance, Colombia. Geneva: World Health Organization; 2010. Available at http://www.who.int/chp/steps/colombia/en/. Accessed January 15, 2015.
55. Werle CO, Wansink B, Payne CR. (2014). Is it fun or exercise? The framing of physical activity biases subsequent snacking. Marketing Letters 2014;1-12. Available at http://tinyurl.com/njdqw2e. Accessed January 19, 2015.
56. Kaplowitz PB. Link between body fat and the timing of puberty. Pediatrics 2008;121(Suppl 3):S208-17.
57. Add Health. The National Longitudinal Study of Adolescent to Adult Health. University of North Carolina Population Center. Available at http://www.cpc.unc.edu/projects/addhealth. Accessed January 19, 2015.
58. Campaign for Tobacco-Free Kids; Web site. Washington, DC; 2014. Available at http://www.tobaccofreekids.org/. Accessed January 19, 2015.
59. Basu S, McKee M, Galea G, et al. Relationship of soft drink consumption to global overweight, obesity, and diabetes: a cross-national analysis of 75 countries. Am J Public Health 2013;103(11):2071-7.
60. From burden to "best buys": reducing the economic impact of non-communicable diseases in low- and middle-income countries. Geneva: World Health Organization; 2011. Available at http://www.who.int/nmh/publications/best_buys_summary.pdf?ua=1. Accessed January 19, 2015.

Wendy Baldwin, PhD, is a social demographer with a background that spans government, academic, and private-sector organizations. At the National Institute of Child Health and Human Development, she led a program to advance research on adolescent sexual and reproductive health, the role of child care in fertility and employment decisions, behavioral aspects of HIV/AIDS, and fertility in the United States. As the Deputy Director for Extramural Research, she advanced policies to support the sharing of research data and streamlined peer review and the electronic submission and processing of grants. At the Population Council in New York City, she led a program of research on poverty, gender, and youth in international settings that emphasized girls' schooling, contraception and childbearing, and access to demographic data and its use by policymakers. At the Population Reference Bureau, Dr. Baldwin focused on developing data about youth risks for non-communicable diseases in developing countries and communicating with policymakers. Now retired, she maintains her professional interests and activities from her home in Blacksburg, VA.

Identifying the Principal Factors Responsible for Improvements in the Health of Populations

By Samuel H. Preston

Abstract

Improvements in health and length of life for the average person have been among the greatest of human achievements. Global life expectancy at birth has risen from about 31 years in 1900 to 70 years today. In this chapter, I describe some important contributions to understanding the principal factors driving improvements in health since the mid-1800s. While this chapter does not represent a comprehensive account of the relevant literature, it does provide some examples and discussion of useful analytic studies that examine the progress that has been made over the past 150-plus years toward improving health and life expectancy, both globally and here in the United States.

Introduction

In the United States, life expectancy has increased from about 48 years in 1900 to 79 years today.[1,2] Nordhaus[3] estimated that the imputed value of gains in life expectancy in the United States between 1900 and 1995 was approximately equal to the entire gain in per capita income over the period. If the value of improvements in morbidity were factored into the assessment, the value of health gains would have exceeded the value of increased consumption of all other goods and services combined (see also Murphy and Topel[4]).

Social scientists have been the key contributors to understanding the broad social forces that have driven these improvements in health. Their central role in these discussions is a product of their concern with accurate measurement, their attention to issues of research design, and their emphasis on understanding population-level phenomena. While medical sciences have focused on molecules, micro-organisms, genes, and physiology, social scientists have attempted to assess what advances in the broad realms of personal income, medicine, and public health have meant for levels of population health. The results have helped to calibrate public investments in health and in other social sectors.

Most studies of these relations have used indicators of mortality. The main reason for using these measures is that they are often available from registries of death that usually span long periods of time and include nearly 100 percent population coverage. Furthermore, death is an unambiguous event whereas measures of self-assessed health and of disease incidence and prevalence are subject to many forms of error and bias. Costa[5] demonstrates for the United States that secular improvements in longevity have been accompanied by huge reductions in the prevalence of major chronic conditions.

The subject of this review is massive, and what follows should be considered a set of illustrations of useful analytic studies rather than a comprehensive account of the relevant literature.

Selected Case Studies

England and Wales

The first serious effort to identify major factors responsible for improvements in longevity was made by a medical historian, Thomas McKeown. McKeown took advantage of the longest series of vital statistics on causes of death for a national population, that pertaining to England and Wales, which dates back to 1838. A series of journal articles culminated in a 1976 monograph.[6] McKeown showed that the mortality decline was primarily attributable to declines in infectious and parasitic diseases and that, for disease after disease, the bulk of the decline had occurred before any medicine or therapy was available to combat the disease. Tuberculosis, the single most important disease, had declined by some 80 percent before effective medical therapy was available. McKeown argued that, if improvements in medical treatment were not responsible for the declines in infectious and parasitic diseases, then improvements in standards of living must be responsible. In particular, he attributed the bulk of the mortality decline to improved nutrition. McKeown presented no direct evidence about nutrition's role, and his process-of-elimination reasoning failed to eliminate some obvious alternatives, particularly improvements in preventive public and personal health practices.

Especially important elaborations of the part played by public health initiatives in England have been provided by Szreter[7] and Woods.[8] In a thorough and well-considered monograph, Mercer[9] assembles a great deal of evidence about the importance of public and personal health practices in the mortality decline in England and Wales. McKeown did make one very important point that has for the most part stood the test of time: specific therapeutic medical treatments have played a minor role in reductions in infectious disease mortality in now-developed countries.[6]

United States

The role of diet appears to be modest at best in the 20th century mortality decline in the United States. The United States was already a very well-fed country by 1900, when life expectancy at birth was only 48 years; dietary reconstructions from direct inquiries suggest that the average daily caloric consumption per adult equivalent was about 3,700, higher than today.[10] Demographers Ewbank and Preston[11] argue that the essential element in U.S. gains in mortality between 1900 and 1930 was an enormous scientific breakthrough—the germ theory of disease. This theory was empirically validated in the 1880s and was beginning to displace the misguided miasma theories by the turn of the century. The development and diffusion of the theory and its practical technologies is a plausible explanation of why the frontier of national life expectancy increased faster between 1880 and 1950 than in any equivalent period, as shown by demographers Vallin and Mesle.[12]

Although the new theory led to few practical drugs, as McKeown had argued,[6] it did lead to an entirely new approach to preventive medicine, practiced both by departments of public health and

by individuals. While public health departments had been concerned about the need to improve water supply and sewage disposal processes before the germ theory, the criteria for success were sight and taste and odor, rather than bacteria counts. All this changed rapidly. Enlightened public health officials were quick to recognize how the germ theory should guide their practice.

In an important article, economists Cutler and Miller[13] investigated the relation between changes in municipal mortality and the timing of water supply and sewage improvements in 13 American cities during the early decades of the 20th century. They concluded that 43 percent of the decline in the death rate in these cities between 1900 and 1936 was attributable to improvements in the cleanliness of water. Economist Richard Easterlin[14] examined the mortality decline in the United States between 1880 and 1950 and concluded that the role of public policy was of critical importance. The market could not be counted on to provide pure water or milk, to control pests, or to disseminate knowledge about hygiene and infant care. Market failure in the face of the new bacteriology prompted a massive expansion of departments of public health.

By the time of the first White House Conference on Infant Mortality, held in 1909, public health officials realized that rapid advances in longevity required that they go beyond their normal domain of public works and attempt to change the personal health practices of individuals. The germ theory provided a number of powerful weapons for doing so. These included boiling bottles and milk, washing hands, protecting food from flies, isolating sick children, and ventilating rooms. Public health officials launched massive campaigns to encourage these practices.[11] Cutler, Deaton, and Lleras-Muney[15] usefully draw the distinction between "macro public health" (filtering and chlorinating water supplies, building sanitation systems, draining swamps) and "micro public health," which comprises features of individuals' behavior that were often influenced by governmental efforts.

The Children's Bureau, created in 1912, adopted a primary focus on child health and producted a pamphlet, Infant Care, that became the largest selling publication in the history of the Government Printing Office, with some 12 million copies sold by 1940. By the 1920s, the Bureau was receiving and answering over 100,000 letters a year from parents seeking child care advice. Economic historian Joel Mokyr[16,17] documents the arduous changes that the germ theory introduced into household practices. The new hygienic practices appear to have been most vigorously adopted among professionals. Ewbank and Preston[11] show that, at the turn of the century, the children of physicians had mortality levels that were scarcely better than those of the average child, suggesting that physicians had few weapons at their disposal to advance survival. By 1924, the mortality of physicians' children was 35 percent below the national average. Children of teachers also advanced rapidly, and children of all professionals made great strides during the period.

Sri Lanka

The materials to study mortality decline in developing countries are much thinner than those in the developed world. Very few countries had vital registration data that could properly measure levels and changes in mortality during the period of rapid, often dramatic, improvements, following World

War II. Sri Lanka (Ceylon) is an exception, thanks to a vital registration system established during British colonization.

Sri Lanka achieved a remarkable decrease in mortality after World War II; the crude death rate fell from 21.5 per 1,000 in 1945 to 12.4 in 1950. This period coincided with an insecticide-spraying anti-malarial campaign. The combination of data availability and dramatic change has drawn the attention of a score of social scientists. The most convincing study is by the economist Peter Newman,[18] who examined the regional distribution of mortality declines in comparison to estimates of pre-campaign malarial endemicity. He concluded that the malaria eradication campaign in 1946 reduced Sri Lanka's crude death rate by 4.2/1,000 between 1936-1945 and 1946-1960, representing some 42 percent of the reduction in mortality during the period. Newman's study gained prominence because control of malaria appears to have been the single leading source of mortality decline in developing countries after World War II.[19] In addition to antimalarial programs, immunization and antibiotics are thought to have played an important role in the rapid mortality declines in developing countries after World War II.[20] Unfortunately, the mortality data are simply too weak to draw a very precise quantitative conclusion about their importance.

International Studies

McKeown's claim that improved standards of living were responsible for the bulk of mortality improvements[6] was put to a test by Preston.[21] Using time series and cross-sectional evidence for nations with relatively good data during the 20th century, I demonstrated that a tight cross-sectional relationship existed between national levels of life expectancy and per capita income during the 1900s, the 1930s, and the 1960s. The correlation between the two series (with per capita income logged) was .88 in both the 1930s and the 1960s. Such a tight relationship implied that living standards, as captured in per capita income, were a dominant determinant of mortality.

The curve relating life expectancy to income shifted decisively during the century, however, as shown in Figure 1. If the contribution of gains in per capita income to improvements in life expectancy were effectively captured by the cross-sectional relations, then the gain in life expectancy attributable to gains in per capita income between the 1930s and the 1960s would have been about 1.3-2.5 years. In contrast, worldwide gains in life expectancy were about 12.2 years during the period. So, I concluded that factors exogenous to a country's current level of income probably account for about 85 percent of the growth in life expectancy for the world as a whole between the 1930s and the 1960s.[21] The proportions were similar in industrialized and developing countries.

I have argued that the shift in the curve, and hence the bulk of advance in life expectancy, was primarily attributable to improvements in health technology and its application.[21] The germ theory of disease was prominently featured in the account. However, no direct evidence was presented on the role of technology; it was a residual explanation. Life expectancy and per capita income were featured in the analysis because social scientists had developed good measures of each; no equivalent measure of the quality of health technology had or has yet been developed.

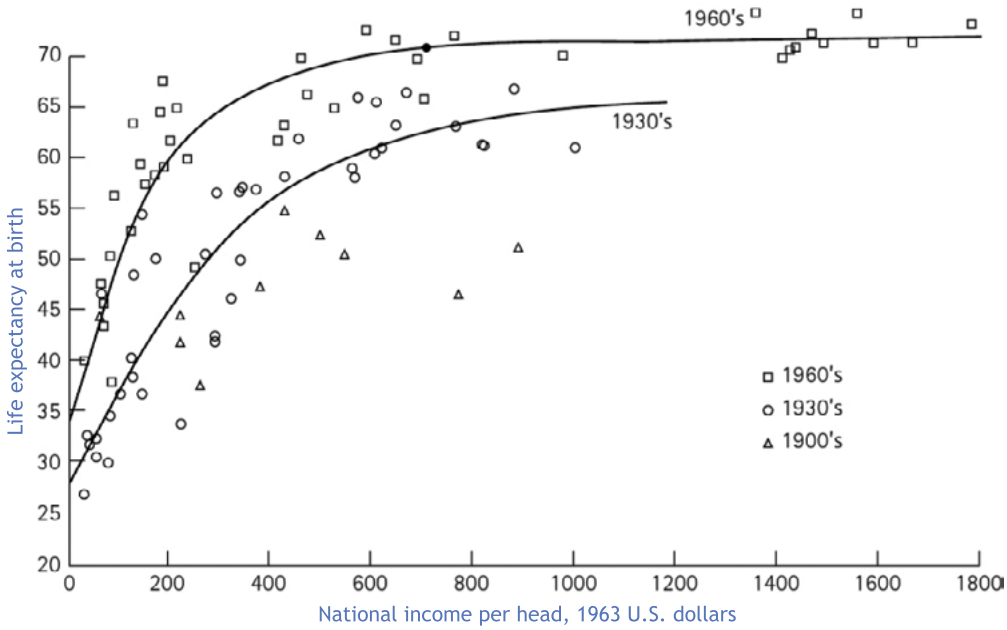


Figure 1. Scatter diagram of relations between life expectancy at birth and national income per head for nations in the 1900s, 1930s, and 1960s.
Source: Preston.[21] Used with permission.

The analytic apparatus in Figure 1—dubbed the Preston Curve by Angus Deaton[22]—has been extensively employed. Maintenance of a strong cross-national relationship between mortality and income and continued shifts in the cross-sectional curve have been demonstrated by the World Bank,[23] Kenny,[24] Pritchett and Viarengo,[25] and Hum et al.[26] Pritchett and Viarengo[24] examine the relationship from 1902 to 2007. They conclude that the slope of the relation (elasticity) has remained quite constant and that "There is no indication the Preston curve is 'breaking down' and no indication from over 100 years of data that a very strong relationship between national income and life expectancy will not persist." They also find that the curve has continued to shift.

One question that has been raised about the curve is the direction of causality that it represents. While richer countries and richer people have the resources needed to extend life, it is also the case that healthier people are more productive. Adding complexity to the question is that lower mortality produces faster population growth, which can produce diminishing returns to labor.[27] A volume addressing the effects of improvements in health on economic productivity concluded that results were "inconclusive."[28,a] The most recent, but surely not the last, paper on this issue concludes that advances in health have not contributed to economic growth.[29] In the absence of compelling evidence to the contrary, it seems reasonable to treat the relation between life expectancy and income as reflecting primarily the effect of income on life expectancy. Confirmation of the causal

[a] It is important to recognize that, in the analyses of the effects of health on economic productivity, the value of health and longevity for their own sake is not included in the value of economic product. If they were to be included, as noted earlier, the impact of health improvements on economic well-being would be enormous.

impact of income on mortality is presented in an important article by economists Prichett and Summers.[30] They used instrumental variables to model the effects of economic advances on infant and child mortality and concluded that country differences in income growth rates over the previous three decades explain roughly 40 percent of the cross-country differences in mortality improvement.

As can be seen in Figure 1, the relationship between income and life expectancy is curvilinear. The curvature suggests that a policy of redistributing income from rich to poor would improve average health outcomes, since the gains in health of those with low incomes would outweigh the losses of those with high incomes. There is little agreement about whether inequality has additional effects on health and mortality beyond those associated with this non-linearity (for opposing views, see Wilkinson and Pickett[31] and Deaton[32]).

Amendments and Extensions

Although the relationship between life expectancy and income has retained its basic shape since the 1960s, there have been several regional departures from the basic form. During the 1990s, the HIV/AIDS epidemic in Sub-Saharan Africa reduced life expectancy by as much as 20 years in the hardest-hit countries. This epidemic, now receding, had the effect of strengthening the relationship between income and life expectancy.[31,33] A second departure occurred in Eastern Europe and the former Soviet Union, where life expectancy has fallen below that implied by levels of income. The most persuasive explanation for the poor performance in this region is excessive consumption of alcohol, especially among males, whose life expectancy has fallen further behind the norm.[34] A third departure is the United States, a wealthy country with a level of life expectancy that has fallen further and further behind its affluent peers since 1980.

As these examples imply, one of the benefits of finding a close association between variables is the lessons that can be drawn from exceptions to the rule. Analyses of reasons for the poor performance of the United States relative to its high level of income have drawn a great deal of attention. It is clear that, above age 50, the history of exceptionally heavy smoking in the United States and its position as a frontrunner in the obesity epidemic are implicated in the country's poor performance.[35] But the poor performance by the United States also extends to ages below 50, including especially high rates of accidental and violent death among young men as well as high levels of infant mortality. In fact, a report demonstrating the pervasiveness of the U.S. disadvantage indicated that the United States had the highest age-specific death rates at all ages below 70 among 17 comparison countries.[36]

The Woolf and Aron report[35] is one of the most heavily cited publications in the history of the National Research Council. It was the subject of nearly 100 newspaper or journal articles and has been downloaded more than 21,000 times. It also has been the subject of congressional hearings, congressional staff briefings, public commentary by members of Congress, and a response from the Department of Health and Human Services (data supplied by press office, National Research Council).

The examination of national exceptions has been extended to developing countries. John Caldwell, an Australian demographer, studied in great detail three populations in developing countries where life expectancy far surpassed what could be expected based on income in those countries: Sri Lanka, Costa Rica, and Kerala state in India.[37] Caldwell concluded that the principal reason for the exceptional performance in these regions was a history of women's empowerment, which manifested itself in women's high levels of schooling, greater authority within the family, and more intensive political participation. Greater political participation in turn helped to generate more egalitarian governance, including great emphasis on public health. Deaton[38] elaborates on the argument that a strong political voice for the poor is an important ingredient in supporting public health programs. Predominantly Muslim countries, with unusually low levels of female autonomy, help to confirm Caldwell's inference by having typical life expectancy levels far below those expected by their levels of income.

Caldwell's research[36] raised the question of whether educational attainment and/or literacy were more important drivers of mortality decline than income. Of course, educational attainment and income are highly correlated, so efforts to separate their effects are vulnerable to measurement error and other sources of imprecision. The addition of literacy or educational attainment to income in cross-national analyses of life expectancy has typically found significant coefficients on both.[19,24] Gakidou et al.[39] argue that increases in women's education over the period 1970-2009 in developing countries could account for 51 percent of the reductions in child mortality over the period. However, the estimated micro-level relation between women's education and child mortality that was used in the macro-level inference is not robust to selectivity biases. The contribution of changes in parental educational attainment to improvements in child mortality remains an important open question. Research designs that focus on exogenous changes in educational requirements, rather than on educational achievements per se, would help to resolve issues of selectivity (e.g., Klemptner et al.[40]).

Developed Countries Since 1960

Aided by sulfa drugs since the 1930s and antibiotics since the 1940s, the bacteriologic revolution had essentially played itself out by 1960 in the developed world. If there were to be any additional gains in life expectancy—and contemporaries were skeptical about the possibilities—such gains would have to come from advances against the chronic diseases associated with older age, especially cardiovascular disease and cancer.

Between 1960 and 2000, age-standardized cardiovascular death rates declined by slightly more than 50 percent in the United States.[15] Cardiovascular disease mortality reductions account for 70 percent of the 7-year increase in life expectancy between 1960 and 2000. Between 2000 and 2009, death rates from cardiovascular disease declined by an additional 31 percent.[41] Other developed countries enjoyed comparable or even faster declines.[42] Cutler[43] matches the results of clinical trials to actual mortality declines and attributes the bulk of the decline in cardiovascular disease mortality—as much as two-thirds of the reduction—to medical advances. Ford et al.[44] look specifically at the 50.2 percent decline in coronary heart disease mortality between 1980 and 2000. They conclude that approximately 47 percent of the decrease can be attributed to improved treatment, including secondary preventive therapies after myocardial infarction or revascularization (11 percent), initial

treatments for acute myocardial infarction or unstable angina (10 percent), treatments for heart failure (9 percent), revascularization for chronic angina (5 percent), and other therapies (12 percent). Approximately 44 percent can be attributed to changes in risk factor distributions, especially reductions in smoking.

The technologies that were developed from the bacteriologic revolution were primarily embodied in public goods, at least during the major mortality declines in developed countries. In contrast, the new health technologies associated with improvements in cardiovascular disease were more like private goods, requiring both individual purchases and, often, skills in self-administration.[5] One might expect them to be more widely diffused and more effective among richer countries and richer people within countries.

Mackenbach and Looman[45] reexamined the Preston curve in a group of European populations. They found that the shift in the curve has declined in magnitude since 1960, and the shift accounted for only one-fourth to one-half of the increase in life expectancy in Europe as a whole since 1960. Most of the gains over the period were attributable to increases in income. Countries with larger increases in income typically had larger gains in life expectancy, although the connection was not statistically significant. The slope of the relation between life expectancy and national income was quite constant between 1960 and 2008. However, the relation between changes in income and declines in cardiovascular mortality was significant in the expected direction.

Within countries, socioeconomic disparities in health and mortality are growing in most places. Kunst et al.[46] found a growing disparity in mortality by income quintile in most European countries, although not in Nordic countries. Glied and Lleras-Muney[47] found that education gradients in mortality in the United States became steeper for causes of death that had experienced greater technological progress. Goldman and Smith[48] found that the health benefits associated with additional schooling in the United States rose over time as measured by self-reported health status. This rise can be attributed to both a growing disparity by education in the probability of having major chronic diseases during middle age as well as better health outcomes for those with each disease.

The growing disparity in health outcomes for those with a disease may in fact be attributable to the greater access of better educated people to superior technology. But their lower prevalence of disease is also partly or largely attributable to better health habits.[49,50] That growing disparities by income class are not inevitable is suggested by the fact that major health improvements can be attained simultaneously with reductions in spending—that is, by smoking less, eating less, and walking instead of driving.

Issues of Timing

Table 1 summarizes a large literature on speed of mortality reduction and the causes responsible for the reduction. It includes studies of the period before 1880, which have not been the focus of this review.[5,15,37,51,52] Since 1945, advances in the practice of medicine have played an important role in both developed and developing countries, although the claim is on shakier footing in the latter case

because of poor data. The forces of improvement were sometimes fighting an uphill battle against unhealthy behaviors that left a clear mark on life expectancy: alcohol abuse in Eastern Europe, smoking and obesity trends in the United States, and practices associated with HIV/AIDS in sub-Saharan Africa.

Annual advances in life expectancy have become a routine expectation in developed countries, and the expectation is nearly always realized. Oeppen and Vaupel[53] examined the time series of the maximum life expectancy in a national population since 1800. They found that the gains have been steady at about 0.25 years per year. Vallin and Mesle[12] excluded weaker data from the calculation and found that the bacteriological era between 1880 and 1950 saw the fastest improvements at the life expectancy frontier. Even so, the frontier has advanced by 0.20 year per year since 1950.

Table 1. Speed of health improvements in different eras and main contributing factors

1700-1880	1880-1945	1945-2000
Developed Countries		
Slow advance	Fast advance	Moderate advance
Standard of living (diet) **Medicine** (smallpox) **Public health** (quarantine)	**Public health** Personal health practices	**Medicine** Antibiotics Cardiovascular treatment
1700-1880	1880-1945	1945-2000
Developing Countries		
No advance	Slow advance	Fast advance
	Public health	**Medicine** Antibiotics Immunization **Public health** Sanitation Antimalarial programs **Standard of living**

Note: Slow advance = gain of approximately 0.1 year of life expectancy per calendar year. Moderate advance = gain of approximately 0.2 year of life expectancy per year. Fast advance = gain of approximately 0.4-0.5 year of life expectancy per year.

Such steady improvements have given rise to a sense of inevitability in the advances in life expectancy. One possible mechanism producing such steady advances is a medical research establishment that continually reorients its aim towards the major health problems of the day. Such a process might be reflected in the "hand-off" from bacteriologic advances to improvements in cardiovascular disease treatment during the 1950-1970 period. But such a hand-off is not automatic. Cancer was a more promising target than cardiovascular disease in 1970, eliciting a national "war on cancer." But age-standardized death rates from cancer declined by only 12 percent between 1970 and 2008, in contrast to a decline of more than 50 percent for cardiovascular diseases.

This chapter has dealt only with contemporaneous relations between health outcomes and causal factors. Many causal factors operate with a lag. One clear example is smoking, which affects lung cancer and all-cause mortality with a substantial lag.[54] Some conditions in childhood also influence mortality throughout the lifetime of birth cohorts. These include infections, developmental problems, and nutrition in utero and during childhood. Studies of these relations are too voluminous to review here. There is no doubt that poor health in childhood is highly correlated with poor adult health.[55] Sweden and England both show clear cohort influences in their patterns of secular mortality decline, most likely a product of the persistent effects of tuberculosis infection on frailty throughout life.[56,57] However, cohort effects appear to be much less influential than period effects over time and space (see Barbi and Vaupel[58] and references therein).

Effects on Policies and Programs

Studies by demographers and economists of the forces that have driven advances in health have created a broad landscape upon which health policies have been constituted. They have demonstrated what is feasible in the world of action and not simply in the rarified world of randomized control trials. The documented successes have created an air of optimism in discussions of both domestic and international health, especially when the achievements are artfully cumulated as by Deaton.[37] And, they have helped to focus discussions on specific health programs and their cost-effectiveness and discouraged the more passive stance that health levels advance primarily as a byproduct of economic development. This re-orientation away from a "development first" approach was explicit in the World Bank's watershed 1993 World Development Report, Investing in Health.[b] This study, written under the supervision of economist Dean Jamison, presented a massive array of evidence on global health problems and cost-effective ways of addressing them.

One constituent of the 1993 report was the first study of the "Global Burden of Disease." This study generated estimates of the burden of premature mortality and disability for 107 diseases and 483 sequelae. This research clarified disease burdens for regions and nations and has been cited

[b] "Wealthier is Healthier," a 1996 paper by a senior economist and former President of the World Bank, Lant Pritchett and Larry Summers, was widely viewed as advocating a "development first" approach to health. As such, it was seen as a refutation or at least an amendment of the 1993 report. However, the authors were more careful than many readers: "By estimating income effects, we are not attempting to compare the efficiency of investment in overall income growth with investments in child mortality judged solely on the basis of the improvement in child mortality. Improved child mortality is not the only benefit of economic growth, so obviously investments specific to child health improvements are expected to be more 'cost effective' in producing health gains than economic growth."[29]

over 4,000 times. Supervised by an economist/demographer/physician, Christopher Murray, it has been updated and expanded twice, most recently for 2010.[59] Among many other accomplishments, this series of studies has brought into bold relief the burden that HIV/AIDS and malaria presented in Africa. One consequence of this and other demographic/epidemiologic efforts was a massive expansion of international aid to combat these diseases. Foreign aid from the United States for health rose from $1.7 billion in FY2001 to $8.9 billion in FY2012. The Congressional Research Service[60] estimates that, from 2006 through 2011, more than 59 million insecticide-treated nets and 11 million malaria treatments were procured with financing from the United States. Malarial deaths declined by roughly 33 percent over that period, from 985,000 in 2000 to 655,000 in 2010.

Private patterns of philanthropy were also affected by the report.[22] Bill Gates said in 2003, "I remember reading the 1993 World Development Report. Every page screamed out that human life was not being as valued in the world at large as it should be. My wife Melinda and I were stunned to learn that 11 million children die every year from preventable causes. That is when we decided to make improving health the focus of our philanthropy" (cited in Jamison et al.[61]). According to its annual report for 2012, the Gates Foundation spent $893 million on global health that year, including principal support for the Institute for Health Metrics and Evaluation.[62]

Domestically, the set of studies described in this review may have their largest policy impact within the Social Security Administration. The U.S. Social Security System is required by law to be in "close actuarial balance" over a 75-year period. In simulations performed by Social Security actuaries, the actuarial balance is more sensitive to the future of longevity than it is to any other index except real wages. So projections of longevity have important fiscal implications, and the relevance of these implications is ensured by legislation.

These projections are made through extrapolations of past changes, with allowance for potential changes in the factors responsible for mortality improvement. There is substantial disagreement over whether the rate of progress observed in the past, as represented in age-specific death rates, can be sustained in the future. Demographers have pointed out how steady such progress has been[52,63] and argued that the decline in smoking prevalence will aid future advances in life expectancy.[64] Actuaries from the Social Security Administration have often taken a more conservative stance, arguing that the elements that produced past gains in longevity cannot be expected to contribute to future improvements. As in most of the research reviewed in this chapter, the focus of discussion in technical advisory committees is on personal behaviors such as smoking, on medical systems, and on advances in standards of living. Literally trillions of dollars are at stake in the actuarial balance of the Social Security Trust Fund. Social scientists' interpretations of the factors that drive mortality change are playing a central role in planning for the Nation's fiscal future.

Acknowledgments

This chapter was supported by a grant from the National Institute on Aging [R01AG040212]. The content is solely the responsibility of the author and does not necessarily represent the official views

of the Agency for Healthcare Research and Quality, the National Institute on Aging, the National Institutes of Health, or the U.S. Department of Health and Human Services.

Author's Affiliation

Samuel H. Preston, PhD, is Professor of Sociology in the Population Studies Center, University of Pennsylvania.

Address correspondence to: Samuel Preston, Population Studies Center, 289 McNeil Building, 3718 Locust Walk. University of Pennsylvania, Philadelphia, PA 19104; email speston@sas.upenn.edu.

References

1. Preston S. Human mortality throughout history and prehistory. In: Simon J. (Ed) The state of humanity, p. 30-6. Blackwell: London; 1995.
2. World population prospects: the 2012 revision. New York: United Nations, Department of Economic and Social Affairs, Population Division; 2013.
3. Nordhaus W. The health of nations: the contribution of improved health to living standards. In: Murphy K, Topel R. (Eds) Measuring the gains from economic research: an economic approach, p. 9-40. Chicago: University of Chicago Press; 2003.
4. Murphy KM, Topel RH. The value of health and longevity. J Political Econ 2006;114(5):871-904.
5. Costa D. Health and the economy in the United States from 1750 to the present. NBER working paper 19685. Cambridge, MA: National Bureau of Economic Research; 2014.
6. McKeown T. The modern rise of population. New York: Academic Press; 1976.
7. Szreter S. The importance of social intervention in Britain's mortality decline 1859-1940. Soc Hist Med 1988;1(1):1–37.
8. Woods R. The demography of Victorian England and Wales. Cambridge: Cambridge University Press; 2000.
9. Mercer A. Micro-organisms, disease and mortality. Rochester, NY: University of Rochester Press; Forthcoming.
10. Preston S, Haines M. Fatal years: child mortality in late nineteenth century America. Princeton, NJ: Princeton University Press; 1991.
11. Ewbank DC, Preston SH. Personal health behavior and the decline of infant and child mortality: The United States, 1900-1930. In: Caldwell J, Findley S, Caldwell P, et al (Eds),What we know about health transition, pp. 116-48. Health Transition Series. Canberra: Australian National University; 1990.
12. Vallin J, Mesle F. The segmented trend line of highest life expectancies. Popul Dev Rev 2009;35(1):159-87.
13. Cutler D, Miller G. The role of public health improvements in health advances: the twentieth-century United States. Demography 2005;42(1): 1-22.
14. Easterlin RA**.** How beneficent is the market? A look at the modern history of mortality. In: Easterlin RA (Ed), The reluctant economist, pp. 101-38. Cambridge: Cambridge University Press; 2004.
15. Cutler DM, Deaton A, Lleras-Muney A. The determinants of mortality. J Econ Perspect 2006;20(3):97-120.
16. Mokyr J. Why was there more work for mother? Knowledge and household behavior, 1870-1945. J Econ Hist 2000;60(1):1-41.
17. Mokyr J, Stein R. Science, health, and household technology: the effect of the Pasteur revolution on consumer demand. In: Bresnahan TF, Gordon RJ (Eds), The economics of new goods*,* pp. 143-206. Chicago: University of Chicago Press; 1996.
18. Newman P. Malaria control and population growth. J Develop Stud 1970;6:133-58.
19. Preston SH. Causes and consequences of mortality declines in less developed countries during the 20th century. In: Easterlin RA (Ed), Population and economic change in developing countries. Cambridge, MA: National Bureau of Economic Research; 1980.
20. Caldwell J. Old and new factors in health transitions. Health Transit Rev 1992;2(Suppl):205-15.
21. Preston S. The changing relation between mortality and level of economic development. Popul Stud 1975;29(2): 231–48.

22. Deaton A. Health in an age of globalization. Brookings Trade Forum; pp. 83-130. Washington, DC: Brookings Institution Press; 2004.
23. World development report 1993: investing in health. Washington, DC: World Bank; 1993.
24. Kenny C. There's more to life than money: exploring the levels/growth paradox in income and health. J Int Dev 2009;24-41.
25. Pritchett L, Viarengo M. Explaining the cross-national time series variation in life expectancy: income, women's education, shifts, and what else? Human Development Research Paper 2010/31. London: United Nations Development Programme; 2010.
26. Hum RJ, Prabbat J, McGahan A, et al. Global divergence in critical income for adult and childhood survival: analyses of mortality using Michaelis-Menten. Elife 2012;1:e00051. DOI: 10.7554/eLife.00051 1 of 17.
27. Acemoglu D, Johnson S. Disease and development: the effect of life expectancy on economic growth. J Polit Econ 2007;115(6):925-85.
28. Spence M, Lewis M (Eds). Health and growth. Washington, DC: Commission on Growth and Development, International Bank for Reconstruction and Development; 2010.
29. Acemoglu D, Johnson S. Disease and development: a reply to Bloom, Canning, and Fink. Working Paper 20064. Cambridge, MA: National Bureau of Economic Research; 2014. Available at http://www.nber.org/papers/w20064. Accessed September 4, 2014.
30. Pritchett L, Summers LH**.** Wealthier is healthier. J Hum Resour 1996;31:4:841–68.
31. Wilkinson R, Pickett K. The spirit level: why more equal societies almost always do better. London: Allen Lane; 2009.
32. Deaton A. Health, inequality, and economic development. J Econ Lit 2003;41:113–58.
33. Bloom D, Canning D. The Preston curve thirty years on: still sparking fires. Int J Epidemiol 2007;36:498–9.
34. Leon DA. Trends in European life expectancy: a salutary view. Int J Epidemiol 2011;40:271–7.
35. Crimmins E, Preston S, Cohen B. Explaining divergent levels of longevity in high income countries. National Research Council, Panel on Divergent Trends in Longevity. Washington, DC: National Academies Press; 2011.
36. Woolf S, Aron L. U.S. health in international perspective: shorter lives, poorer health. Washington, DC: National Research Council; 2013.
37. Caldwell JC. Routes to low mortality in poor countries. Popul Dev Rev 1986;12(2):171–220.
38. Deaton A. The great escape: health, wealth, and the origins of inequality. Princeton, NJ: Princeton University Press; 2013.
39. Gakidou E, Cowling K, Lozano R, et al. Increased educational attainment and its effect on child mortality in 175 countries between 1970 and 2009: a systematic analysis. Lancet 2010;376:959–74.
40. Klemptner D, Jurges H, Reinhold S. Changes in compulsory schooling and the causal effect of education on health: evidence from Germany. J Health Econ 2011; 30(2): 340-54.
41. American Heart Association. Heart disease and stroke statistics–2013 update. Circulation 2013;127: e6-e245.
42. Glei D, Mesle F, Vallin J. Diverging trends in life expectancy at age 50: a look at causes of death. In: Crimmins E, Preston S, Cohen B (Eds), International differences in mortality at older ages: dimensions and sources. Washington, DC: National Research Council; 2011.
43. Cutler DM. Your money or your life. Oxford: Oxford University Press; 2004.
44. Ford E, Ajani UA, Croft JB, et al. Explaining the decrease in U.S. deaths from coronary disease, 1980–2000. N Engl J Med 2007;356:2388-98.
45. Mackenbach J, Looman C. Life expectancy and national income in Europe, 1900-2008: an update of Preston's analysis. Int J Epidemiol 2013;42:1100-10.
46. Kunst AE, Bos V, Lahelma E, et al. Trends in socioeconomic inequalities in self-assessed health in 10 European countries. Int J Epidemiol 2005;34(2):295-305.
47. Glied S, Lleras-Muney A. Technological innovation and inequality in health. Demography 2008;45(3):741-61.
48. Goldman DP, Smith JP. The increasing value of education to health. Soc Sci Med 2011;72(10):1728-37.
49. Jha P, Peto R, Zatononski W, et al. Social inequalities in male mortality, and in male mortality from smoking: indirect estimation from national death rates in England and Wales, Poland, and North America. Lancet 2006;368:367-70.
50. Martikainen P, Makela P, Peltonen R, et al. Income differences in life expectancy: the changing contribution of harmful consumption of alcohol and smoking. Epidemiology 2014;25:182-90.
51. Fogel RW. Economic growth, population theory, and physiology: the bearing of long-term processes on the making of economic policy. Am Econ Rev 1994;84(3):369-95.
52. Fogel RW. The escape from hunger and premature death, 1700–2100. Cambridge: Cambridge University Press; 2004.
53. Oeppen J, Vaupel JW**.** Demography: broken limits to life expectancy. Science 2002;296(5570):1029–31.
54. Preston SH, Wang H. Changing sex differentials in mortality in the United States: the role of cohort smoking patterns. Demography 2006;43(4):413-34.

55. Case A, Lubotsky D, Paxson C. Economic status and health in childhood: the origins of the gradient. Am Econ Rev 2002;92(5):1308-34.
56. Finch CE, Crimmins EM. Inflammatory exposure and historical changes in human life-spans. Science 2004;305:1736-9.
57. Hobcraft J, Menken J, Preston S. Age, period, and cohort effects in demography: a review. Population Index 1982;48(1):4-43.
58. Barbi E, Vaupel J. Comment on: Inflammatory exposure and historical changes in human life-spans. Science 2005;308:1743a-4a.
59. The global burden of disease: generating evidence, guiding policy. Seattle, WA: Institute for Health Metrics and Evaluation; 2013.
60. U.S. global health assistance: background and issues for the 113th Congress. Washington, DC: Congressional Research Service; June 21, 2013.
61. Jamison DT, Summers LH, Allleyne G, et al. Global health 2035: a world converging within a generation. Lancet 2013;382(9908):1898-955.
62. 2012 Annual report: Bill and Melinda Gates Foundation. Avalable at http://www.gatesfoundation.org/~/media/GFO/Documents/Annual%20Reports/2012_Gates_Foundation_Annual_Report.pdf
63. Lee RD, Carter L. Modeling and forecasting the time series of U.S. mortali*ty.* J Am Stat Assoc 1992;87:659-71.
64. Bongaarts J. Trends in causes of death in low-mortality countries: implications for mortality projections. Popul Dev Rev 2014;40(2):189–212.

Samuel Preston, PhD, is Professor of Sociology at the University of Pennsylvania, where he also has served as Dean of the School of Arts and Sciences, Director of the Population Studies Center, and several other leadership roles. His major area of interest has been the health of populations, including methods for measuring and analyzing health and mortality. He has devoted special attention to cause-of-death patterns across space and time, the impact of cigarette smoking on aggregate mortality, and the factors that have driven significant improvements in mortality over the past century. Dr. Preston's current research is focused on reducing the substantial uncertainty about the mortality risks associated with obesity.

Influence of Policies Focused on Behavioral Risk Factors

Changing Population Behavior and Reducing Health Disparities: Exploring the Potential of "Choice Architecture" Interventions

Theresa M. Marteau, Gareth J. Hollands, and Michael P. Kelly

Abstract

Much cancer and most diabetes would be avoided if people did not smoke, moderated alcohol use, consumed less food, and became physically active. Achieving such change in these behaviors across populations would also reduce much of the gaps in health and life expectancy between the rich and the poor. In this chapter we first describe attempts to change these four aspects of behavior over the last century, focusing on the United States and the United Kingdom. We then elaborate on one particular and recent set of interventions, altering "choice architecture" or micro-environments, that can change behavior, often without conscious awareness. We argue that these show greater potential for achieving change across populations than hitherto predominant approaches that are delivered individually and/or rely on information and persuasion. We outline the conceptual and empirical research needed to estimate the contribution choice architecture interventions could make to changing behavior in populations at the scale and pattern needed to prevent noncommunicable chronic diseases in the poorest as well as the richest. In addition to this scientific challenge, we note the need to address the political challenge that stems from free market economies built on over-consumption, including what can be considered the main vectors of noncommunicable disease in the 21st Century: tobacco, alcohol, processed foods, and transport powered by fossil fuels.

Introduction: The Problem

The majority of deaths worldwide are due to four noncommunicable diseases: cancer, cardiovascular disease, type 2 diabetes, and respiratory diseases; most of these deaths are potentially preventable. Four sets of behaviors contribute to the high and growing burden of noncommunicable disease: consumption of tobacco, alcohol, and highly processed foods, as well as physical inactivity. In the United States and the United Kingdom, most of these behaviors are more common in those who are poor, thereby contributing to the large and growing gap between the rich and the poor in premature mortality and years lived without disease or disability.

We show the changes in tobacco and alcohol consumption in the United States and United Kingdom over 100+ years [1,2] in Figure 1.

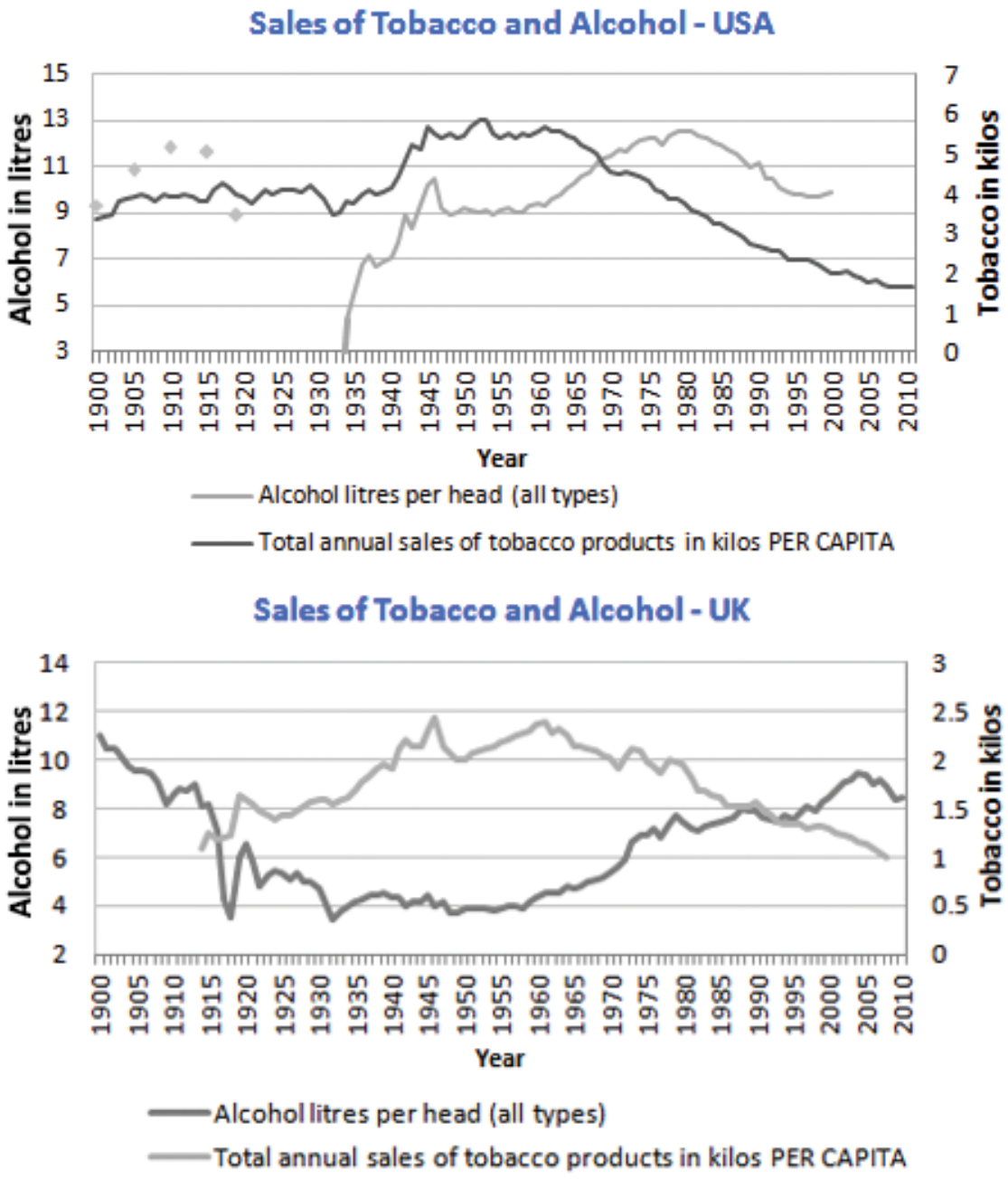


Figure 1. Sales of tobacco and alcohol in the United States and United Kingdom: 1900 – 2010

Source: Tobacco data, United States, Prof. Cristine Delnevo, Rutgers, School of Public Health (personal correspondence, 2014). Used with permission. Alcohol Data (Nephew, Williams, Hoy, et al, 2003); Tobacco Data (Forey, Hamling, Hamling, 2012); United Kingdom.

Alcohol data, United Kingdom Beer and Pub Association (Personal Correspondence, 2014).

Although sales of tobacco have declined dramatically, the decline has been less for those who are more materially and socially deprived, making smoking the most important contributor to health inequalities (Figure 2). [3,4]

Rates of physical activity in the United States and United Kingdom are shown in Figure 3.

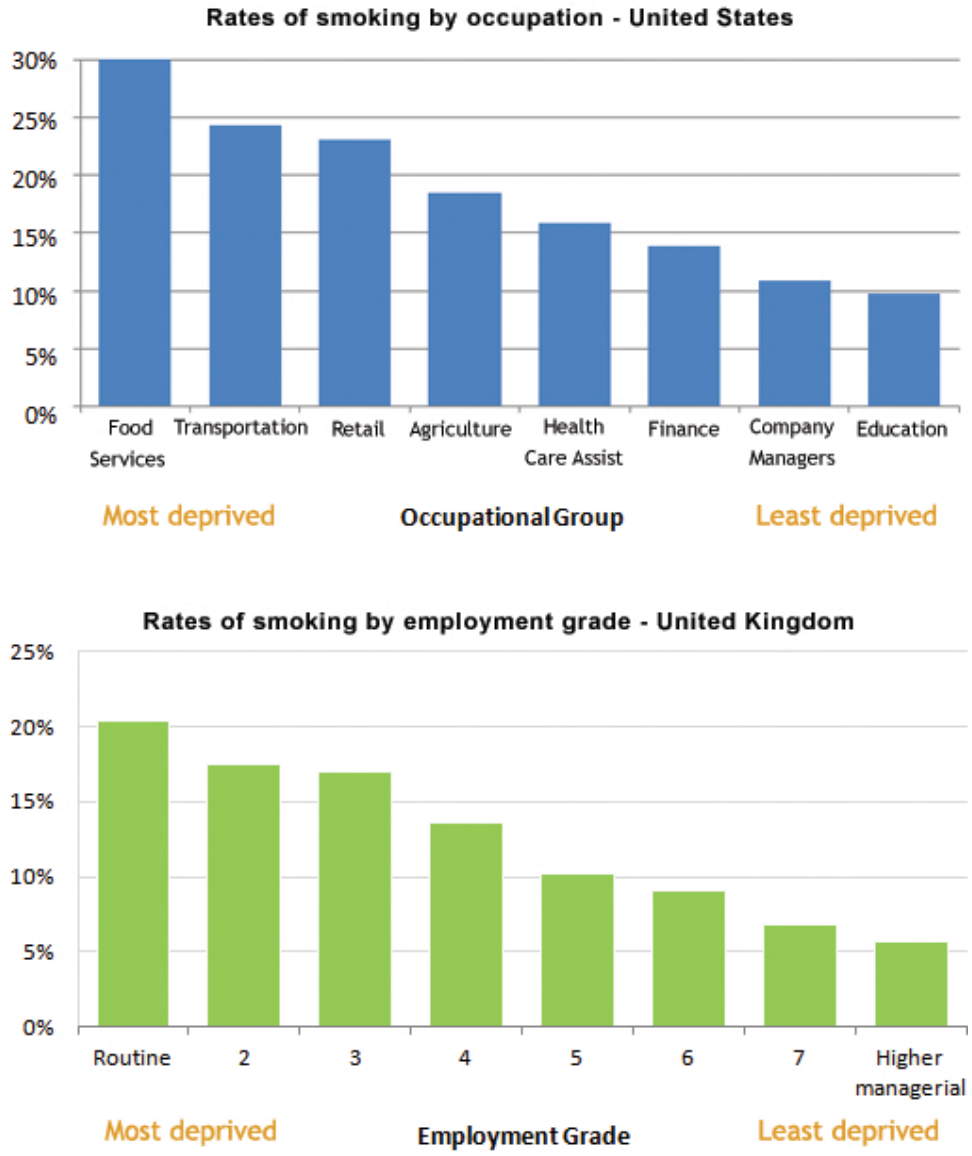
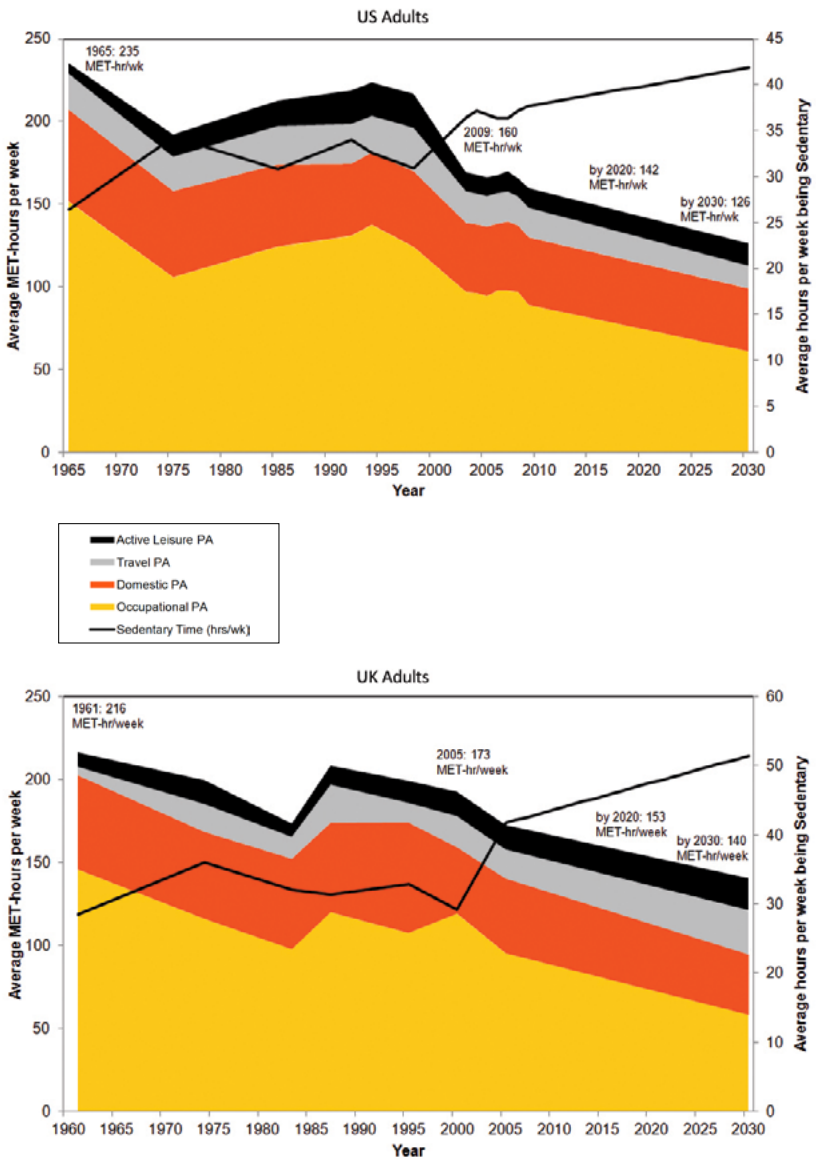


Figure 2. Rates of smoking by socioeconomic status in the United States and United Kingdom

Sources: National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention; 2011. Action on Smoking and Health (ASH) factsheet; 2013.

Rapid changes in rates of obesity, as a consequence of changes in both diet and physical activity, are shown in Figure 4. This shows variation in the rates of obesity by geographical region in the United States and United Kingdom, reflecting and reinforcing disparities in health outcomes. [5]



U.S. adults MET-hours per week of all physical activity, and hours/week of time in sedentary behavior: measured for 1965–2009, forecasted for 2010–2030. Source: Multinational Time Use Studies (MTUS) v.5.52 (1965, 1975, 1998) v.5.8 (1985, 1992, 1995), and American Time Use Survey 2003–2009; Applying Compendium of Physical Activity MET-intensity values based on reported time spent across 41 MTUS-coded activities and by occupation. Forecasting for 2010–2030 based on 2003–2009 slopes.

U.K. adults MET-hours per week of all physical activity, and hours/week of time in sedentary behavior: measured for 1961–2005, forecasted for 2006–2030. Source: Multinational Time Use Studies v.5.52 (1961, 1983, 1987), and v.5.8 (1974, 1995, 2000, 2005); Applying the Compendium of Physical Activity MET-intensity values based on reported time spent across 41 MTUS coded activities and by occupation. Forecasting for 2006–2030 based on 1961–2005 slopes.

Figure 3. Trends over time in energy expenditure related to the activity domains of active leisure, occupation, travel, home, and sedentary behavior in the United States and United Kingdom

Source: Ng & Popkin, 2012. Reproduced with permission from John Wiley and Sons.

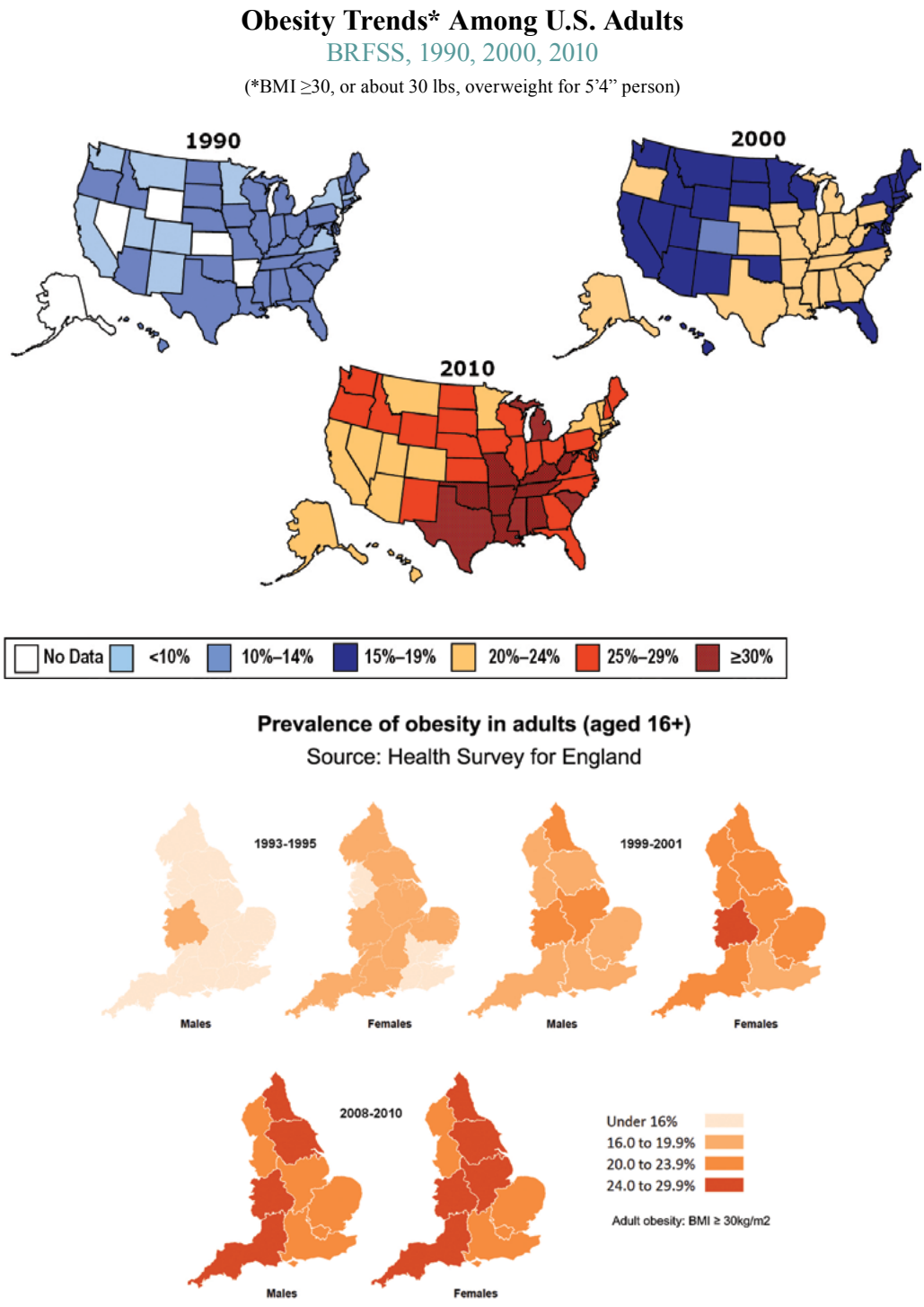


Figure 4. Rates of obesity in adults in the United States and United Kingdom by region: 1990-2010
Source: Behavioral Risk Factor Surveillance System, CDC.

The prize for changing these health-related behaviors across populations is impressive: the majority of type 2 diabetes (75 percent) and two-fifths of cancers (40 percent) would be avoided, and the gap in life expectancy between the rich and the poor halved were people to stop smoking, consume alcohol and food in modest quantities, and become physically active.[6]

We start with a brief history of organized attempts to change health-related behavior in populations over the last 100 years. This provides some evidence concerning more and less effective approaches. We argue, however, that there is a relatively limited evidence base from which to meet 21st century health challenges posed by unhealthy behaviors. Altering the environments in which these behaviors occur, or "the choice architecture," seems to hold more promise than interventions based on information and persuasion. The theoretical and evidential basis for this proposition, together with future research needs, form the final portion of this chapter.

A Brief History of Organized Efforts to Change Population Behavior to Improve Health

Governments across time have had an interest in the health of their populations to ensure, among other things, a productive workforce and sufficient numbers of fit and able people to join the military. The health focus of the U.S. and U.K. governments differed in the first compared to the second half of the 20th century. This reflected changes in the pattern of diseases affecting their populations, as well as changes in employment and growing affluence. We outline the emerging policy focus through that century towards the four sets of behavior that today form the cornerstone of policies aimed at preventing chronic noncommunicable diseases (Table 1).

The Early Years: 1900 – 1950

In 1900, life expectancy in the United States (among the white population) was 50.8 years, and in England it was 48.2 years.[12] The early 20th century was characterized by concerns about the general fitness of the population. The major threats to life, particularly for children, were those from infectious diseases, and so attempts to control their spread were the policy focus. Sanitary reforms in the 19th century to ensure clean water and safe sewage disposal had already done much to reduce infant mortality, along with the creation of public health institutions operating most often locally. These paved the way for a particular focus on the promotion of children's health (see Schorb, 2013 for a review).[10]

More Recent Efforts: 1950 – 2014

By 1950, life expectancy among the white population in the United States and all populations in the United Kingdom had risen dramatically to 69.4 and 69.2 years, respectively.[12] This period has been characterized as reflecting "an epidemiological transformation."[8,13] The leading causes of disease and death were no longer the infectious diseases of cholera, pneumonia, influenza, tuberculosis, and diarrhea that characterized the 19th century and first part of the 20th century but instead were the noncommunicable diseases of cancer, heart disease, stroke, and diabetes.

Table 1. Threats posed by diet, alcohol, tobacco, physical activity and policy responses during the first and second halves of the 20th century

1900 – 1949		1950 – 2014
The main diet-related threats to health early in the 20th century were those of malnutrition and starvation, as well as food adulteration, exacerbated by the Great Depression in the 1930s. Policies to address these problems in the United States included the food stamp program (http://www.nga.gov/exhibitions/sarg1b.shtm) and the distribution of agricultural surplus through relief programs and school lunches.[7] Britain established child welfare programs providing clean milk, vitamins, and health care.[8]	Diet	The main threat to health during this period was from excessive consumption of food in relation to need. The 1980s onwards saw the publication of many policy documents targeting obesity largely through the provision of information about diet and physical activity.[9,10] Additional interventions included nutritional labeling of food and reformulation of foods to reduce calories. While rates of obesity in the United Kingdom and United States are beginning to plateau, it is unclear which if any interventions have contributed to this change.[9]
Excessive consumption of alcohol was deemed a problem in the United States and United Kingdom in the first half of the 20th century. Prohibition in the United States resulted in a nationwide ban on the sale and consumption of alcohol in 1920, halving consumption but increasing organized crime, the latter contributing to the repeal of this legislation in 1933. Alcohol consumption was reduced in Britain during the First World War following legislation that increased the price of beer and reduced licensing hours.[11]	Alcohol	Consumption of alcohol increased in the United States and United Kingdom during this period, reflecting reduced relative price as well as increased availability and marketing. Evidence accumulated to show that policies that targeted the first two of these could reduce consumption, with the body of evidence regarding the connection between marketing and consumption being small and weak.
Tobacco during this period was generally seen as harmless, with some claims even being made for its health promoting qualities as exemplified by the use of doctors advertising particular brands of cigarettes. Government interest was primarily in tobacco taxes as a source of revenue. During the two wars in this period (1914-1918 and 1939-1945) tobacco companies vied with each other to provide troops with cigarettes as part of their rations.	Tobacco	In the second half of the 20th century, tobacco was recognized as a lethal product. Multiple tobacco control interventions in the United States and United Kingdom, often as part of the World Health Organization (WHO) Framework Convention for Tobacco Control, reduced rates of smoking. These included increasing price and restricting smoking in public. Nonetheless, tobacco consumption remains the leading cause of preventable death in both countries and is a key contributor to health disparities.
Many of the populations in the United States and United Kingdom in the first half of the 20th century were engaged in nonsedentary jobs, with few owning motor cars. Exercise, one form of physical activity, was largely viewed as a means of achieving and maintaining fitness for military action or sporting achievement.	Physical Activity	Levels of physical activity declined globally during the latter half of the 20th century including in the United States and United Kingdom, principally reflecting reductions in occupational and travel-related activity, and increases in time spent in sedentary behavior.[5] Policy interventions that favor active travel, more evident in the United Kingdom than in the United States, had some impact on increasing physical activity, but the scale of progress is deemed insufficient to stem the tide of physical inactivity and its associated widespread health-harming effects.[5]

It was during this period that lifestyle or health-related behavior was increasingly recognized as a risk factor for disease. The Framingham study was the first to use the term "risk factor" in 1961 to describe the role of diet in heart disease.[14] Smoking emerged as a key health threat in the United States and United Kingdom with the publication of landmark reports in both countries in the 1960s.[15,16] The transformation of common behaviors into risks led directly to the increasing focus on health-related behaviors and efforts to change these during the latter part of the 20th century. This shift was reflected in the growing number of citations to "health behavior" in the American Journal of Public Health (Figure 5). There was also growing awareness of health disparities.[17-20]

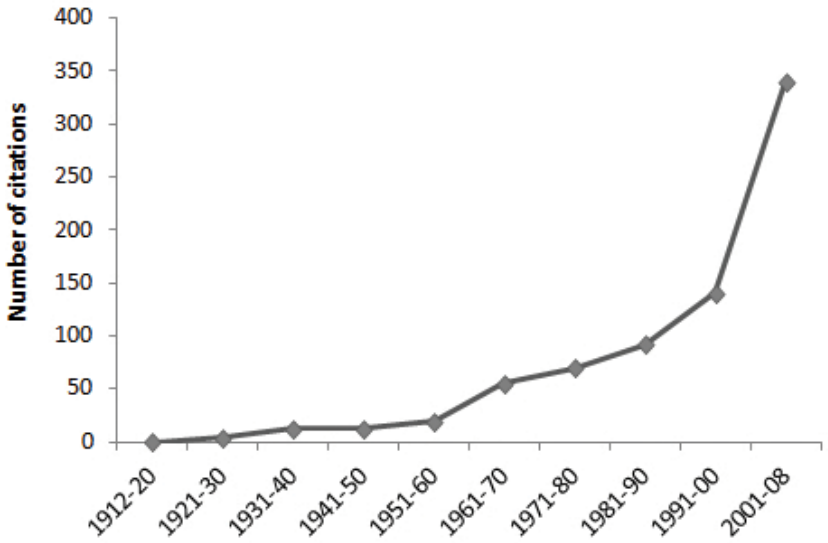


Figure 5. Citations to "health behavior" in the American Journal of Public Health: 1912-2008
Source: Armstrong, 2009. Reproduced with permission from Sage Journals.

Intervening to Change Health Behavior to Prevent Disease

Interventions that characterized attempts to change behavior at the population level varied with the target behavior, but mainly involved one of two approaches: those that required regulation, by altering the price and availability of two of the three sets of products increasingly associated with health harms (tobacco and alcohol), and educational or information-based approaches in which information was provided to persuade individuals to change their behavior. A wide variety of materials have been used to convey information including messages on television, leaflets, billboards, and direct communication in health education or counseling contexts being aimed at a range of audiences. The target for these interventions was usually particular behaviors, such as smoking, but with a gradual shift towards interventions that targeted multiple behaviors using education, counseling, or both.

There is a growing recognition of the limited effectiveness of such information-based interventions, with a recent systematic review of the evidence concluding that one-to-one, family oriented or worksite-based interventions are ineffective in general populations in reducing cardiovascular mortality or clinical events.[21] The authors stated: "Health protection through national fiscal and

legislative changes that aim to reduce smoking, dietary consumption of fats, 'hidden' salt and calories, and increase facilities and opportunities for exercise, should have a higher priority than health promotion interventions applied to general and workforce populations."[21] These findings need to be read against a divide which had by then opened up in the public health community between those who favored an approach based on individual behavior change and those who advocated changing the legal, fiscal, or regulatory climate.[22] Nonetheless, they capture the broad trend away from educational and information-based approaches as effective means to change behavior in populations at the scale and with the reach needed. The reason for this is evident in the more recent scientific literature on behavior and behavior change. This highlights the potentially greater effectiveness for changing behavior across populations of interventions that involve changing environments (with the potential to influence many people) and that focus on nonconscious processes (which appear more potent in eliciting many health-related behaviors).

The theoretical perspectives that guided the development of interventions began to shift over the latter part of the 20th century. For much of this period, the dominant models of behavior and behavior change had their roots in subjective expected utility (SEU) models.[23] In brief, these models predicted a choice based on the option with the highest expected utility, this being a function of the personal utility of an outcome and the likelihood of its occurrence. In keeping with this, communications about risks to health commonly included a message about the benefits of a targeted behavior (e.g. fruit and vegetable consumption to improve general health, or stopping smoking to reduce risk of heart disease) and the perceived probability of its occurrence. More elaborate models incorporated self-regulatory concepts of self-control and acknowledged that while these models may predict intentions to change behavior, their ability to predict actual behavior was less accurate. The dominant models included the Theory of Planned Behavior,[24] Protection Motivation Theory,[25] the Health Action Process Approach,[26] and Stages of Change.[27] Aspects of these dominant models were implicit—and occasionally explicit—in policy documents and academic writing that focused attention on behavior as a reflective, conscious choice to the exclusion of nonconscious behaviors cued by environments.[10]

By the end of the 20th century, it was apparent that there had been some successes in achieving healthier behavior in populations, most notably in relation to tobacco, likely a result of multiple interventions that included information as well as fiscal regulation and controls on advertising. But it was also clear that far more was needed to reverse the high and rising burden of noncommunicable diseases and health disparities.

Two distinct responses were discernible among behavioral researchers. The first comprised concerted efforts to continue developing interventions using existing models of behavior and behavior change but to apply these with more theoretical integrity and methodological rigor. The Behavior Change Consortium, set up in 1999 as a collective across 15 of the National Institutes of Health (NIH) institutions, encouraged evaluations that tested more than one model or involved attempts to change multiple behaviors.[28] While acknowledging the importance of combining individual and environmental approaches, particularly for increasing physical activity, the dominant models driving interventions remained those that focused on goal-directed, self-regulatory processes central to the conscious reflective systems guiding behavior. Working across these models, a group of largely U.K.-based psychologists developed a Behavior Change Technique typology as a basis for more systematic, theoretically-coherent conduct and reporting of studies.[29]

These and other initiatives will doubtless provide much needed improvement to the evidence base, particularly based more on individuals and directed at conscious processes.

The second distinct response was a recognition that the dominant paradigm, with its reliance on models that emphasized the conscious, intentional control of behavior, produced results with, at best, small effect sizes and high degrees of variance. The failure was pinned on the absence of nonconscious routes to behavior within the models that dominated the behavior change literature.[30,31] Nonconscious processes had played a central role in explanations of human behavior at the end of the 19th century and the first part of the 20th century as reflected in the writings of James, Wundt, and Freud and a little later in the work of animal learning theorists.[32,33] The cognitive revolution of the 1960s shifted the focus towards more conscious processes as evident in the models that came to dominate health behavior research.[34] Some of the most compelling evidence drawing attention to the need to embrace nonconscious processes was provided in a meta-analysis of the experimental evidence concerning the causal link between intentions and behavior.[31] The authors concluded that "...intentional control of behavior is a great deal more limited than previous meta-analyses of correlational studies have indicated." They went on to argue that future attempts to change behavior should pay greater attention to automatic, nonconscious routes to action.

Further concerns about interventions targeting individuals and those targeting conscious processes stem from their potential to increase health disparities either through accessibility or differential effectiveness. In a systematic review assessing these effects, evidence was lacking on the potential for educational interventions to affect disparities, but mass media campaigns were identified as one type of intervention associated with increased disparities and fiscal interventions associated with a decrease.[35] There also has been considerable unease about the values implicit in approaches directed to individuals, notably that a focus on individuals too readily elides their responsibility for change with blame for not changing.[36]

We focus here on the potential of one particular set of interventions that involve altering environments to cue healthier behavior, also known as "Choice Architecture." In addition to capitalizing on contemporary models of behavior that emphasize the nonconscious shaping of behavior by cues in the environment, this approach complements those based on sociologically informed understanding of health disparities, which highlights the importance of the social and material conditions that constrain the way people live.

Potential of Choice Architecture

Before we consider the potential of choice architecture to change population behavior, we first define the term and present a provisional typology and map of the evidence, as a means to identify the conceptual space inhabited by this set of interventions. "Choice architecture" is a term used to denote an approach designed to change behavior unobtrusively in predictable ways. It draws on the psychological and sociological understanding of the relationship between people and the environments they inhabit. Choice architecture interventions work by altering features of these environments to change the way in which the choices people could make are presented, thereby altering the likelihood of different behavioral responses.

Although only recently brought to the wider attention of the public and policymakers by the book Nudge,[37] the approach is not new, reflecting historical themes across psychology and neuroscience, economics, sociology, and the population health sciences. Links can be drawn, for example, to psychological perspectives on situationism—a position that regards behavior as principally determined by external, environmental factors rather than internal characteristics or motivations—and recognition of the automatic or nonconscious bases of much of our behavior.[30] It is also informed by behavioral economics with its focus on explaining why people's behavior deviates from classical economic theory.[32] From a sociological perspective, choice architecture can be conceptualized as an interaction between individual human agency and both the immediate and broader environment that make up the social structure, following a long line of sociological thought going back to Durkheim and Marx.[38] It is consistent with a population health sciences perspective because it is a population-level approach, given features of the environment that have the potential to exert influence on everyone who comes into contact with them.

In relation to changing health behavior, this approach has typically been applied to the physical aspects of small-scale or micro-environments—principally those within buildings, such as workplaces and shops—rather than macro-environments, such as city design or transport infrastructure. Given this context, we have defined choice architecture interventions in the following way:[39]

- Choice architecture interventions are those that involve altering the properties or placement of objects or stimuli within micro-environments with the intention of changing health-related behavior. Such interventions:
  - Are implemented within the same micro-environment as that in which the target behavior is performed.
  - Typically require minimal conscious engagement.
  - Can in principle influence the behavior of many people simultaneously.
  - Are not targeted or tailored to specific individuals.

A Typology of Choice Architecture Interventions

Informed by this definition, we conducted a large-scale systematic scoping review of evidence of the effects of choice architecture interventions on diet, physical activity, and alcohol and tobacco use.[39,40]

From this we developed a typology of choice architecture interventions in micro-environments, displayed on the left side of Figure 6.[41] This groups the available evidence into nine types of intervention that share common characteristics. These can be placed within two higher level classes of intervention that involve altering: (1) the properties of objects or stimuli and/or (2) the placement of objects or stimuli. As can be seen in Figure 6, some interventions involve both.

PROVISIONAL TYPOLOGY OF CHOICE ARCHITECTURE INTERVENTIONS IN MICRO-ENVIRONMENTS		MAPPING OF AVAILABLE EVIDENCE BY INTERVENTION TYPE AND TARGET BEHAVIOR			
Intervention class	**Intervention type**	**Number of study reports (combining primary research and reviews)**			
		Diet 309/440=70.2%	**Physical activity** 84/440=19.1%	**Alcohol** 32/440=7.3%	**Tobacco** 15/440=3.4%
Primarily alter **properties** of objects or stimuli	**AMBIENCE** - alter aesthetic or atmospheric aspects of the surrounding environment	33	10	14	
	FUNCTIONAL DESIGN - design or adapt equipment or function of the environment	27	11	5	
	LABELLING – apply labelling or endorsement information to product or at point-of-choice	78		7	10
	PRESENTATION - alter sensory qualities or visual design of the product	21			2
	SIZING - change size or quantity of the product	66			1
Primarily alter **placement** of objects or stimuli	**AVAILABILITY** - add behavioural options within a given micro-environment	28	6		
	PROXIMITY – make behavioural options easier (or harder) to engage with, requiring reduced (or increased) effort	21	1		
Alter both **properties and placement** of objects or stimuli	**PRIMING** - place incidental cues in the environment to influence a non-conscious behavioural response	9	1	5	1
	PROMPTING – use non-personalised information to promote or raise awareness of a behaviour	26	55	1	1

Figure 6. Provisional typology of choice architecture interventions in micro-environments (left side) and mapping of available evidence (right side)

Source: Hollands et al, 2013. Originally published by BioMed Central; Used with permission.

Mapping the Existing Evidence Base

We mapped the existing evidence base corresponding to these intervention types and the health behaviors to which they have been applied, represented on the right side of Figure 6. Although we did not exhaustively identify all of the available research, the systematic approach taken in generating the figure means that it is likely to broadly reflect the overall distribution of evidence.

Most of the evidence base comprises studies of interventions to change purchasing and consumption of food, with far fewer studies having focused on interventions to change physical activity or alcohol or tobacco consumption. Reflecting our earlier observation that persuasion through providing information has historically been the principal approach to behavior change, the most common types of choice architecture intervention are those that involve giving verbal and numeric information to be read, thus being farthest from the spirit of our definition of choice architecture. These include point-of-choice labeling, such as nutritional labeling on foods, and the use of prompts such as posters to encourage behaviors such as stair-climbing. In the Box below we provide a short summary of where the evidence is accumulating and some examples of the types of interventions that have been used to change each of four sets of behavior.

Summary of accumulating evidence for choice architecture interventions by behavior

Importantly this summary describes where the evidence is accumulating, as indicated by a scoping review,[39] but it does not provide a synthesis of effects (for which in-depth systematic reviews are needed).

Diet

Interventions to change diet-related behaviors, including the purchasing and consumption of food, make up the great majority of the choice architecture evidence-base with multiple examples for every intervention type, including:

- *Proximity*: altering layouts to increase or decrease the distance of products from routes of passage, such as choosing the area where unhealthy foods are dispensed or individual products within a salad bar are placed, nearer or farther away from customers.[42]
- *Sizing*: increasing or decreasing the portion sizes of foods served.[43]
- *Priming*: using conspicuously placed recipe posters to prime healthier eating behaviors.[44]
- *Labeling*: using nutritional labeling on the packaging of food products to indicate nutrient composition.[45]

Physical Activity

Interventions to increase physical activity were mainly those using prompting, although some studies using other interventions were found:

- *Prompting*: the use of signs or posters to encourage stair vs. lift or escalator use posted on surfaces proximal to stairwells and on stair risers.[46,47]
- *Availability*: reducing the availability of lift or escalator options, by increasing the time taken for elevator doors to close[48] or reducing the number of working escalators.[49]

Alcohol Use

There is relatively little evidence. Some examples include:

- *Ambience*: altering the type or volume of music played in shop or bar environments.[50]
- *Functional design*: changing the shape of drinking vessels.[51,52]

Tobacco Use

There is relatively little evidence. Some examples include:

- *Sizing*: altering the size of cigarettes.[53]
- *Presentation*: changing the design of cigarette packaging.[54]

Estimating the Potential of Choice Architecture Interventions

Both primary and secondary research are needed to assess the potential of choice architecture interventions to change behavior across populations. Although the typology we present in Figure 6 is provisional, primarily descriptive, and may benefit from additional analysis and development, it can, nonetheless, function to frame and inform research to populate this space with estimates of the effect sizes of choice architecture interventions. By providing a structure for knowledge of the nature and parameters of the evidence base, it can inform both primary research and the design of reviews aiming to identify and interpret the available evidence.

Figure 6 also highlights that the evidence relating to some interventions and behaviors is limited in quantity (in particular, the paucity of studies focused on physical activity and alcohol and tobacco use), indicating significant potential for primary research. In some areas, there appears to be sufficient evidence for systematic reviews to generate reliable estimates of the effects of a given intervention.

Developing Systematic Reviews of the Effects of Choice Architecture Interventions

Systematic reviews seek to collate all evidence that fits prespecified eligibility criteria in order to address a specific research question, aiming to minimize bias in estimating effects by using explicit, systematic methods.[55] A review is part of an iterative cycle of evidence generation governed by the

research question, in which findings have the potential to inform policy and practice decisions, as well as decisions about the primary research that could contribute most to addressing the given research question. Further primary research may then feed back into an updated systematic review, and the cycle continues.[56]

In the case of choice architecture interventions, the approach has gained traction in policy circles and been devalued in others, but its effectiveness remains largely unknown.[57] It is therefore premature to make judgments on likely effectiveness, since systematic reviews featuring formal critical appraisal and synthesis would be necessary to produce reliable assessments of the likely direction and magnitude of intervention effects, as well as the factors that may moderate those effects.

We are, for example, conducting a systematic review of one choice architecture intervention identified within the typology, namely "sizing," for which we found substantial evidence that has yet to be rigorously synthesized.[58] This review aims to estimate the effects of manipulating portion, package, or tableware sizes on the selection or consumption of food, alcohol, or tobacco products. In addition, the review aims to estimate the extent to which these effects may be modified by both the characteristics of the intervention—important when such interventions are likely to display significant heterogeneity in their characteristics—and of the participants. The latter is central to understanding the potential of interventions to reduce or increase existing health disparities and to determine whether there is evidence of differential effects across populations.

Choice Architecture and Health Disparities

Reducing health disparities is an important goal of health policies in the United States and United Kingdom. The social patterning of health-related behavior has been estimated to explain between 20 percent and 70 percent of the variation in life expectancy between the least and most materially and socially deprived in high-income countries.[59] Choice architecture interventions have the potential to reduce health disparities by focusing on two sets of interacting influences on these disparities: (1) those concerning the environments in which people live (physical, social, and economic), and (2) those concerning the cognitive resources of individuals to resist these environments.

Environments that house those who are more deprived contain more cues to unhealthier behaviors and fewer cues to healthier behaviors than environments that contain more advantaged residents. For example, the density of fast food outlets is greater in deprived neighborhoods,[60,61] as is the density of alcohol and tobacco retail outlets.[62,63] In addition, walkability is a feature not usually found in more deprived neighborhoods.[64]

The ability to resist tempting environmental cues, often without awareness, depends in part on two sets of cognitive resources: the individual's executive functioning, and his or her immediate cognitive capacity. There is now growing evidence to suggest that both these resources may be reduced in those who experience greater degrees of deprivation. Executive functioning is a theorized behavioral control network linked to the prefrontal cortex. One of its core functions is the inhibition of impulsive responses. The strength of executive functioning predicts obesity and alcohol and tobacco consumption, as well as physical inactivity, and there is increasing evidence of its links to environment factors experienced in one's early years, particularly poverty.[65-67] We can

think of executive function as being a relatively fixed trait, albeit with some plasticity. Immediate cognitive capacity is also influenced by poverty as evident in decisionmaking related to economic and non-economic problems, hypothesized to arise from the cognitive demands associated with poverty.[68-71]

In sum, those who are most deprived face a double hit: living in environments that provide stronger cues to unhealthy behaviors, coupled with a reduced capacity to resist impulses to these cues. These two sets of contributors to a higher likelihood of unhealthy behaviors in those who are most materially and socially deprived provide several points for intervention. These include interventions designed to protect children from environments that undermine the development of executive functioning, such as early years parenting programs[72] and income supplementation of the poorest families.[62] The second point for intervention is the environment. As highlighted in our definition of choice architecture, these interventions typically require little or no conscious engagement, and so they do not draw upon executive functioning.[30] An approach that involves the removal of cues for unhealthier behaviors and the addition of ones for healthier behaviors has the potential to shape the behavior of all those exposed, regardless of their executive functioning. Although plausible, this potential remains an untested hypothesis.

Implications for Practice

The evidence base for choice architecture interventions is at an early stage, so specific recommendations for practice are unreliable at this time. We would, however, encourage practitioners to continue to follow developments in what is a rapidly growing body of evidence. Where interventions are implemented, this should be within a rigorous evaluative framework so that it can enhance the existing evidence base.

Implications for Research

We offer the following priorities for research to build the evidence base on choice architecture interventions.

Primary research:

- Estimating the effect sizes of choice architecture interventions to change each of the four sets of behavior across populations and in those who are most deprived. In particular, there is a general paucity of studies of the effects of interventions to change levels of physical activity and alcohol and tobacco use.
- Testing the behavioral and neuroscience bases of choice architecture interventions by, for example, (1) assessing the extent to which their impact requires minimal conscious engagement with the intervention and (2) assessing the extent to which their impact is not moderated by strength of cognitive resources, including executive functioning.
- Testing components of the micro-environments identified here and exploring their links to broader social structures.

Evidence synthesis:

- Conducting systematic reviews of the effects of choice architecture interventions, where it is determined there is sufficient evidence.

Concluding Comment

The changing patterns of health-related behaviors over the last century provide examples of unhealthy behavior that has declined, albeit more slowly in those who are most deprived (smoking), unhealthy behavior that has increased and fallen in response to some effective but unpopular interventions (excessive alcohol consumption), unhealthy behavior that has increased enormously but is perhaps now plateauing (excessive consumption of food), and unhealthy behavior that is rising with few signs of abating (physical inactivity). While a wide range of interventions have contributed to the success of tobacco control, the narrower range and scale of interventions targeted at the latter two types of behavior (alcohol use and inactivity) have been inadequate due to the nature and scale of the problem. It is increasingly acknowledged that the physical, material, and social environments in which we live cue these behaviors, often without awareness; this acknowledgment is shifting focus away from largely ineffective interventions that focus on encouraging individuals to resist these environments towards interventions that focus on changing environments to cue healthier behavior. As part of this, choice architecture interventions are a promising addition to the range of interventions needed to change behavior at the scale needed, including those involving altering physical macro-environments, social environments, and economic environments.

We now face two key challenges, one scientific and one political. The scientific challenge is to generate high quality evidence to estimate the magnitude of effect of choice architecture interventions, singly and in combination, for sustained change in the four sets of behavior that contribute most to premature deaths, disability, and disparities arising from chronic, noncommunicable diseases. The political challenge stems from the tension between health and wealth creation, made more acute by the recent economic recession. The United States and United Kingdom have economies that are built on excessive consumption—including consumption of tobacco, alcohol, and food, and indeed, the use of transport powered by fossil fuels. Attempts to reduce consumption, particularly of the first three products, in the form of evidence-based policies that might include restricting their availability, reducing their appeal through changing their presentation, or increasing their prices are met with intense lobbying of governments and the public, and in some cases, lead to litigation. Meeting these challenges will require vision and commitment from research funders alongside an alignment of public and political wills to implement the growing evidence base on effective interventions that could start to turn the tide on the huge and growing global burden of potentially avoidable noncommunicable diseases.

Acknowledgments

The Behaviour and Health Research Unit, University of Cambridge, is funded as part of the Department of Health Policy Research Program (Policy Research Unit in Behaviour and Health [PR-UN-0409-10109]). The Department of Health had no role in the writing of this paper, the decision to publish, or preparation of the manuscript. Theresa Marteau was funded by the National Institutes of Health (NIH) to attend the NIH's Office of Behavioral and Social Sciences Research meeting (March 26-27, 2014), at which an earlier draft of this paper was presented. We thank Virginia Berridge for acquainting us with the historical context of current policy foci on health-related behaviors. We also thank Andrew Hoy who compiled the data on which Figures 1 and 2 are based.

Authors' Affilliations

Theresa M. Marteau, Gareth J. Hollands, and Michael P. Kelly, Behaviour and Health Research Unit, University of Cambridge, U.K.

Address correspondence to: Theresa M. Marteau, Behaviour and Health Research Unit, University of Cambridge School of Clinical Medicine, Box 113 Cambridge Biomedical Campus, Cambridge, CB2 OSR; email tm388@cam.ac.uk.

References

1. Nephew TM, Williams GD, Yi H, et al. Surveillance report #59: Apparent per capita alcohol consumption: national, state, and regional trends, 1977–2000 (Vol. Alcohol Epidemiologic Data System). Rockville, MD: NIAAA Division of Biometry and Epidemiology; 2003.
2. Forey B, Hamling J, Hamling J, et al. International smoking statistics (Web edition). A collection of worldwide historical data. Sutton, UK: P.N. Lee Statistics & Computing Ltd; 2012.
3. Syamlal G, Mazurek JM. Current cigarette smoking prevalence among working adults- United States, 2004-2010. National Center for Chronic Disease Prevention and Health Promotion, CDC. MMWR 2011; 60(38):1305-9. Available at http://www.cdc.gov/mmwr/preview/mmwrhtml/mm6038a2 htm#tab2. Accessed October 2, 2014.
4. Action on Smoking and Health (ASH). ASH factsheet. Smoking statistics who smokes and how much. Action on Smoking and Health; 2013. Available at http://ash.org.uk/files/documents/ASH_106.pdf [online]. Accessed October 9, 2014.
5. Ng SW, Popkin BM. Time use and physical activity: a shift away from movement across the globe. Obes Rev 2012; 13(8), 659-80.
6. Report by the Secretariat. Prevention and control of noncommunicable diseases. Outcomes of the high-level meeting of the General Assembly on the prevention and control of non-communicable diseases and the first global ministerial conference on healthy lifestyles and noncommunicable disease control. Sixty-Fifth World Health Assembly, 26th April 2012, Geneva.
7. Rosen G. A history of public health. New York: MD Publications, Inc; 1958.
8. Gorsky M. Public health in the West since 1800: the responses. In Berridge V, Gorsky M, Mold A (Eds.), Public health in history. New York: McGraw-Hill Education; 2011.

9. Jebb SA, Aveyard PN, Hawkes C. The evolution of policy and actions to tackle obesity in England. Obes Rev 2013; 14(Suppl 2):42-59.
10. Schorb F. Fat politics in Europe: theorizing on the premises and outcomes of European anti-"obesity-epidemic" policies. Fat Studies 2013; 2(1):3-16.
11. House of Commons, Health Committee. Report on Alcohol. London: House of Commons; 2010. Available at http://www.publications.parliament.uk/pa/cm200910/cmselect/cmhealth/151/15102 htm. Accessed August 25, 2014.
12. Livi-Bacci M. A concise history of world population (3rd Ed). Oxford: Blackwell; 2001.
13. McLeroy KR, Crump CE. Health promotion and disease prevention: a historical perspective. Generations 1994; 18(1):9-17.
14. Kannel WB, Dawber TR, Kagan A, et al. Factors of risk in the development of coronary heart disease--six year follow-up experience. The Framingham Study. Ann Intern Med 1961;55:33-50. PMID: 13751193.
15. A report of the Royal College of Physicians on smoking in relation to cancer of the lung and other diseases. London: Royal College of Physicians; 1962.
16. Smoking and health: Report of the Advisory Committee of the Surgeon General of the Public Health Service. Rockville, MD: U.S. Public Health Service, Office of the Surgeon General; 1964.
17. Antonovsky A. Social class, life expectancy and overall mortality. Milbank Q 1967; 45(2):31-73.
18. Marmot M. Fair society, healthy lives: a strategic review of health inequalities in England post-2010. London: UCL Institute of Health Equity; 2010.
19. Townsend P, Davidson N, Whitehead M. Into the Divide. Inequalities in health: the Black Report and the health divide. London: Penguin; 1992 (revised).
20. Armstrong D. Origins of the problem of health-related behaviours: a genealogical study. Soc Stud Sci 2009; 39(6):909-26.
21. Ebrahim S, Taylor F, Ward K, et al. Multiple risk factor interventions for primary prevention of coronary heart disease. Cochrane Database Syst Rev 2011; CD001561.
22. Capewell S, Graham H. Will cardiovascular disease prevention widen health inequalities? PLoS Med 2010;7(8):e1000320.
23. Savage L. The foundations of statistics. New York: Wiley; 1954.
24. Fishbein M, Ajzen I. Belief, attitude, intention and behavior: an introduction to theory and research. Reading, MA: Addison Wesley; 1975.
25. Rogers RW, Prentice-Dunn S. Protection motivation theory. In Gochman DSE (Ed.), Handbook of health behavior research 1: personal and social determinants (pp. 113-32). New York: Plenum Press; 1997.
26. Schwarzer R. Self-efficacy in the adoption and maintenance of health behaviors: theoretical approaches and a new model. In Schwarzer R (Ed.) Self-efficacy: thought control of action (pp. pp. 217-43.). Washington, DC: Hemisphere; 1992.
27. Prochaska JO, Velicer WF, Rossi JS, et al. Stages of change and decisional balance for 12 problem behaviors. Health Psychol 1994;13(1):39-46.
28. Ory MG, Jordan PJ, Bazzarre T. The behavior change consortium: setting the stage for a new century of health behavior-change research. Health Educ Res 2002;17(5):500-11.
29. Michie S, Richardson M, Johnston M, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013;46(1):81-95.
30. Marteau TM, Hollands GJ, Fletcher PC. Changing human behavior to prevent disease: the importance of targeting automatic processes. Science 2012;337(6101):1492-5.
31. Webb TL, Sheeran P. Does changing behavioral intentions engender behavior change? A meta-analysis of the experimental evidence. Psychol Bull 2006;132(2):249-68.
32. Hull CL. Principles of behavior: an introduction to behavior theory. New York: Appleton-Century Company; 1943.
33. Skinner BF. The behavior of organisms; an experimental analysis. New York: Appleton-Century Company; 1938.
34. Conner M, Norman PE. Predicting health behaviour. Buckingham, UK: Open University Press; 1996.
35. Lorenc T, Petticrew M, Welch V, et al. What types of interventions generate inequalities? Evidence from systematic reviews. J Epidemiol Community Health 2013;67(2):190-3.
36. Crawford R. Healthism and the medicalization of everyday life. Int J Health Serv 1980;10(3), 365-88.
37. Thaler RH, Sunstein C. Nudge: improving decisions about health, wealth, and happiness. New Haven: Yale University Press; 2008.

38. Kelly MP. The axes of social differentiation and the evidence base on health equity. J Royal Soc Med 2010;103(7):266-72.
39. Hollands G, Shemilt I, Marteau T, et al. Altering choice architecture to change population health behaviour: a large-scale conceptual and empirical scoping review of interventions within micro-environments Cambridge: University of Cambridge; 2013.
40. Shemilt I, Simon A, Hollands G, et al. Pinpointing needles in giant haystacks: use of text mining to reduce impractical screening workload in extremely large scoping reviews. Res Synth Methods 2014;5(1):31-49.
41. Hollands GJ, Shemilt I, Marteau TM, et al. Altering micro-environments to change population health behaviour: towards an evidence base for choice architecture interventions. BMC Public Health 2013;13(1):1218.
42. Rozin P, Scott S, Dingley M, et al. Nudge to nobesity I: Minor changes in accessibility decrease food intake Judgm Decis Mak 2011;6(4):323-32.
43. Rolls BJ, Roe LS, Meengs JS. The effect of large portion sizes on energy intake is sustained for 11 days. Obesity 2007;15(6):1535-43.
44. Papies EK, Hamstra P. Goal priming and eating behavior: enhancing self-regulation by environmental cues. Health Psychol 2010;29(4):384-8.
45. Dumanovsky T, Huang CY, Nonas CA, et al. Changes in energy content of lunchtime purchases from fast food restaurants after introduction of calorie labelling: cross sectional customer surveys. BMJ 2011;343.
46. Soler RE, Leeks KD, Buchanan LR, et al. Point-of-decision prompts to increase stair use. A systematic review update. Am J Prev Med 2010;38(2 Suppl):S292-300.
47. Webb OJ, Eves FF, Kerr J. A statistical summary of mall-based stair-climbing interventions. J Phys Act Health 2011;8(4):558-65.
48. van Houten R, Nau P, Merrigan M. Reducing elevator energy use: a comparison of posted feedback and reduced elevator convenience. J Appl Behav Anal 1981;14(4):377-87.
49. Faskunger J, Poortvliet E, Nylund K, et al. Effect of an environmental barrier to physical activity on commuter stair use. Scand J Food Nutr 2003;47(1):26-8.
50. Guéguen N, Jacob C, Le Guellec H, et al. Sound level of environmental music and drinking behavior: a field experiment with beer drinkers. Alcohol Clin Exp Res 2008;32(10):1795-8.
51. Attwood AS, Scott-Samuel NE, Stothart G, et al. Glass shape influences consumption rate for alcoholic beverages. PLoS ONE 2012;7(8):e43007.
52. Wansink B, van Ittersum K. (2003). Bottoms up! The influence of elongation on pouring and consumption volume. J Consum Res 2003;30(3):455-63.
53. Russell MA, Sutton SR, Feyerabend C, et al. Smokers' response to shortened cigarettes: dose reduction without dilution of tobacco smoke. Clin Pharmacol Ther 1080;27(2):210-8.
54. Stead M, Moodie C, Angus K, et al. Is consumer response to plain/standardised tobacco packaging consistent with Framework Convention on Tobacco Control Guidelines? A systematic review of quantitative studies. PLoS ONE 2013;8(10):e75919.
55. Higgins JPT, Green SE. Cochrane handbook for systematic reviews of interventions, version 5.0.2 [updated September 2009]. The Cochrane Collaboration; 2009. Available from http://handbook.cochrane.org/.
56. Wilson E, Abrams K. From evidence-based economics to economics-based evidence: using systematic review to inform the design of future research. In: Evidence-based decisions and economics: health care, social welfare, education and criminal justice. Oxford: Wiley-Blackwell; 2010.
57. Marteau TM, Ogilvie D, Roland M, et al. Judging nudging: can nudging improve population health? BMJ 2011;342:228.
58. Hollands GJ, Shemilt I, Marteau TM, et al. Portion, package or tableware size for changing selection and consumption of food, alcohol and tobacco. Cochrane Database Syst Rev 2014;(4):Art. No. CD011045.
59. Stringhini S, Dugravot A, Shipley M, et al. Health behaviours, Socioeconomic status, and mortality: further analyses of the British Whitehall II and the French GAZEL prospective cohorts. PLoS Med 2011;8(2):e1000419.
60. de Vogli R, Kouvonen A, Gimeno D. 'Globesization': ecological evidence on the relationship between fast food outlets and obesity among 26 advanced economies. Crit Public Health 2011;21(4):395-402.

61. Obesity and the environment: fast food outlets. Public Health Observatories in England: National Obesity Observatory; 2013. Available at http://www noo.org.uk/uploads/doc/vid_15683_FastFoodOutletMap2.pdf. Accessed October 2, 2014.
62. Cooper K, Stewart K. Does money affect children's outcomes? A systematic review. York, England: Joseph Rowntree Foundation; 2013. Available at www.jrf.org.uk. Accessed October 2, 2014.
63. Shareck M, Frohlich KL, Poland B. Reducing social inequities in health through settings-related interventions -- a conceptual framework. Glob Health Promot 2013;20(2):39-52.
64. Walking and cycling: local measures to promote walking and cycling as forms of travel or recreation. London: National Institute for Health and Care Excellence; 2012. Available at http://guidance nice.org.uk/PH41/Guidance/pdf/English. Accessed October 2, 2014.
65. Marteau TM, Hall PA. Breadlines, brains, and behaviour: targeting executive functioning and environments may loosen the link between demography and destiny. BMJ 2013;347:f6750.
66. Moffitt TE, Arseneault L, Belsky D, et al. A gradient of childhood self-control predicts health, wealth, and public safety. Proc Nat Acad Sci 2011;108(7):2693-8.
67. Raver CC, Blair C, Willoughby M. Poverty as a predictor of 4-year-olds' executive function: new perspectives on models of differential susceptibility. Dev Psychol 2013;49(2):292-304.
68. Mani A, Mullainathan S, Shafir E, et al. Poverty impedes cognitive function. Science 2013;341(6149):976-80.
69. Shah AK, Mullainathan S, Shafir E. Some consequences of having too little. Science 2012;338(6107):682-5.
70. Spears D. Economic decision-making in poverty depletes behavioral control. B E J Econ Anal Policy 2011;11(1).
71. Vohs KD. The poor's poor mental power. Science 2013;341(6149):969-70.
72. Vogel C, Brooks-Gunn J, Martin A, et al. Impacts of early Head Start participation on child and parent outcomes at ages 2, 3, and 5. Monogr Soc Res Child Dev 2013;78(1):36-63.

Theresa M. Marteau, PhD, is Director of the Behaviour and Health Research Unit in the Clinical School at the University of Cambridge, and Fellow and Director of Studies in Psychological and Behavioural Sciences at Christ's College, Cambridge. She studied social psychology at the London School of Economics and abnormal psychology at the University of Oxford. Her research interests include the development and evaluation of interventions to change behavior (principally diet, physical activity, tobacco and alcohol consumption) to improve population health and reduce health inequalities; risk perception and communication particularly of biomarker-derived risks and their weak links with behavior change; and the role of evidence in policy. She is a Fellow of both the Academy of Medical Sciences and the Academy of Social Sciences.

Gareth Hollands, PhD, is a Senior Research Associate in the Behaviour and Health Research Unit at the University of Cambridge. He studied Experimental Psychology at the University of Oxford, followed by work in clinical and research settings in neurological rehabilitation, mental health, and HIV/AIDS. He then completed an MSc (University College London) and a PhD (Institute of Psychiatry, King's College London), both in Health Psychology. His research interests include the effects of communicating risk information on health behaviour, development and testing of interventions that target non-conscious processes, and systematic reviews of interventions to change health behaviour.

Michael P. Kelly, PhD, is Honorary Senior Visiting Fellow in the General Practice and Primary Care Research Unit at the Institute of Public Health, the University of Cambridge and Honorary Professor at University College London. Professor Kelly is a public health practitioner, researcher, and academic. His research interests are in evidence-based approaches to health improvement, health inequalities, behavior change, methodological problems in public health research, evidence synthesis, coronary heart disease prevention, chronic illness, disability, physical activity, public dental health, and community involvement in health promotion. From 2005-2008, he was the co-leader of the Measurement and Evidence Knowledge Network of the World Health Organization's Commission on the Social Determinants of Health. From 2005-2014 he was Director of the Centre for Public Health at the National Institute for Health and Care Excellence (NICE).

Application of Behavior Change Theory to Preventing Unintentional Injuries

David A. Sleet and Andrea Carlson Gielen

Abstract

The use of behavioral and social science theories and methods has been critical to progress made in disease prevention and health promotion; their application to preventing unintentional injury is equally important. Injuries are largely predictable and preventable. Behavioral, psychosocial, and sociocultural factors are known contributors to injuries and can be modified using sound theoretical frameworks. While structural and environmental approaches have traditionally been associated with the greatest potential, it is rarely feasible to achieve injury reduction without some element of behavior change. In this chapter, we describe several individual and community-level theories and models with examples of research that use theory to address injury problems. For practitioners, the use of theory can both improve the effectiveness of interventions and accelerate the diffusion process. Ecological approaches that consider the dynamic interaction between behavior and the environment hold the most promise for reducing injury.

Introduction

Injury is among the most under-recognized public health problems facing the world today. Worldwide, 8 of the 15 leading causes of death for people ages 15 to 29 years are injury-related.[1] Injuries of all types account for 9 percent of global mortality and are a threat to health in every country around the world. For every death, it is estimated that there are dozens of hospitalizations and hundreds of emergency department visits.[2] Injury—which includes unintentional causes, suicides, and homicides—represents 79 percent of all deaths for individuals aged 1-30 in the United States.[3] Approximately one-third of all emergency department visits[4] and 6 percent of all hospital stays[5] are due to injuries. The costs of injuries are staggering[6]; in 2005 alone, injuries from both unintentional and violence-related causes cost the United States $355 billion. This figure includes estimates of $77 billion in medical care costs and $278 billion in work-loss costs.[7]

Globally, unintentional injuries—such as those caused by motor vehicle crashes, falls, burns, poisoning, and drowning—account for the largest injury burden, and young people are among the most vulnerable populations.[8] In the 53 countries of the World Health Organization (WHO) European Region, every year intentional and unintentional injuries kill nearly 800,000 people and are the leading cause of death for people under 45 in the developed world.[9]

From 1990-2010, global deaths from road traffic injuries increased by almost 50 percent,[10] and traffic injuries are projected to move from being the ninth leading cause of death in the world to

the third leading cause of death by the year 2020. Half of the world's road traffic deaths occur among vulnerable road users—motorcyclists (23 percent), pedestrians (22 percent), and cyclists (5 percent).[11]

In the United States, unintentional injuries are responsible for most deaths of those ages 1-30.[3] Nearly 126,500 lives are lost each year due to unintentional injuries, including car crashes, poisoning, drowning, fires, and falls.[7] Millions more survive but are left with physical, emotional, and financial scars. Because of the significance of unintentional injuries to the overall injury problem in the United States, this chapter focuses on the application of behavioral and social science theory to preventing unintentional injuries.

Science of Injury Prevention and Role of Behavior Change

The science of injury prevention teaches us that injuries are not accidents. Like most diseases, injuries are often both predictable and preventable. Today we know more about the causes and consequences of injury, and the effectiveness of injury prevention strategies, than ever before. Behavioral, psychosocial, and sociocultural factors associated with lifestyle behaviors are known contributors to injury morbidity and mortality, along with characteristics of products and environments.

While the rationale for using structural or environmental interventions to change injury patterns might seem straightforward, there is rarely an environmental change that does not require behavioral adaptation. For every technological advance, there are behavioral components that need to be addressed. Children need to wear helmets while bicycling; parents need to correctly install child safety seats and booster seats; homeowners need to check their smoke alarms and change the batteries; parents with four-sided fences around their backyard pool need to ensure that the gate to the pool is always closed; and occupants alerted by a smoke alarm still need to find their way to safety. Even the more passive approach to poison prevention through the use of child-resistant closures—one of the great successes in injury control—requires active individual effort in replacing lids correctly.[12,13]

Integrating knowledge about behavioral science into the mainstream of injury prevention research and practice will help avoid the false dichotomy between active and passive strategies and reduce the tendency to choose one over the other. In Haddon's epidemiological approach to injury, the host's role in injury reflects only personal risk at the level of the individual.[14] Much of the research on injury behavior change has been on individuals whose behavior puts them at risk, such as the drinking driver[15] or the child pedestrian.[16] However, because so many of the effective injury countermeasures are policy oriented in nature, practitioners may find behavioral change strategies useful to modify injury prevention policy at the community level.[14,17,18] Finding effective ways to activate individuals to become advocates for safer products, policies, and environments represents a new opportunity for behavior change to contribute to injury prevention.[9]

Safer products and environments require behavior change, too, on the part of manufacturers (such as toy makers) and environmental designers (such as city planners), as well as policymakers who regulate exposure to hazards and those who mandate and enforce safety behaviors (such as

legislators, judges, or police).[19,20] Cataldo et al.[21] emphasized this point when they said "Ultimately, injury control must entail some degree of behavior change, requiring the establishment and maintenance of appropriate safety behavior—by parents, legislators, judges and juries, police, health educators, physicians, reporters, and the like" (p. 233).

It is rarely feasible to achieve injury reduction without some element of behavior change. Behavior change is integral to any comprehensive approach to injury prevention. We define behavioral interventions as the development and application of behavioral science theory, knowledge, strategies, and techniques to the understanding and modification of injury risk behaviors and harms. Behavioral science applications have lagged behind other approaches to injury prevention, despite repeated calls for more behavioral science research in injury prevention.[22,23] Historically, little scholarly attention has been paid to understanding determinants of injury-related behaviors or how to initiate and sustain injury behavior changes. In the past, interventions to change injury behaviors were often based on simplistic assumptions that changing individuals' awareness about an injury problem would lead to changes in behavior. We know the process is more complex than this, and that behavior change strategies cannot rely on information and education alone. Yet many practitioners in public health still approach behavior change with this assumption in mind, rather than taking a more holistic approach, such as that suggested by an ecological model.[24]

Injury Prevention and Ecological Approaches

Researchers and practitioners have begun to recognize the importance of taking an ecological perspective to understand and intervene on contemporary public health problems, such as injury prevention.[25] In the Institute of Medicine report "Promoting Health," Smedley and Syme[26] summarized the importance of taking an ecological approach this way:

> "Perhaps the most significant contribution of behavioral and social sciences to health research is the development of strong theoretical models for interventions (p. 9). The committee …… found an emerging consensus that research and intervention efforts should be based on an ecological model"(p. 2).

The ecological model states that health and well-being are affected by a dynamic interaction among biology, behavior, and the environment, and this interaction changes over the life course.[22,27] This definition conveys the notion of multiple levels of influence on health and makes clear the importance of both individual level and community level factors in shaping health related behaviors. Reducing the number of motor vehicle deaths is an example of improving population health through interventions at multiple levels of influence. Legislative policies, educational programs, and changes in the physical and social environment all contribute to changes in injury and injury risks.[28] Thus, an ecological model has utility in both describing influencing factors and developing prevention programs.[29]

Ecological approaches have been used in efforts to modify risk factors for tobacco use, obesity, heart disease, diabetes, cancer, and HIV.[30,31] A complete discussion of injury prevention and ecological models is beyond the scope of this chapter; however, Simons-Morton, McLeroy, and Wendel[32] have provided rich detail on applying social-ecological perspectives across a variety of topics in public health, and Green and Gielen[33] have provided an in-depth analysis comparing

ecological approaches to reductions in motor vehicle crashes and tobacco use in the United States. Their analysis is elaborated further in Gielen and Green[34] in a discussion of policy and environmental interventions to promote health and safety.

Injury Prevention and Behavioral Science

Researchers and practitioners have begun to recognize the value of using behavioral and social science approaches for injury prevention and control.[18,35] Behaviors that give rise to injuries are amenable to preventive interventions. Products, practices, and programs are available that can save lives, yet many people have neither heard about them, nor accepted and adopted them. Many people do not see the need to change, do not perceive themselves to be at risk, or do not have access to affordable safety products or programs that could save their lives. Behavioral scientists can help remedy this. Behavioral and social sciences also can help by documenting behavioral and social risk factors, developing and evaluating interventions, influencing social norms, assisting in post-injury recovery from psychological harm, and by shaping individual and community preventive behaviors.[36]

Historically, much of the behavioral science applications to injury issues have appeared outside of traditional population health in disciplines such as psychology, anthropology, child development, human ecology, transportation, human factors, geriatrics, and others. These contributions have helped grow interest and relevance to public health applications for injury control.[37] However, the models, theories and behavior change strategies used to address other public health problems have been sorely underrepresented in the injury literature.[38,39] A focus on behavior can complement other approaches, since there is rarely an injury that does not have a behavioral component.

Many authors have noted the need to improve behavioral interventions by using better empirical data about behavioral determinants and by employing modern health behavior change theories, frameworks, and research methods.[40-42] Methods for rigorous basic behavioral science research and evaluation of behavior change interventions in injury prevention are discussed in detail by Thompson[43] and Lowe et al.[44] In the following sections, we discuss theories and examples that can help facilitate the change process among individuals at risk and among those in a position to influence policy and environmental change.

Theories From Behavioral Science

In recent years, there has been growing national interest in the contributions of theoretical models from the behavioral sciences to public health.[32,45,46] The limited success of behavioral change efforts in modifying injury-related behaviors, however, can be traced, in part, to failure to apply these theories to develop, implement, and test injury-prevention behavior change interventions.[47,48] When rigorous research methods are used, theories can not only help us understand causes of problems, but—because they also allow us to identify mechanisms of change—they can also help us determine why programs succeed or fail and help us build better prevention programs. Selecting the most appropriate theory is situation-specific and depends on the particular audience, setting, and characteristics of the behavior to be changed.[32]

Theories influence change across the levels of the ecological model. In translating an ecological model to action programs, Glanz, Rimer, and Viswanath[49] describe three levels of influence for change. One, intra-personal change refers to influencing an individual's knowledge, attitudes, and beliefs on his or her behavior. Theories of cognition, perception and motivation are relevant here. Two, inter-personal change refers to the influence of significant others such as families, friends, and co-workers; relevant here is the modifying effect of social influence and social norms on individual behavior. Three, community-level change includes the influence of organizational settings (such as workplaces, schools, churches) and their influence on behavior. On a larger societal level is the influence of social and health policies (such as those related to welfare reform) and other societal influences (such as poverty and disenfranchisement) that influence injury risk behaviors. Examples of models applied to the community level include community mobilization, organizational change, and inter-sectoral action. Theories and models can help explain community and individual change processes in an ecological context. For example, simple changes in community zoning and urban planning can dramatically impact injuries related to the built environment, such as bicycling and pedestrian safety.

Different intervention strategies and methods are available when working with individuals and with communities.[50,51] For example, at the individual level, intervention strategies typically include a variety of behavioral, educational, counseling, skills development, and training methods. Innovative new technologies such as computer-tailored messaging and behavioral prescriptions, Web-based learning, and motivational interviewing are promising approaches to strengthen the impact of injury prevention interventions at the individual level.[52,53] When interventions focus on organizations, communities, and policies, the use of social marketing, mass media, and media advocacy are important,[54] as are coalition building, social planning, and community development.[55]

Application of Theories

A complete enumeration of the theories and their applications in the field of health behavior change is beyond the scope of this chapter, although interested readers are referred to relevant textbooks[32,56] and reports[26,57] and recent articles using theory to influence a variety of injury prevention behaviors such as: booster seat use,[58] legislator behavior,[59] all-terrain vehicle (ATV) helmet use,[60] driver behavior,[61] and smoke alarm maintenance behaviors.[62] Behavior change theories, research methods, and applications in injury prevention have been described previously by Gielen, Sleet and DiClemente.[35] Here, we describe several examples of well-respected behavior change theories or models that have been applied to injury problems.

Individual-Level Theories and Models

The Health Belief Model, Theory of Reasoned Action, and Applied Behavioral Analysis have an extensive literature supporting their utility, and each has been used for understanding injury problem behavior. In this section, we briefly describe the key constructs of each of these models and provide an example of research that utilized them to address an injury problem.

Health Belief Model (HBM). This model says that preventive behaviors are a function of individuals' beliefs about their susceptibility to the health problem, the severity of the health

problem, the benefits versus costs of adopting the preventive behavior, and experiencing a cue to action.[63] In recent years, the concept of self-efficacy was added to the model. Self-efficacy, a concept originally from Bandura's work,[64] refers to one's confidence in his or her ability to perform a specific behavior. An illustration of this model in injury prevention comes from the study by Peterson and colleagues'[65] study of the beliefs and safety practices of 198 parents with children aged 8 to 17. They used the HBM to predict how parents' attitudes might influence their injury prevention teaching and environmental modifications. Parents generally were not very worried about injuries to their children (i.e., low perceived susceptibility). The HBM constructs most strongly associated with parental safety efforts were beliefs that their actions would be effective (i.e., benefits), a realistic appraisal of the costs of action (i.e., costs), and feeling knowledgeable and competent to perform the behaviors (i.e., efficacy). In this case, the authors[65] suggest that practitioners use interventions influencing parents' beliefs about their child's susceptibility to injury through education, while simultaneously increasing parents' competency to intervene through specific behavior change strategies. Health education methods and strategies might include direct communications to address susceptibility and skills training and improved access to safety products to address competence.

Theory of Reasoned Action (TRA) and Theory of Planned Behavior (TPB). This model describes behavior as a function of behavioral intention, subjective norms, and attitudes.[66] The model focuses on the individuals' intention to perform a behavior as predictive of their actual behavior. Intention is a function of attitudes and subjective norms. Ajzen[67] later modified the TRA to be called the Theory of Planned Behavior and included the concept of perceived behavioral control, which reflects how easy or difficult the individual perceives the behavior to be. The TPB has been used to study safe swimming practices,[68] storage of household firearms,[69] the use of safety gear by in-line skaters,[70] and drinking and driving.[71]

In practical use, the TRA was used as the conceptual framework for a survey of parents' beliefs and practices regarding use of car safety seats.[72] Attitude toward car seat use was found to be the single best variable for distinguishing between car seat users and non-users. This variable consisted of responses to six items measuring beliefs about the consequences of the behavior (e.g., using a car seat would be a hassle; your child would be better behaved in a car seat). Respondents who believed that their spouse would approve of using a car seat (a measure of subjective norm) were also more likely to report using one. These results can help practitioners develop public and patient education materials using salient messages with credible spokespersons. For example, media messages might communicate the ease with which car seat use becomes a habit with positive consequences such as child comfort and spouse approval.

The HBM and TPB were compared, along with the Locus of Control Theory (LCT), in a study of 965 Finnish youth ages 12-19.[73] Structural equation modeling was used to examine these youth's intentions to use bicycle helmets, and the authors found that both the TPB and the LCT offered good fit models for the data. Lower intention to use a helmet was associated with perceived negative peer opinions, inconvenience, not having a helmet, and believing that a helmet could not improve their safety.

Applied Behavior Analysis (ABA). ABA is a specific subfield within psychology that uses the technology of behavior modification and operant conditioning to facilitate change. Behavior is viewed as learned, and principles of stimulus control, feedback, reinforcement, and punishment

shape the acquisition, maintenance, and extinction of behavior.[74] Applied behavioral analysis or behavioral safety addresses the '"ABCs" of behavior by manipulating the Antecedents, Behaviors, and Consequences associated with behavior. Antecedents occur before the behavior (such as cues in the environment), behaviors include the context in which the behavior occurs, and consequences are what follow the behavior.

Understanding the ABCs that control a behavior can help the practitioner intervene by shaping the behavior and the environment to bring about change. Removing roadside billboards that remind drivers of drinking, increasing prompts and cues in the drinking environment that discourages drinking and driving, and selecting a designated driver can all be ways to modify the antecedents of drinking and driving. Slowing the rate of alcohol consumption, learning drinking or binge drinking refusal skills, server interventions in the drinking environment, and feedback from blood alcohol detection devices can be used to modify the drinking behavior. Social and peer support for not drinking and driving, positive feedback, incentives or rewards from bartenders or friends, and fines, license suspension, and jail time can all be used to modify consequences in a positive or negative way.[75]

Applications of ABA in road safety have effectively increased the use of safety belts[76] and child restraints,[21,77] reduced vehicle speeding,[78] and improved bicycle helmet use.[42] In other areas of injury prevention, applied behavior analysis has been used to reduce children's fall-related behavior on playgrounds[79] and change safety behaviors in a fire in public buildings.[80]

Behavioral safety approaches have had a long history of use in promoting occupational health and safety[81] and have been employed more recently to increase the use of personal protection devices, such as hard hats and ear protection, to reduce injuries on the job[82] and to increase worker productivity and morale.[83] Applying behavioral safety to the design of a "culture of safety" is a new approach at the organizational level. This is a recent understudied research area with different theoretical perspectives taken, suggesting that human factors engineering with behavioral monitoring can play an important role in injury reduction on the job.[84]

Integrating Individual-Level Models

In 1991, the National Institute of Mental Health convened a theorists' workshop to bring together creators of behavioral theory to develop a unifying framework to facilitate health behavior change.[85] Their discussions led to an enumeration of five theories that, taken together, contain virtually all the variables that have been used in attempts to understand and change human behaviors: the Health Belief Model, Social Cognitive Theory, Theory of Reasoned Action, Theory of Self-regulation and Self-control,[86] and the Theory of Subjective Culture and Interpersonal Relations.[87] Considering all five theories and their many variables, eight variables appear to account for most of the variations in health-related behaviors: (1) intentions, (2) environmental barriers, (3) skills, (4) outcome expectancies (or attitudes), (5) social norms, (6) self-standards, (7) emotional reactions, and (8) self-efficacy. It is likely that these same eight variables might also regulate and predict change in injury risk behavior (Dr. Martin Fishbein, Personal Communication, January 23, 2003).

In translating this guidance to action, Fishbein and colleagues[88] concluded that, generally speaking, for a person to perform a given behavior, one or more of the following conditions must be met:

1. The person forms a strong positive intention or makes a commitment to perform the behavior.
2. There are no environmental barriers that make it impossible to perform the behavior.
3. The person possesses the skills necessary to perform the behavior.
4. The person believes that the advantages of performing the behavior outweigh the disadvantages.
5. The person perceives more normative pressure to perform than not to perform the behavior.
6. The person perceives that performance of the behavior is consistent with his or her self-image or values.
7. The person's emotional reaction to performing the behavior is more positive than negative.
8. The person perceives that he or she has the capabilities to perform the behavior under different circumstances.

The first three factors are viewed as necessary and sufficient for producing any behavior, while the remaining five are viewed as modifying variables, influencing the strength and direction of intentions. By way of a hypothetical example, we can apply these notions to the injury control behavior of testing the functionality of a residential smoke alarm. If a homeowner is committed to testing the smoke alarm every month, has access to the alarms in the home, and has the skills necessary to successfully test the alarm, we would predict that there is a high probability he or she will perform the behavior. The probability that the individual will test his or her smoke alarm monthly would be predicted to increase even more if the homeowner also believes that testing is worth the time and trouble, knows that neighbors all test their alarms, believes that testing is consistent with his or her values as a responsible homeowner, has no negative emotional reaction to testing, and can test the alarms under different conditions in the home. According to this notion, the probability of testing monthly would be predicted to reach nearly 100 percent under these conditions. In practice, this integrated model has not been applied to this or any other injury-related behavior, but it holds promise as an innovative approach to program development, at least until such time as sufficient research is available on specific theories as they relate to injury prevention.

Community-Level Theories and Models

Community-based injury prevention occurs when people and organizations collaborate as communities to design and implement strategies to keep citizens safe.[89] A community can be defined either geographically or on the basis of common interests. Community organization and mobilization and community-based participatory research focus on the active participation and development of the community to enable community members to better evaluate and solve their own health and social problems.[90]

Gielen and Collins[91] and McLoughlin et al[92] described the difference between 'community-wide' interventions and 'community-based' programs, highlighting the importance of treating the community as the source and not simply the site of prevention programs. Green and Kreuter[93] have described the necessary components of community interventions this way: "Given reasonable resources, the chances are that a community intervention will succeed if the practitioner (1) builds from a base of community ownership of the problems and the solution; (2) plans carefully; (3) uses sound theory, meaningful data, and local experience as a basis for problem decisions; (4) knows what types of interventions work best for specific populations and circumstances; and

(5) has an organizational and advocacy plan to orchestrate multiple intervention strategies into a complementary cohesive program" (p. 261).

The application of community-level theories and models in injury prevention can be seen in the World Health Organization (WHO) "Safe Communities" movement, initiated in Sweden in the 1980s,[94,95] and as of June, 2014, 336 designated international safe communities were in operation worldwide. The program, which combines top-down with bottom-up strategies, was developed in eight steps: (l) epidemiological mapping, (2) selection of risk groups and hazardous environments, (3) formation of coalitions or interdisciplinary workgroups, (4) joint action planning involving many sectors, (5) implementation, (6) evaluation, (7) program modification from feedback, and (8) transfer of program success to others. It is structured around a broad coalition of community partners, involving business, civic organizations, local government, non-profit groups and organizations, and local residents.[89] In the United States, the Safe Community Model has been applied mostly to the traffic safety sector by the National Safety Council and has been adopted by the National Highway Traffic Safety Administration as a part of its support to the Governor's Offices of Highway Safety Programs in many parts of the country. While the model presents a unique approach to community injury prevention, research findings on its impact on reducing injury rates have been equivocal.[96,97]

In the United States, Hingson et al.[98] describe a community-based program to reduce drinking and driving in which intervention cities reduced fatal crashes by 25 percent and fatal crashes involving alcohol by 42 percent, relative to the rest of the State of Massachusetts, during the 5 years of the program, compared with the previous 5 years without the program. This community-level approach to drinking and driving prevention has appeal among many communities struggling to reduce their injury burden.[99]

Implications for Practice

For practitioners, the use of theory can both improve the effectiveness of interventions and accelerate the diffusion process. At both the individual and community levels, using theory can help clarify assumptions on which interventions are selected and, when used in conjunction with thorough needs assessments, can contribute to more successful injury prevention programs.

In the Community Trials Project to reduce alcohol-involved trauma,[100] a research-in-practice partnership was formed to focus on changes in the social and structural contexts of alcohol use. The researchers worked to implement prevention policies and activities that were evidence-based and asked communities to customize and prioritize their initiatives based on local concerns and interests. Specific components of the mobilization effort were directed toward responsible beverage service, drinking and driving, underage drinking, and alcohol access. Coalitions, task forces, and media advocacy were used to raise awareness and support for effective policies with the public and decisionmakers. Researchers conducted an evaluation of the impact of these efforts and demonstrated significant reductions of 6 percent in the reported quantity of alcohol consumed, 51 percent in driving over the legal alcohol limit, 10 percent in nighttime injury crashes, 6 percent in alcohol-related crashes, and 43 percent in alcohol-related assault injuries seen in emergency departments.[101]

Effective injury prevention will require involvement by multiple stakeholders. We need to involve engineers, social workers, pediatricians, and developmental psychologists to help uncover everyday problems and solutions that can be implemented in the community and in clinical practice. Manufacturers of consumer products that carry the risk of an injury need to help us understand the hazards that might lead to an injury; in the process, they also will help to protect themselves from recalls and costly product liability litigation.

While the goal of most individual and community-based interventions is a reduction in injury, they can also serve intermediate goals, such as increasing knowledge and awareness of safe practices (e.g. replacing batteries in smoke alarms), changing behaviors that reduce risk (e.g. wearing helmets), and improving skills (improving the installation of child safety seats).

Implications for Future Research

Behavioral sciences applications to injury prevention must rest on a solid foundation of rigorous research methods. Some of the challenges faced in past injury research are related to research design and sample size. Often, sample sizes are too small to test population effects. A related problem is that some of the interventions are not intensive enough, lack penetration in the community, or are too expensive to implement at the population level.[102]

Methods used to evaluate outcomes of injury-related interventions often do not take into account the unique attributes of the community or are not designed with community input.[103] For example, behavioral interventions can contribute to the success of community-based participatory research (CBPR)—a collaborative approach to research that equitably involves all partners in the research process. CBPR begins with a research topic of importance to the community with the aim of combining knowledge and action for social change to improve community health and eliminate health disparities.[90,104] Although participatory research can be effective in addressing multiple public health problems, it is perhaps especially important for problems that relate to individual behaviors influencing injury.

Greater use of an array of quantitative and qualitative methods is justified, including analytical approaches such as randomized control trials, quasi-experimental designs, cohort studies, case-control and case cross-over designs, ecological studies, and mixed methods designs, all of which will strengthen the evidence base for injury prevention effectiveness. The use of modern analytic methods such as video data analysis, multilevel modeling, spatial regression, geographic information system (GIS) analysis, and social network analysis can improve the precision of injury data.[105]

Injury prevention research can continue to advance our understanding of injury causes, mechanisms, risk factors, and outcomes through efforts to improve surveillance systems at the State and local levels. Improved surveillance capacity will also enable us to uncover new and emerging threats, such as the recent rise in prescription drug overdose deaths or sports-related concussion.

Research designed to evaluate the impact of policies and programs on personal behavior will require extreme sensitivity to the ethical issues surrounding the protection of individual autonomy. Identifying priority injury problems in community settings can be facilitated using needs assessment strategies.

Decisionmaking about program design and evaluation and intervention strategy selection is a shared responsibility, and it will be embraced when it remains consistent with the core values of the community.[106] Finally, cost benefit, cost utility, and cost-effectiveness studies will be needed to assess the return on investment for implementing injury prevention practices and policies and to support the value of injury prevention as a public health strategy.

Conclusions

The use of behavioral theories and methods has been critical to progress in improving public health and injury prevention. Behavioral interventions, in conjunction with structural approaches and environmental change, can enable and reinforce efforts to prevent injuries in ways that can ultimately protect whole populations.

The application of behavior change theory to unintentional injury prevention should enable practitioners and researchers to more easily identify potentially useful strategies for many injury problems. Researchers and policymakers have highlighted the need for more effective educational and behavioral approaches to injury control.[35,38,107,108] Educational approaches must focus not only on awareness, but also on skill development and training that leads to measurable behavior change and risk reduction.[47,109-112]

Because of the wide range of types of injury, preventive behaviors, and various target groups and community characteristics, there remains a strong need for additional research using accepted behavioral theories and models. Wider application of interventions using rigorous study designs, health promotion and policy change theory, multifaceted community-based interventions with community input, intervention tailoring, economic evaluations, and implementation and dissemination research is needed.[103]

More attention is also needed to the issues of training researchers and community practitioners in the application of relevant theories.[113] Training more behavioral scientists in the epidemiology of injury and the science of injury control is an urgent first step. Likewise, enhancing the behavioral and social science training of injury practitioners and researchers is necessary.

Theoretical research is needed to clarify the mechanisms by which change occurs across levels of ecological models. Applied research can help us understand and modify risk perceptions, social norms, and other psychosocial factors associated with behavior and behavior change. Child developmental research is needed to tailor interventions to accommodate differences in patterns of physical and cognitive growth. Community-level research is necessary to understand mechanisms for influencing large populations through behavioral and environmental strategies.

While evidence from a single study can provide useful information about program efficacy, the study must be repeated and replicated in different populations and in different settings. Ultimately, what is needed is substantial research on both the determinants of behavior and the efficacy and effectiveness of programs, so that recommendations can be made to practitioners about best practices. Noticeably absent from injury/behavior literature are longitudinal study designs and mediator models of analysis, both of which would aid in understanding behavior over time and the influencing factors that

account for any changes.[114] We believe these are important steps for strengthening the application of behavioral science to injury control, which in turn can contribute to changing individual behaviors, environmental conditions, and social structures in ways that prevent injuries and improve population health.

Acknowledgments

This paper was commissioned at the request of the National Institutes of Health, Office of Behavioral and Social Sciences Research for an Expert Meeting on Behavioral and Social Sciences Research Opportunities: Innovations in Population Health Metrics, March 26-27, 2014, Bethesda, MD. Portions of this chapter have been adapted from Gielen AC, Sleet D. Application of behavior change theories and methods to injury prevention. Epidemiol Rev 2003;25:65-76. The opinions presented herein are those of the authors and do not necessarily represent the official views of the Agency for Healthcare Research and Quality (AHRQ), the Centers for Disease Control and Prevention (CDC), the National Institutes of Health (NIH), or the U.S. Department of Health and Human Services.

Authors' Affiliations

David Sleet, PhD, FAAHB, National Center for Injury Prevention and Control, Centers for Disease Control and Prevention; Andrea Carlson Gielen, ScD, ScM, FAAHB, Johns Hopkins Bloomberg School of Public Health, Johns Hopkins Center for Injury Research and Policy.

Address correspondence to: David A. Sleet, PhD, Associate Director for Science, Division of Unintentional Injury Prevention, National Center for Injury Prevention and Control, Centers for Disease Control and Prevention, 4770 Buford Highway, NE, F-62, Atlanta, GA 30096; email Dds6@cdc.gov

References

1. Krug EG, Sharma GK, Lozano R. The global burden of injuries. Am J Public Health 2000;90(4):523-7.
2. Injuries. Geneva: World Health Organization; 2015. Available at http://www.who.int/topics/injuries/about/en/. Accessed March 25, 2015.
3. Haegerich TM, Dahlberg LL, Simon TR, et al. Prevention of injury and violence in the USA. Lancet 2014;384(9937):64-74.
4. Pitts SR, Niska RQ, Xu J, et al. National Hospital Ambulatory Medical Care Survey: 2006 Emergency Department Summary. Hyattsville, MD: National Center for Health Statistics; 2008.
5. Bergen G, Chen LH, Warner M, et al. Injury in the United States: 2007 chartbook. Hyattsville, MD; National Center for Health Statistics; 2008.
6. Finkelstein EA, Corso PS, Miller TR. Incidence and economic burden of injuries in the United States. New York: Oxford University Press; 2006.
7. Web-based Injury Statistics Query and Reporting System (WISQARS). Atlanta, GA: Centers for Disease Control and Prevention, National Center for Injury Prevention and Control. Available at www.cdc.gov?ncipc/wisqars. Accessed February 3, 2015.
8. Chandran A, Hyder AA, Peek-Asa C. The global burden of unintentional injuries and an agenda for progress. Epidemiol Rev 2010;32(1):110-20.
9. Sethi D, Mitis F. APOLLO policy briefing: Using advocacy for injury prevention. Centre for Research and Prevention of Injuries (CEREPRI) of the University of Athens. Rome: World Health Organization, Regional Office for Europe; 2009.
10. GBD 2010: Global Burden of Disease Study. Lancet 2012;380(9859):2053-60.
11. Peden M, Scurfield R, Sleet DA, et al (Eds). World report on road traffic injury prevention. Geneva: World Health Organization; 2004.
12. DiLillo D, Peterson L, Farmer J. Injury and poisoning. In Boll TJ, Johnson SB, Perry NW (eds), Handbook of clinical health psychology; Vol 1: Medical disorders

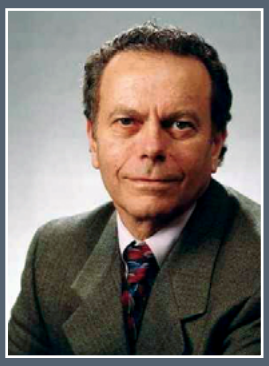
and behavioral applications. Washington, DC: American Psychological Association; 2002.
13. Shields J. Have we become so accustomed to being passive that we've forgotten to be active? Inj Prev 1997;3:243-6.
14. Runyan CW. Using the Haddon matrix: introducing the third dimension. Inj Prev 1998;4(4):302-7.
15. Geller ES, Elder JP, Hovell MF, et al. Behavior change approaches to deterring alcohol-impaired driving. In Ward WB, Lewis FM (eds), Advances in Health Education and Promotion: a research annual 1991. London: Jessica Kingsley Publishers; 1992.
16. Cross D, Hall M, Howat P. Using theory to guide practice in children's pedestrian safety education. Am J Health Ed 2003;34(5 Suppl):S-42-7.
17. Gielen AC. Health education and injury control: integrating approaches. Health Educ Q 1992;19:203-
18. Gielen AC, Girasek DC. Integrating perspectives on the prevention of unintentional injuries. In Schneiderman N, Speers MA, Silva JM, et al (eds), Integrating behavioral and social sciences with public health (p. 203-30). Washington DC: American Psychological Association; 2001.
19. McGinnis JM, Williams-Russo P, Knickman JR. The case for more active policy attention to health promotion. Health Aff 2002;21(2)78-93.
20. Shaw FE, Sleet DA, Ogolla C, et al. Law and injury prevention. Encyclopedia of primary prevention and health promotion New York: Springer; 2014.
21. Cataldo MF, Dershewitz RA, Wilson M, et al. Childhood injury control. In Krasnegor NA, Arasteh JD, Cataldo MF (eds), Child health behavior: a behavioral pediatrics perspective. New York: John Wiley and Sons; 1986.
22. Bonnie RJ, Fulco CD, Liverman CT (eds). Reducing the burden of injury: advancing prevention and treatment. Washington, DC: National Academies Press; 1999.
23. Injury prevention: meeting the challenge. The National Committee for Injury Prevention and Control. Am J Prev Med 5(3 Suppl):1-303;1989.
24. Simons-Morton B. Health behavior in ecological context. Health Educ Behav 2012;40(1):6-10.
25. Allegrante JC, Hanson D, Sleet DA, et al. Ecological approaches to the prevention of unintentional injuries. Ital J Public Health 2010;7(2):24-31.
26. Smedley BD, Syme SL (eds). Promoting health: intervention strategies from social and behavioral research. Washington, DC: National Academies Press; 2000.
27. Schneiderman N, Speers MA, Silva JM, et al (eds). Integrating behavioral and social sciences with public health. Washington DC: American Psychological Association; 2001.
28. Mercy J, Mack KA, Steenkamp M. Changing the social environment to prevent injury. In Doll L, Mercy J, Bonzo S, et al (eds), Handbook of injury and violence prevention. Secausus, NJ: Springer; 2007.
29. Green, LW, Richard L, Potvin L. Ecological foundation of health promotion. Am J Health Promot 1996; 10(4):270-81.
30. Green LW. Health education's contribution to public health in the twentieth century: a glimpse through health promotion's rear-view mirror. Annu Rev Public Health 1999;20:67-88.
31. Holtgrave DR, Doll LS, Harrison J. Influence of behavioral and social science on public health policy making. Am Psychol 1997;52:154-66.
32. Simons-Morton B, McLeroy KR, Wendel ML. Behavior theory in health promotion practice and research. Burlington, MA: Jones & Bartnett Learning; 2012.
33. Green LW, Gielen AC. Evidence and ecological theory in two public health successes for health behavior change. In Kahan S, Gielen A, Fagan P, et al (eds), Health behavior change in populations. Baltimore, MD: Johns Hopkins University Press; 2014.
34. Gielen AC, Green LW. The impact of policy, environmental, and educational interventions: a synthesis of the evidence from two public health success stories. Health Educ Behav 2015;42(1 Suppl):20S-34S.
35. Gielen AC, Sleet DA, DiClemente R. Injury and violence prevention: behavioral science theories, methods and applications. San Francisco: Jossey-Bass; 2006.
36. Sleet DA, Hammond R, Jones N, et al. Injury and violence prevention in the community. In Rozensky RH, Johnson NG, Goodheart CD, et al (eds), Building healthy communities: psychology's role in public health (p. 185-216). Washington, DC: American Psychological Association; 2004.
37. Sleet D, Hopkins K (eds). Bibliography on behavioral science research and unintentional injury prevention (CD-ROM). Atlanta, GA: National Center for Injury Prevention and Control, Centers for Disease Control and Prevention; 2004.
38. Shields WC, Perry EC, Szanton SL, et al. Knowledge and injury prevention practices in homes of older adults. Geriatr Nurs 2013;34(1):19-24.
39. Trifiletti LB, Gielen AC, Sleet DA, et al. Behavioral and social sciences theories and models: are they used in unintentional injury prevention research? Health Educ Res 2005;20(3):296-307.
40. Geller ES, Berry T, Ludwig T, et al. A conceptual framework for developing and evaluating behavior change interventions for injury control. Health Educ Res 1990;5(2):125-37.
41. Runyan CW. Progress and potential in injury control. Am J Public Health 1993;83(5):637-9.
42. Thompson NJ, Sleet D, Sacks JJ. Increasing the use of bicycle helmets: lessons from behavioral science. Patient Educ Counsel 2002;46(3):191-7.

43. Thompson N. Study methods for understanding injury behavior. In Gielen A, Sleet DA, DiClemente R (eds), Injury and violence prevention: behavioral science theories, methods and applications. San Francisco: Jossey Bass; 2006.
44. Lowe J, Yang J, Heiden E, et al. Intervention research and program evaluation. In Gielen A, Sleet DA, DiClemente R (eds), Injury and violence prevention: behavioral science theories, methods and applications. San Francisco: Jossey Bass; 2006.
45. Gielen AC, Sleet DA, Parker EM. Unintentional injury and behavior change. In Kahan S, Gielen A, Fagan P, et al (eds), Health behavior change in populations. Baltimore, MD: Johns Hopkins University Press; 2014.
46. Glanz K, Bishop D. The role of behavioral science theory in development and implementation of public health interventions. *Annu Rev Public Health* 2010;31: 399-418.
47. Liller KD, Sleet, DA (eds). Injury prevention (Special Issue*)*. Am J Health Behav 2004;28(suppl 1):S1-72.
48. McGlashan AJ, Finch CF. The extent to which behavioral and social sciences theories and models are used in sport injury prevention research. Sports Med 2010;40(10):841-58.
49. Glanz K, Rimer B, Viswanath K. Health behavior and health education: theory, research and practice (4th Ed). San Francisco, CA: Jossey-Bass; 2008.
50. Bartholomew LK, Parcel GS, Kok G, et al. Intervention mapping. Mountain View, CA: Mayfield Publishing; 2001.
51. Bensley RJ, Brookins-Fisher J. Community health education methods: a practical guide; 2nd ed. Mount Pleasant, MI: Central Michigan University; 2003.
52. Dunn C, DeRoo L, Rivara F. The use of brief interventions adapted from motivational interviewing across behavioral domains: a systematic review. Addiction 2001;96:1725-42.
53. Kreuter MW, Jacobsen HA, McDonald EM, et al. Developing computerized tailored health messages in community health education methods: a practical guide (2nd edition). Sudbury, MA: Jones and Bartlett; 2003.
54. Wallack L, Dorfman L, Jernigan D, et al. Media advocacy and public health: power for prevention. Washington, DC: Sage Publications; 1993.
55. Bracht N, Rice RE. Community partnership strategies in health campaigns. In Rice RE, Atkin CK (eds), Public communication campaigns (4th ed, chapter 20, p 289-304). Thousand Oaks, CA: Sage; 2012.
56. Kahan S, Gielen A, Fagan P, et al (eds). Health behavior change in populations. Baltimore, MD: Johns Hopkins University Press; 2014.
57. Institute of Medicine, Committee on Health and Behavior. Health and behavior: the interplay of biological, behavioral and societal influences. Washington, DC: National Academies Press; 2001.
58. Bruce BS, Snowdon AW, Cunningham C, et al. Predicting parents' use of booster seats, Inj Prev 2011;17:313-8.
59. Tung G, Vernick JS, Reiney E, et al. Legislator voting and behavioral science theory: a systematic review. Am J Health Behav 2012;36(6):823-33.
60. Adams LE, Aitken ME, Mullins SH, et al. Barriers and facilitators to all-terrain vehicle helmet use. J Trauma Acute Care Surg 2013;75:S296-300.
61. Carey RN, MDermott DT, Sarma KM. The impact of threat appeals on fear arousal and driver behavior: a meta-analysis of experimental research 1990-2011. PLoS One 2013;8(5):e62821.
62. Miller TR, Bergen G, Ballesteros MF, et al. Increasing smoke alarm operability through theory-based health education: a randomised trial. J Epidemiol Community Health 2014;68(12):1168-74.
63. Janz NK, Becker MH. The health belief model: a decade later. Health Educ 1984;11:1-47.
64. Bandura A. Perceived self-efficacy in the exercise of personal agency. Psychol Bull Br Psychol Soc 1989;10:411-24.
65. Peterson L, Farmer J, Kashani JH, Parental injury prevention endeavors: a function of health beliefs? Health Psychol 9(2):177-91;1990.
66. Fishbein M, Ajzen I. Belief, attitude, intention, and behavior: an introduction to theory and research. Reading, MA: Addison-Wesley; 1975.
67. Ajzen I. The theory of planned behavior. Org Behav Hum Decis Process 1991;50:179-211.
68. White K, Hyde M. Swimming between the flags: a preliminary exploration of the influences on Australians' intentions to swim between the flags at patrolled beaches. Accident Anal Prev 2010;42(6):1831-8.
69. Johnson RM, Runyan CW, Coyne-Beasley T, et al. Storage of household firearms: an examination of the attitudes and beliefs of married women with children. Health Educ Res 2008;23(4):592-602.
70. Deroche T, Stephan Y, Castanier C, et al. Social cognitive determinants of the intention to wear safety gear among adult in-line skaters. Accident Anal Prev 2009;41(5):1064-9.
71. Moan IS, Rise J. Predicting intentions not to "drink and drive" using an extended version of the theory of planned behavior. Accid Anal Prev 2011;43:1378-84.
72. Gielen AC, Eriksen MP, Daltroy LH, et al. Factors associated with the use of child restraint devices. Health Educ Q 1984;11(2):195-206.
73. Lajunen T, Rasanen M. Can social psychological models be used to promote bicycle helmet use among teenagers? A comparison of the Health Belief Model, Theory of Planned Behavior, and the Locus of Control. J Safety Res 2004;35(1):115-23.
74. Hovell MF, Elder JP, Blanchard J, et al. Behavior analysis and public health perspectives: combining

paradigms to effect prevention. Educ Treat Child 1986;9(4):287-306.
75. Sleet DA, Lonero L. Behavioral strategies for reducing traffic crashes. In Breslow L (ed), Encyclopedia of public health (p. 105-7). New York: McMillan; 2002.
76. Streff FM, Geller ES. Strategies for motivating safety belt use: the application of applied behavior analysis. Health Educ Res 1986;1:47-59.
77. Sleet DA, Hollenbach K, Hovell M. Applying behavioral principles to motor vehicle occupant protection. Educ Treat Child 1986;9(4):320-33.
78. Foss R. Addressing behavioral elements in traffic safety: a recommended approach. In: Improving traffic safety culture in the United States: the journey forward (p. 149-63). Washington, DC: AAA Foundation for Traffic Safety; 2007.
79. Heck A, Collins J, Peterson L. Decreasing children's risk taking on the playground. J Appl Behav Anal 2001;34:349-52.
80. Leslie J. Behavioral safety: extending the principles of applied behavioral analysis to safety in fires in public buildings. In Proceedings of the 2nd International Symposium on Human Behavior in Fire. Boston, MA: March 2001.
81. Margolis BL, Kroes WH (eds). The human side of accident prevention: psychological concepts and principles which bear on industrial safety. Springfield, IL: Charles C Thomas; 1975.
82. Krause TR. The behavior-based safety process: managing involvement for an injury-free culture. New York, NY: Van Nostrand Reinhold; 1997.
83. Geller ES. Psychological science and safety: Large-scale success at preventing occupational injuries and fatalities. Curr Dir Psychol Sci 2011;20 (2):109-14.
84. Macchi L, Pietikainen E, Liinasuo M, et al. Safety culture in design. Final report. Roskilde, Denmark: Nordic Nuclear Safety Research, NKS; 2013. Available at http://www.iaea.org/inis/collection/NCLCollectionStore/_Public/44/054/44054114.pdf. Accessed February 5, 2015.
85. Fishbein M, Bandura A, Triandis HC, et al. Factors influencing behavior and behavior change. Final report – theorists workshop. Washington, DC. National Institute of Mental Health; 1991.
86. Kanfer R, Kanfer FH. Goals and self-regulation: applications of theory to work settings. In Maehr ML (ed), Advances in motivation and achievement, vol 7 (p. 287-326). Greenwich, CT: JAI Press; 1991.
87. Triandis HC. Values, attitudes and interpersonal behavior. In Howe HE, Page MM (eds), Nebraska symposium on motivation 1979 (p. 197-259). Lincoln, NE: University of Nebraska Press; 1980.
88. Fishbein M, Triandis HC, Kanfer FH, et al. Factors influencing behavior and behavior change. In Baum A, Tevenson TA, Singer JE (eds), Handbook of health psychology. Mahwah, NJ: Lawrence Erlbaum Associates; 2001.
89. Coggan C, Bennett S. Community-based injury prevention programs. In McClure R, Stevenson M, McEvoy S (eds), The scientific basis of injury prevention and control. Melbourne: ISA Press; 2004.
90. Minkler M, Wallerstein N. Community-based participatory research for health. 2nd ed. San Francisco, CA: Jossey-Bass; 2008.
91. Gielen AC, Collins B. Community-based interventions for injury prevention. Fam Community Health 1993;15(4):1-11.
92. McLoughlin, E, Vince, C, Lee, A and Crawford, J. Project Burn Prevention: outcome and implications. Am J Public Health 1982;72:241-7.
93. Green LW, Kreuter MW. Health promotion planning: an educational and environmental approach (2nd ed). Mountain View: CA: Mayfield Publishing; 1991.
94. Moller J. An introduction to community-based injury prevention. In Ozanne-Smith J, Williams F (eds), Injury research and prevention: a text. Victoria: Monash University Accident Research Center; 1995.
95. Svanstrom L. Building safe communities – a safety promotion movement entering 2000. In 4th IUHPE European Conference on the Effectiveness and Quality of Health Promotion. Best practices. May 16-19, 1999. Helsinki, Finland and Talinn, Estonia.
96. Rahimi-Movaghar V. Controlled evaluation of injury in an international safe community: Kashmar, Iran. Public Health 2010;124(4):190-7.
97. Spinks A, Turner C, Nixon J, et al. The 'WHO Safe Communities' model for the prevention of injury in whole populations. Cochran Database Syst Rev 2009;3:CD004445.
98. Hingson R, McGovern T, Howland J, et al. Reducing alcohol-impaired driving in Massachusetts: the Saving Lives Program. Am J Public Health 1996;86:791-7.
99. Spinks A, Turner C, McClure R, et al. Community-based programs to promote use of bicycle helmets in children aged 0-14 years: a systematic review. Int J Inj Contr Saf Promot 2005;12(3):131-43.
100. Treno AJ, Holder HD. Community mobilization: evaluation of an environmental approach to local action. Addiction 1997;92(Suppl 2):173-87.
101. Holder HD, Gruenewal PJ, Ponicki WR, et al. Effect of community-based interventions on high-risk drinking and alcohol-related injuries. JAMA 2000;284(18):2341-7.
102. Prinz RJ. Parenting and the prevention of childhood injuries. In Doll L, Bonzo S, Mercy J, et al (eds), Handbook of injury and violence prevention. New York: Springer; 2007.
103. Oldenburg BF, Brodie AM. Trends and challenges in intervention research methods. In Doll L, Mercy J, Bonzo S, et al (eds), Handbook of injury and violence prevention (p. 359-80). Secausus, NJ: Springer; 2007.
104. Green LW, Mercer SL. Can public health researchers and agencies reconcile the push from funding bodies and the pull from communities? Community-

based participatory research. Am J Public Health 2001;91(12):1926-43.
105. Li G, Baker SP (eds). Injury research: theories, methods, and approaches. New York: Springer; 2012.
106. Stallones L, Thoreson S. Community readiness as a tool for engaging community and researchers in injury prevention programs. In Doll L, Mercy J, Bonzo S, et al (eds), Handbook of injury and violence prevention. Secaucus, NJ: Springer; 2006.
107. Grossman D, Johnston B. Behavioral approaches to injury control. Participants' notebook and workshop materials. Seattle: Harborview Injury Prevention and Research Center; January 23, 2003.
108. Zaza S, Thompson RS (eds). The guide to community preventive services: reducing injuries to motor vehicle occupants: systematic reviews of evidence, recommendations from the Task Force on Community Preventive Services, and expert commentary. Am J Prev Med 2001;21(4 Suppl):1-90.
109. Gielen AC (ed). Injury and domestic violence prevention. Patient Educ Counsel 2002;46(3):163-8
.
110. Ludwig TD, Geller ES, Mawhinney TC. Intervening to improve the safety of occupational driving: a behavior-change model and review of empirical evidence. J Organ Behav Manage 2000;19(4):11-2.
111. Schwartz R (ed). Focal point issue: injury prevention and control. Health Promot Pract 2003;4:85-196.
112. Sleet DA, Bryn S (eds). Injury prevention for children and youth. (Special Issue) Am J Health Educ 2003;34(5):S1-64.
113. The contributions of behavioral and social sciences research to improving the health of the nation: a prospectus for the future. Rockville, MD: National Institutes of Health, Office of Behavioral and Social Sciences Research; 2007.
114. Sleet DA, Trifiletti LB, Simons-Morton B, et al. Individual level models: applications of behavioral theory. In Gielen AC, Sleet DA, DiClementi R (eds), Injury and violence prevention: behavior change theories, methods, and applications. San Francisco, CA: Jossey-Bass; 2006.

David Sleet, PhD, is the Associate Director for Science for the Division of Unintentional Injury Prevention at the National Center for Injury Prevention, Centers for Disease Control and Prevention. Previously, he taught and conducted research in public health at San Diego State University, directed the Road Accident Research Unit at University of Western Australia, and spent 4 years as an IPA with the U.S. Department of Transportation. He has published extensively and is on the editorial boards of seven scientific journals. Dr. Sleet is a Fellow of the American Academy of Health Behavior, and received the Fries Foundation Prize for Health Education in 2015.



Andrea Carlson Gielen, ScD, is Professor, Department of Health, Behavior and Society and Director, Johns Hopkins Center for Injury Research and Policy at the Johns Hopkins Bloomberg School of Public Health. Her research focuses on the development, implementation, and evaluation of behavior change interventions to promote health and safety in clinic and community settings. Dr. Gielen has expertise in child injury prevention and domestic violence. She has led federally funded research grants on motor vehicle safety; home safety, including fires and burns, poisonings, and falls; and the relationship between intimate partner violence and HIV risk in women. Dr. Gielen's work has been recognized by the Research Laureate Award from the American Academy of Health Behavior and the Award of Excellence from the American Public Health Association.

Cigarettes: The Rise and Decline But Not Demise of the Greatest Behavioral Health Disaster of the 20th Century

David B. Abrams, Allison M. Glasser,
Andrea C. Villanti, and Raymond Niaura

Abstract

Tobacco use is expected to prematurely kill 1 billion people globally during the 21st century. More must be done than the current status quo approaches to tobacco control to avert this preventable global disaster. In contrast to a longstanding (status quo) and very effective focus of the U.S. tobacco control community on eliminating all tobacco/nicotine use behavior, the 50th anniversary U.S. Surgeon General's Report, for the first time, articulated a potentially game-changing harm-minimization strategy. The evidence is incontrovertible that the behavior of inhaling lethal smoke from combusting tobacco, primarily from the mass-produced, widely appealing, and heavily marketed cigarette, is responsible for the majority of deaths. The possible change to a harm minimization strategy is in part due to the urgent imperative to further reduce the massive disease burden resulting from cigarette use. This chapter selectively explores, at a broad brush or macro level of granularity, some of the individual and population perspectives that influenced arguably the greatest human engineered, behavioral lifestyle, public health disaster of the 20th century. We will briefly review the equally dramatic public health success of limiting tobacco use to date and discuss the rapidly changing present landscape with an increasingly diverse range of emerging tobacco and tobacco-derived nicotine products. Products that, if responsibly made and marketed and prudently managed by regulators and policymakers, could provide—for the first time in over a century—a way out for those unable or unwilling to stop using cigarettes and thus make cigarettes obsolete. We will conclude with a key question for science, practice, and policy: Should the endgame be the elimination of use of all forms of tobacco/nicotine, or is public health better served by speeding the obsolescence or outright banning of the lethal combustible products by supporting safer forms of nicotine use for those smokers who do not want to or are unable to quit all use?

Epidemic of Tobacco Use and Its Effects on Health

Overview of the Epidemic of Tobacco Use

It has now been more than 50 years since the original 1964 U.S. Surgeon General's report[1] on smoking and health. In 2014, tobacco use, primarily the combustible tobacco products dominated

by cigarettes, is estimated by the Surgeon General to prematurely kill even more users than ever before: 480,000 adults annually; furthermore, it is estimated that 5.6 million children alive in 2014 will die prematurely of a tobacco-related disease unless more is done.[2] The Surgeon General also reports that cigarettes have killed over 20 million users, more American deaths than in any of the wars since the founding of the Nation. Moreover, the effects of combustible tobacco have been underestimated because analysts combined disease-specific causes of death rather than considering all-cause mortality.[3]

Despite 50 years of significant progress that has cut U.S. cigarette use by more than 50 percent (demonstrating that one can change human behavior on a large, population-wide scale), further and more rapidly reducing smoking and its exorbitant cost of over $200 billion annually to society, remains a priority.[2] While there are many other pressing challenges in U.S. health care (obesity, eliminating disparities, universal/affordable care, escalating costs), smoking remains the primary cause of preventable death. Although stopping use, or not initiating use, is the best advice, the reality is that some youngsters will start, and many current smokers don't want to or cannot easily quit. Thus, for those already using tobacco and having difficulty stopping despite all that has been done to further change smoking behavior (prevention, treatment, and policy efforts), the use of alternatives and avoiding or desisting use of any combustible tobacco product (cigarettes, cigars, hookah, or pipe) at the earliest age possible may dramatically minimize the cumulative morbidity, mortality, and damage from combustible products.[4,5]

The disease burden, premature deaths, and high costs to the Nation that can be attributed to smoking will continue well into the next century unless we pick up the pace to reduce smoking prevalence more rapidly than the current trajectory. The results of several modeling studies projecting population smoking prevalence document the need for more aggressive action than the prevention and cessation interventions and tobacco control policies currently in place.[6,7] Projections indicate that the prevalence of adult smoking could likely remain above the Healthy People 2020 objective of 12 percent, even by mid-century, if there is little change to current strategies.[7,8]

More aggressive action is needed because over 44 million Americans still smoke lethal cigarettes, despite the mountain of evidence showing adverse consequences.[2] Cigarettes have become more appealing and addictive than ever before, presumably as a result of the tobacco industry's relentless product enhancements.[2] Further, a widening chasm in prevalence by socioeconomic gradient means smoking is a primary driver of overall health disparities.[9-11] The unprecedented resilience, persistence, and survival of the lethal cigarette is astounding and is instructive across biological, behavioral, and social science domains, in terms of understanding the complex systems forces at work. This is especially true, given that the basic product has largely been unaltered and has dominated the retail market for over 100 years, a period aptly termed the "Cigarette Century" and the "Golden Holocaust."[12,13] "The cigarette is also a defective product, meaning not just dangerous but *unreasonably* dangerous, killing half its long-term users. And addictive by design" (p. i27)."[14]

Tobacco use behavior is embedded and embodied within sociocultural and neurobiological reciprocal interactions. This includes the behavior of individuals and the "behavior" of human-created and dynamically changing multi-level contexts (i.e., the collective behavior of various group aggregations, such as family and friendship networks, neighborhoods and culture, and organizations, corporations, governments, advocates, and policymakers). This goes along with the biological and neuro-behavioral underpinnings of the psychoactive effects of nicotine and other ingredients inhaled in combusting tobacco smoke that result in the sensorimotor satisfaction and appeal of smoking, its addiction liability, and the unavoidable toxic exposure from combustion by-products. The complexity of these pathways of influence on patterns of individual and collective behavior yields an important perspective on the need for an integrative strategy to understand and perhaps better intervene in the space of the bio-behavioral, sociocultural, economic, and population-level pathways and across scientific disciplines to improve the health of whole populations.[15-17]

At a deeper level of analysis, given what we have known about the lethality of cigarettes since at least the publication of the first 1964 Surgeon General's report,[1] the question that needs to be answered is why and how what might be termed "a mass produced weapon of global mass destruction" has been permitted to continue to flourish and expand throughout the world from developed to developing nations? What story does this tell us about understanding human behavior change, the conceptual models and tools we use, and where the gaps are (what is missing or unexplained)? Perhaps one answer, although unsatisfying, is that it is complicated: the explanation lies not in one factor or element (each element is quite well understood), but in the interaction of various components in a complex dynamic interplay at multiple levels, from cells to society. Some of the interactions and trajectories over time are still poorly specified or appreciated, especially the influences on behavior of so-called long- and short-term feedback loops (of vicious and virtuous cycles), as well as the influences of systems within systems (see later in this chapter).[18,19] Lack of true systems integration has been proposed as the single most important missing ingredient in understanding the transformative and resilient cigarette epidemic and the failure of societies and governments worldwide to eliminate mass cigarette production and marketing from the consumer product marketplace, despite evidence that it's a defective product.[20,21]

The myriad factors influencing smoking behavior have led first to an increase, then a decline, and now a stubborn persistence of cigarette use over the last century (Figure 1). On the one hand, this story plays out within individuals over time, from the marketing of products that appeal to youth—like mentholated cigarettes—to the mere seconds it takes a teenager to "decide" to engage in a risky behavior, to the minutes it takes to ingest the smoke and the seconds after inhalation of the first puff when the complex chemosensory and neuro-active (e.g., cool menthol flavoring, other additives, nicotine, monoamine oxidase) chemicals hit the brain's reward/satisfaction pathways, to the months of experimental use and years of progression to established daily use, to the years and decades of sustained use, to the multiple cessation attempts or actual cessation.[22] This goes on within and across generations.[23] If nothing more, tobacco use provides a messy, sophisticated primer about biology and behavior in a societal and global context.[16]

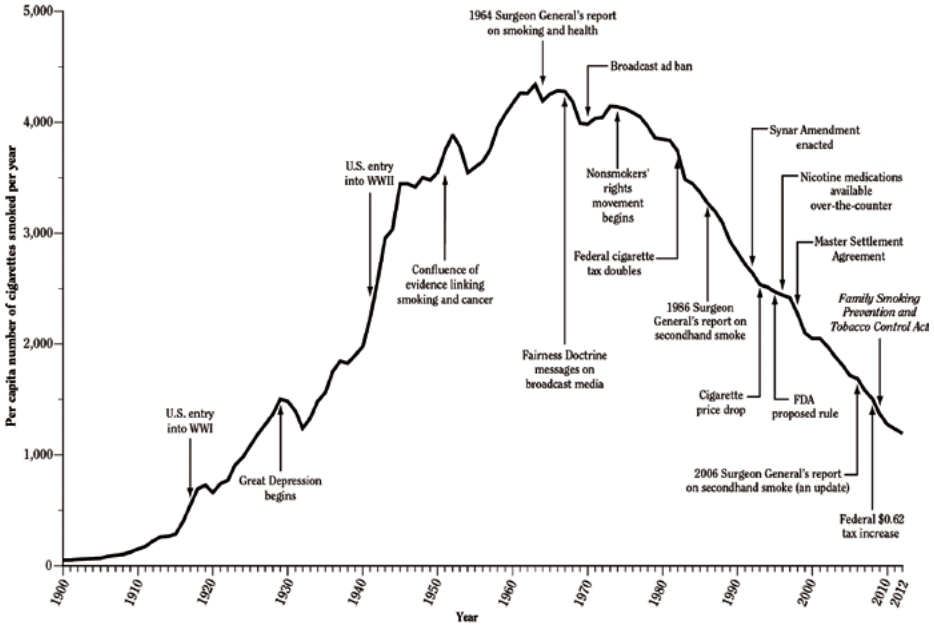


Figure 1. Adult per capita cigarette consumption and major health events, United States 1900-2012

Source: The health consequences of smoking—50 years of progress. Atlanta, GA: Centers for Disease Control and Prevention; January 2014.

For any number of reasons, society continues to tolerate the sale of cigarettes. There is a product for every taste, social class, pocketbook, and group/tribal identity. Lifespan developmental psychology has also contributed to our understanding of bio-behavioral and now epigenetic factors associated with initiation and persistence of smoking in some young people but not in others. Many risky behaviors are experimented with during youth, but thankfully most people grow out of extreme risk-taking behaviors.[24-27] But, the targeted marketing, appeal satisfaction, self-identity formation, and addiction to the nicotine in cigarettes lead to a journey and a lifelong struggle ending in disability and premature death. Given the premature deaths, industry spokesmen for tobacco use have said that youth are their "replacement customers."[28]

Contextual factors also play out over time and interact with individual factors. Human ingenuity, along with other events, have transformed patterns and norms of behavior both rapidly (e.g., world wars) and slowly (e.g., identifying cigarettes as the cause of lung cancer). Leaders and regulators worldwide have in effect also favored a concentration of powerful corporations with huge economic might; profits have overruled the priority for people's quality of life. Creation and maintenance of consumer demand also follows human nature's propensity (and vulnerability) to consume and to seek out satisfying immediate and relatively inexpensive "pleasant" experiences orchestrated by marketing and creating consumer demand.[29]

While this description is oversimplified, the cigarette as a device for combusting of tobacco yields a unique chain of stimuli and responses culminating in that nicotine and chemosensory satisfaction engineered together with marketing, packaging, placement, and cost. The cigarette has, perhaps more than any other product ever developed by humans, managed to slip through every layer of the multiple levels of the cells to interlaced societal systems that interact and result in specific patterns of individual and collective behavior. The cigarette has proven to be an ideal vehicle that conspires with micro and macro socio-behavioral and economic forces that capitalize on the vulnerability inherent in the human "limbic brain."

The limbic brain embodies a core of powerful drive and control systems that interact with the prefrontal, sensorimotor, and visual cortex for the ultimate goal of survival (e.g., the drive to seek food and water and belong to a tribe). The same areas of the brain involve the seeking of pleasure and the avoidance of discomfort or pain (in adolescents especially via emotional "hot cognition/impulsivity"). Other factors that include the pleasure of social bonding and feeling of belonging to a group influence behavior. The way cigarettes influence biological and behavioral pathways of seeking and craving to repeat a highly rewarding experience, in some respects, overrides rationality in critical decisionmaking contexts.[30] There are also vast individual differences driven by genetic and epigenetic factors that determine why some individuals rapidly develop an addiction while others can take or leave the cigarette at will, and why still others can use cigarettes for a period of time and then stop (e.g., during adolescence and early adulthood), and still others find the product to be aversive and cannot use it at all.

The influences on behavior have become more immediate and direct and can now rapidly spread through society by the emergence of the ubiquitous digital and social media environments. Thus, the brain maintains and adapts to change via multi-level proximal and distal "contextual factors" tied to social networks of family, peers, and members of communities to sustain cigarette use.[31] No one dimension adequately and rationally explains the dominance of the cigarette for over 100 years, but each piece accumulates and interacts with other components to produce the present status quo.

The Rise of Cigarettes and the Start of the Disease Epidemic

In this section, we will examine macro trends in tobacco use behavior and associated mortality from the perspective of human ingenuity, disruptive technology, mass production, marketing, and unintended consequences (Figures 1 and 2).

In the 1880s, the cigarette was an expensive, hand-made luxury item used by few. The vast majority of tobacco was consumed in non-combusted forms (e.g., snuff or chewing tobacco), and lung cancer was an extremely rare disease. An automated cigarette rolling machine, invented by 18-year-old James Bonsack, revolutionized production and proved to be a truly disruptive and ultimately lethal technological innovation.[32] The machine produced 70,000 cigarettes over 10 hours in its trial run and rolled what 48 employees could roll by hand in a day. Ironically, another technological innovation, the refrigerator, was adopted by almost all households between 1930 and 1980 and, without much other intervention, reversed the then devastating toll of stomach cancer on the U.S. population.

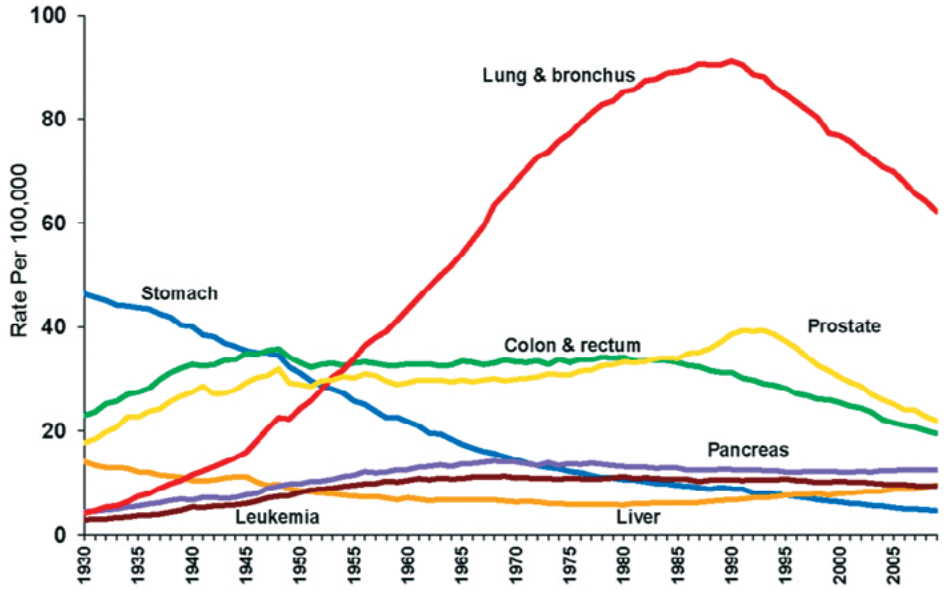


Figure 2. Trends in age-adjusted cancer death rates by site, males, U.S., 1930-2011

Source: US Mortality Volumes 1930 to 1959, US Mortality Data 1960 to 2011, National Center for Health Statistics, Centers for Disease Control and Prevention. ©2015, American Cancer Society, Inc., Surveillance Research. Used with permission.

Notes: Cancer deaths are adjusted to the 2000 U.S. standard population. Mortality rates for pancreatic and liver cancers are increasing. Due to changes in International Classification of Diseases (ICD) coding, numerator information has changed over time. Rates for cancer of the liver, lung, and bronchus and colon and rectum are affected by these coding changes.

Among men, for example, the two prevalence curves of lung cancer and stomach cancer crossed over in the 1950s. Lung cancer went from a rare cancer in 1930 to the most prevalent cancer by 1990, and at its peak, the death rate for lung cancer (per 100,000) was about twice as high compared to stomach cancer at its peak (Figure 2). Stomach cancer was as rare in 1990 as lung cancer was in 1930, both changes resulting from population-level adoption of disruptive technology within 60 years. Most of the increase in overall cancer death rates for men prior to 1990 was attributable to the rapid increase in lung cancer deaths due to the tobacco epidemic. However, since 1990, the lung cancer death rate in men has been decreasing. This decline has accounted for nearly 40 percent of the overall decrease in cancer death rates in men.[33] Even our most conservative estimate indicates that reductions in lung cancer resulting from reductions in tobacco smoking over the last half century account for about 40 percent of the decrease in overall male cancer death rates and have prevented at least 146,000 lung cancer deaths in men during the period 1991-2003. The death rate for stomach cancer, which was the leading cause of cancer death among men early in the 20th century, has decreased by 90 percent since 1930.

At first, adoption of cigarette smoking by the population was slow; companies were reluctant to invest in the cigarette rolling machine, believing consumers preferred hand-rolled cigarettes, cigars, snuff, and chewing tobacco (Figure 3). It is important to note that while they are not harmless,

of course, these non-combustible forms of tobacco (snuff and chewing tobacco) dominated total tobacco sales in the late 1800s, and their use clearly did not cause the devastating diseases associated with pulmonary inhalation of toxic smoke (cardiovascular disease, lung cancer, and chronic obstructive pulmonary disease) that account for the vast majority of preventable deaths just 100 years later.[34,35]

This issue of using less harmful forms of tobacco—like new versions of the smokeless products that dominated in the 1880s—to obtain the nicotine without the lethal tars of combustion has taken center stage again in the early 21st century (discussed in more detail later in this chapter). Early proponents of harm minimization first suggested going back to the use of smokeless products to accelerate stalled reductions in cigarette use prevalence in the 1990s. Harm minimization has, with the introduction of e-cigarettes as a potential disruptive technology, now become perhaps the most hotly debated issue seen since the cigarette rolling machine produced the epidemic in the first place.[36-38] The invention and rise of the e-cigarette and the use of low nitrosamine Swedish snus to reduce the harms from inhaled cigarette smoke in Sweden[35,39] suggest a "back to the future" scenario where cigarettes are made obsolete by new non-combusting technology yielding use of safer tobacco-derived nicotine delivery not seen since the 1890s (Figure 3).

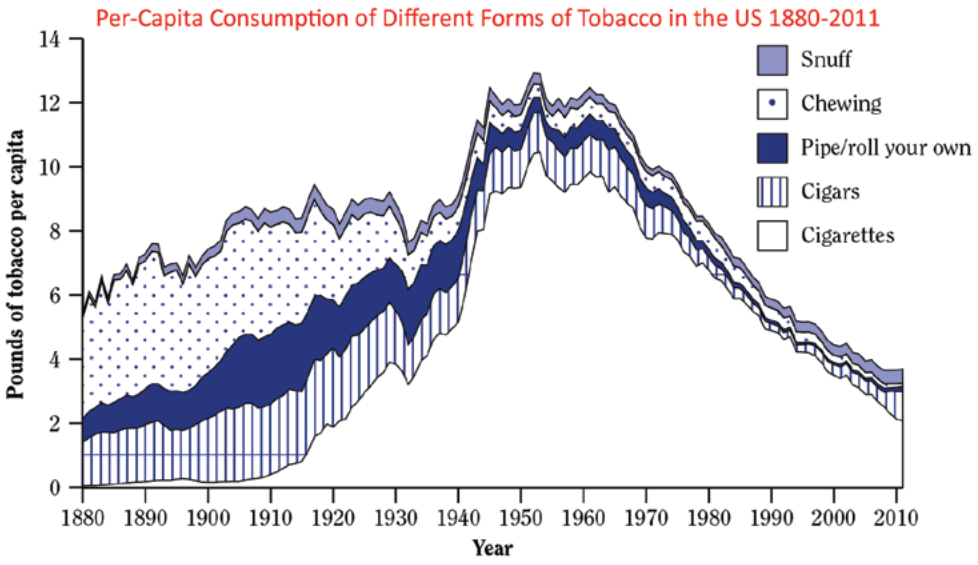


Figure 3. Trends in type of tobacco use

Source: U.S. Department of Treasury, 2012.

Within two decades of the invention of the rolling machine, however, the volume of sales catapulted to 10 billion by 1910, and the retail price was halved. American Tobacco took advantage of the new technology so successfully that it was broken up in 1911 because of U.S. antitrust law. The uptake in patterns of adoption was driven over 50 years, first by predominantly male smoking, the two World Wars, and a decline during the great depression (see macro trends and key events in Figures 1-3), and then by increasing marketing and adoption by women beginning with the infamous Torches of Freedom public relations (now termed marketing) campaign, still cited as a classic in marketing

textbooks.[40,41] Smoking prevalence by 1964, when the first Surgeon General's warnings were issued, exceeded 50 percent of the male population, but by the 1980s, smoking rates were almost identical for men and women.[42-44]

Consequences of Tobacco Use

By the early 20th century, articles addressing the health effects of smoking began to appear in medical journals. In 1930, researchers in Cologne, Germany, discovered a statistical correlation between cancer and smoking. In 1938, Dr. Raymond Pearl of Johns Hopkins University reported that smokers did not live as long as non-smokers. By the 1950s, the evidence mounted that smoking (i.e., inhaled smoke from the burning or combusting of tobacco) had caused the epidemic in lung cancer; this rarest cancer in 1900 exploded to overtake all other cancer death rates by 1955 and progressed to double any other cancer in death rate, responsible for over 30 percent of all cancer deaths by 1990 (Figure 2). These findings led to the landmark 1964 U.S. Surgeon General's report[1] and a dramatic tipping point in the reversal in cigarette sales and smoking prevalence from over 50 percent in men to less than 20 percent in 2012.[2]

It is at this point that the story becomes quite complicated, pushed along by a range of forces working to promote cigarette use while others pushed against it. During the past 100 years, the basic structure and function of the cigarette have remained relatively unchanged in many respects, except for the optimization of the ingredients to enhance chemosensory appeal, satisfaction, and the rapid pulmonary absorption of smoke and nicotine to reach the brain's reward pathways within 7 seconds, thus ensuring the highest degree of addiction and enhanced performance.[45] Over the decades, tobacco companies have masterfully re-engineered, optimized, innovated, adapted, and brilliantly packaged, priced, and marketed cigarettes. In the process, they have created brand loyalty and consumer demand for a defective product that nobody really "needed."

In the last 50 years in particular, the cigarette industry and marketing have withstood massive challenges from numerous sectors of society. These include public education campaigns about the harms of cigarettes, U.S. Surgeon General's Reports (1964-2014), bans on television advertising, restrictions on sales and marketing to children, State and Federal taxes and clean air laws, a U.S. Master Settlement Agreement of over $250 billion, being adjudicated by U.S. Federal court to have engaged in fraudulent behavior, and now being regulated under the 2009 U.S. Tobacco Control Act.[2] However, everything done so far by tobacco control, although resoundingly successful from the 1960s to the 1990s, has reached diminishing returns. The resilience and sustained market share of cigarettes continues globally. If nothing disrupts the reign of the ubiquitous cigarette over the previous century, by 2100, this mass-produced retail product will prematurely kill 1 billion people worldwide. The cigarette has survived and thrived despite being clearly labeled for decades as lethal: "…a defective product – unreasonably dangerous, killing half its users and addictive by design."[12,14]

What Makes Cigarette Smoking Dangerous: Nicotine vs. Tars

Tobacco has been used for millennia without wreaking the harm that it now does on population health. Nicotine is neither necessarily very harmful nor very addictive; it depends on the mode

of delivery (e.g., rapid pulmonary absorption from combustion or slow infusion via a transdermal nicotine patch) and the dose. Nicotine, per se, is therefore not the enemy, although not harmless, especially to some vulnerable populations such as the unborn fetus, young developing brains, and when delivered in modalities that increase its addiction liability.[45] The enemy is the combustion or burning of tobacco that produces the gas phase tars, toxins, and carbon monoxide that kill, while also delivering nicotine in its most addictive form. As Russell succinctly stated almost 40 years ago, "people smoke for the nicotine but die from the tar."[46] Not surprising, then, it was not until the early 1900s that the harms from mass production and marketing of cigarettes exponentially escalated to epidemic proportions and became the leading cause of over 480,000 preventable premature deaths annually.[2,35] Cigarettes have approximately 600 ingredients; when burned, they create carbon monoxide, particulate matter, and more than 7,000 chemicals, commonly referred to as tars. At least 69 of these chemicals are known to cause cancer, and many are poisonous. This toxic mix and exposure to it for years cause the overwhelming majority of preventable morbidity and mortality from inhaled smoke.[35,47]

It is puzzling that a product known for over 50 years to be so lethal has escaped either being taken off the market by State or Federal governments empowered to protect public health or being limited by the companies themselves in line with how other industries recall defective products (e.g., Ford Explorers and Firestone Tires) or risk severe consequences and loss of reputation if they do not—as in the recent debacle of the decade-long use of a defective ignition switch in General Motors automobiles that appears to be responsible for at least 13 deaths.[48-50] These contrasting scenarios make for an interesting lesson for the social and behavioral sciences by asking: How do these contradictory yet obvious injustices become accepted in a civil society and come to live side by side?

New breakthrough ideas and strategies are urgently needed to disrupt this human created and "systems within systems" epidemic. An epidemic that itself was begun with a "disruptive technology"—the invention and patenting of the cigarette rolling machine—can perhaps be halted through the introduction of another disruptive technology that will alter the way in which nicotine is delivered to the brain. Simulations of plausible futures indicate that even if policies are aggressively implemented, reductions in the prevalence of cigarette smoking will be frustratingly slow and reductions in smoking-related deaths even slower.[6,7] We cannot accept a second cigarette century. As articulated in the 2014 Surgeon General's 50th anniversary Report: "Death... is overwhelmingly caused by cigarettes and other combustibles... promotion of e-cigarettes and other innovative products is... likely to be beneficial where the appeal, accessibility, and use of cigarettes are rapidly reduced" (pp. 15-17).[2] The ultimate goal should be to use safer forms of nicotine delivery to speed the obsolescence of lethal combustibles for those who cannot stop using some form of nicotine, while ensuring that new products are kept out of the reach of youth as much as possible.

Harm Minimization and the Rise of the E-Cigarette

Is There a Safe Enough Replacement for the Lethal Cigarette?

Harm minimization, as a tobacco control strategy, includes achieving a society that rejects all tobacco use as the first priority. However, harm minimization recognizes that, pragmatically, some youth and adults will become addicted to combustible products/cigarettes that cause the greatest

harm by far. Thus, a second priority is to focus on eliminating cigarette/combusted product use immediately. Obtaining nicotine without burning tobacco is safer and can be made less addictive. For anyone who cannot stop all use of combustibles/cigarettes, a harm minimization approach is a viable option. Harm minimization adopts a continuum of risk, which places cigarettes (or any burning of tobacco) at one extreme (most harm) and FDA-approved nicotine replacement therapy (NRT) at the other extreme (least harm).[51]

Electronic cigarettes ("e-cigarettes")—or electronic nicotine delivery systems (ENDS)—lie along this continuum. In 1963, Herbert A. Gilbert was granted a patent to produce a prototype of an e-cigarette that delivered nicotine through heated steam.[52] E-cigarettes were not commercially viable until Hon Lik, a Chinese pharmacist, developed a more advanced product in 2003 and received an international patent in 2007.[52] Since then, the products have evolved from disposable, "cigalike" products to larger rechargeable products, also known as "personal vaporizers." These devices deliver an aerosol mist from liquid containing tobacco-derived nicotine, flavorings, and other ingredients.[53,54]

Evidence from user surveys and interviews indicate that new e-cigarette users begin with "cigalike" products, but experienced and long-term e-cigarette users (i.e., "vapers") appear to transition to larger models (tank-like or open systems) and prefer to order nicotine liquid online/purchase from "vaping stores" or mix their own concentrations.[55-59] There is wide variability in nicotine content vaporization efficacy by brand and potential delivery to the user,[60,61] but exposure to toxicants is significantly lower for e-cigarettes than for conventional cigarettes.[62-71] E-cigarettes are generally perceived to be less harmful than regular cigarettes[56,72-74] and are commonly used as a smoking reduction/cessation aid.[55,72-77]

National cross-sectional data in the United States indicate that trial of e-cigarettes has increased in both youth and adult populations, from 4.7 percent in 2011 to 11.9 percent in 2013 among high school students,[78-80] from 6.9 percent in 2011 to 7.8 percent in 2013 among young adults ages 18-24,[81] and from 6.2 percent in 2011 to 8.5 percent in 2013 among adults 18 and older.[81,82] Current use has also increased, but most recent available data show that overall prevalence remains low: 4.5 percent (2013) among high school students, 2.4 percent (2013) among young adults, and 1.9 percent (2013) among adults overall.[76,83] The majority of ever and current e-cigarette users are current cigarette smokers. Among youth in the United States in 2012, 76.3 percent of current e-cigarette users were also current cigarette smokers.[79,80] A national survey of U.S. adults found only a 1.2 percent trial of e-cigarettes among never smokers in 2013, compared to a 36.5 percent trial among current smokers.[81] However, trends could be evolving, with the Monitoring the Future Study finding in 2014 that 17 percent of youth were trying e-cigarettes, also with 75 percent of them already using cigarettes.[84]

Despite considerable promise, concerns about e-cigarette use largely focus on the potential that they will serve as a new entry point into other tobacco products, particularly among youth, and that they will delay or halt cessation altogether via prolonged dual use of e-cigarettes and cigarettes without reduced harm exposure. While studies of youth have advanced the notion that use of e-cigarettes may encourage cigarette use,[85] exploration of alternate hypotheses must also be considered.[37,86,87] As noted by Niaura and colleagues, "It is equally plausible that use of combustible cigarettes

leads to use of e-cigarettes because they are perceived as a less harmful alternative for smokers who are addicted to nicotine. The cross-sectional survey data do not prove that this is the process that explains the association, but they are just as consistent with it…"[87] Additionally, there is very limited evidence from the longitudinal observational studies to determine how e-cigarette use influences other patterns of tobacco use,[77,88] and this is further complicated by the low population prevalence of e-cigarette use and limitations of the selected nature of the populations included in observational studies. Evaluating the potential impact of e-cigarette use on youth cigarette uptake, for example, would require identification of youth who would start with e-cigarettes and move to combustible products over and above those youth who would smoke combustible products anyway. In short, the data required to document potential harms (or benefits) of a specific product at the population level will be difficult to acquire. However, the increased prevalence of e-cigarette use in U.S. youth is occurring at the same time when overall smoking by teens showed a decrease, from 10.6 percent in 2012 to record lows of 8.0 percent in 2014.[89] Thus, it appears that e-cigarettes may not be promoting more cigarette use or delaying the cessation of cigarette use.[90,91] Thus, it does not appear to be the case that increased e-cigarette use is leading to more conventional cigarette use.

A key question regarding e-cigarettes is their potential role in effectively facilitating cessation of combustible cigarettes in current smokers or significant harm reduction, both in individual users and in terms of overall population impact. There have been several studies on the use of e-cigarettes to quit smoking combustible cigarettes. Two randomized controlled trials have been conducted, indicating that e-cigarettes are effective in helping some adult smokers to quit or to reduce their cigarette consumption.[92,93] Other prospective observational studies with comparison groups in population samples report any use of e-cigarettes (it is not clear in some studies that e-cigarette use was expressly for the purpose of trying to quit) may be associated with no change in smoking status or a reduced likelihood of cessation.[74,94,95] Thus, these negative associations are impossible to interpret and may be due to other factors, especially selection bias (e.g., smokers who are less able to quit but more motivated to try are more likely to experiment with e-cigarettes out of curiosity or to have more difficulty with cessation).[96] A longitudinal study that established use of e-cigarettes as part of a serious quit attempt found that in two U.S. metropolitan areas, after 2-3 years, intensive e-cigarette users (used daily for at least a month) had the highest rate of smoking cessation (20.4 percent), compared to intermittent users (more than once or twice but not daily for a month or more; 8.5 percent) and triers (used at most once or twice)/non-users (12.4 percent).[97] Smokers who used e-cigarettes for at least a month were six times as likely to be abstinent at 2-year followup as triers/non-users. Another large cross-sectional survey in the United Kingdom found that among adults who had smoked within the previous year and made at least one quit attempt during that period, e-cigarette users were more likely to report abstinence than either those who used NRT or no aid.[98] It has also been noted by these authors that in the United Kingdom, e-cigarette use for cessation has surpassed use of NRTs, thus making their potential for larger scale population impact quite promising. More research – especially independent, high quality randomized controlled trials with appropriate control groups – is needed to further determine whether and how e-cigarettes can be an effective cigarette cessation or harm reduction aid; nevertheless, results to date are promising for some adult smokers.

Whether e-cigarettes can help people quit combustible cigarettes may also depend in part on their effectiveness at delivering nicotine at or near the level of cigarettes, as well as their overall satisfaction for users, based on the complex chemosensory experience that cigarettes now provide.

Acute examinations demonstrate that nicotine delivery is dependent on the e-cigarette device and liquid concentration, as well as the rate at which the nicotine is delivered and the user's experience with e-cigarette use (i.e., naïve or not naïve).[99-105] As mentioned previously, more experienced users prefer larger personal vaporizers over the "cigalike" disposable devices. This is likely because, as studies have demonstrated, there are higher plasma nicotine levels after use of these new generation products, approaching the levels achieved after combustible cigarette use, although the levels peak after a longer duration of use.[59,101] E-cigarettes have also been shown to reliably decrease adverse symptoms related to tobacco abstinence (e.g., cravings and urges to smoke, irritability).[99,100,102,103,106,107]

More research is needed to determine whether the levels of nicotine delivered support the potential for e-cigarettes to be sufficiently satisfying to displace combustible tobacco products and to determine what level of nicotine dependence may be acceptable at the population level. It is also unclear whether a cleaner form of nicotine, while not harmless and still potentially addictive, warrants as much public health concern if the nicotine delivery is de-coupled from the toxicity in combusted products (i.e., a delivery system that is as satisfying as a cigarette and may have addiction liability but confers significantly less harm than cigarettes). The net public health benefits versus harms would need to be determined by the degree to which the nicotine delivery system can successfully compete with combusted tobacco. That is, the benefits of a product with high addiction liability and with low harm (associated with cleaner nicotine devices) would outweigh harms of the high addiction liability provided that product strongly encouraged complete switching away from combustible cigarettes (a public health benefit if the nicotine addiction liability is compared to lethal cigarettes and not to a placebo or nothing). This would contrast with a lower addiction liability product that resulted not in complete switching but rather prolonged dual use.

Clearly, the emergence of new products like e-cigarettes raises a critical issue regarding whether society is now willing to tolerate a safer form of what might be termed "recreational" nicotine use, even if it has some or even quite a large addiction liability for some users, in order to speed the obsolescence of cigarettes. If so, then for the first time in over 100 years, it may be possible that the e-cigarette, a disruptive technology, could dethrone the cigarette rolling machine that has dominated since the 1880s and fueled the global multinational tobacco industry. Could this be a David and Goliath moment?[37,108]

Interpreting the Science to Improve Public Health While Minimizing Harms

Whether e-cigarettes are a disruptive technology with the potential to completely replace conventional cigarettes depends on consumers, how they perceive and use new and emerging products, how the products are regulated, and how they are made and marketed, both by tobacco companies (who have a vested interest in dual or poly use of all their products, including lethal cigarettes) and by independent manufacturers who do not sell other tobacco products and thus have less conflict and are able to focus on making cigarettes obsolete.[37,52,86] This complex set of dynamics in a rapidly changing landscape calls for a systems science view of the challenges involved (more on this later in the chapter). The U.S. Food and Drug Administration (FDA) Center for Tobacco Products (CTP) regulates tobacco products using a variety of perspectives to inform a "public health

standard" that requires the regulatory agency to weigh the harms versus benefits of products to both individuals and the population as a whole.[109,110] The "public health standard" also must consider several "likelihoods" of (1) beginning use of a product; (2) preventing, delaying, or accelerating cessation of the use of combustible tobacco; and (3) in the case of e-cigarettes, promoting or discouraging product switching or dual use with no reduction in combustible cigarette use.[109]

Simulation modeling provides an opportunity to understand the multitude of influences of e-cigarettes on population health. The proof-of-concept study by Cobb et al.[109] used a Markov model to make explicit the dynamic states and transitions of use over time, using data from a longitudinal cohort of U.S. young adults. The authors found that the largest changes over 10 years occurred in non-current use, cigarette use, and former use states, with dual use and e-cigarette use states remaining at low prevalence over the years; these findings demonstrate a projected minimal impact of e-cigarettes on patterns of cigarette use. This type of information could be used to guide regulatory policy through informing the "likelihoods" in considering whether e-cigarettes, or other potentially reduced-harm products, uphold the "public health standard."

In terms of economics, an argument can be made that companies may still want to market and addict generation after generation of youth; an outcome that should be avoided even if the products are much safer than cigarettes. This may be true of tobacco companies, since only the withdrawal of their lethal combusted products from the marketplace would eliminate the risk of future use among youth. Sweanor[52] and other harm minimization proponents argue that in the case of cigarettes alone, smokers worldwide currently pay the equivalent of roughly $800 billion a year (in U.S. dollars) on a product many are aware is defective and likely to kill them. A non-combustible product line, for which there is already consumer demand among current smokers, could capture an increasing fraction of the enormous cigarette market, thus presenting a huge business opportunity while saving millions of lives and without needing to recruit another generation of adolescents or naïve users of any age. This incentive would be more appealing to independent e-cigarette makers, who if they acted in a responsible fashion, could drive a wedge between combustible and non–combustible product use and accelerate the demise of the lethal cigarette. Responsible behavior (i.e., not targeting youth or marketing in a way that appeals to youth) could be supported by tobacco control, FDA CTP, regulators and other policymakers. E-cigarettes could deliver enormous gains for public health if the tobacco companies would also take the initiative to withdraw combustible products and act in a trustworthy manner as stewards of public health rather than just for shareholder profits. Then, governments could work with rather than against entrepreneurs and collaborate on a pragmatic harm minimization approach.

The introduction of new products, along with changes in marketing and regulation, traditional tobacco control policies, and individual and group behavior represent a dynamic and complex interaction that could have a number of potential unintended effects. For example, this interaction could undermine efforts to end the tobacco problem or it could hasten efforts to eliminate cigarettes. It is possible that new tobacco products may reduce consumption and population harms by displacing combustible tobacco products. On the other hand, the new products may increase harms by promoting the use of multiple products or delaying cessation.[111]

At present, the tobacco landscape continues to evolve. The global and U.S. tobacco industry has stated in various ways that it plans to undergo a major paradigm shift towards making and marketing a wider range of tobacco-derived nicotine products with a purported mission of reduced harm. Innovations in the new products pipeline and the creative marketing that will be emerging during the next decades of the 21st century focus on non-combustible, next-generation products, especially more efficient pulmonary delivery of aerosolized nicotine, which is claimed to be more satisfying and more efficient at delivering "clean(er)" nicotine than the current e-cigarette. These game-changing moves will likely blur historic boundary lines drawn over the last 50 years between the tobacco control community, the tobacco industry, and the pharmaceutical industry.

Changes to product lines set forward in the tobacco industry playbook, if they come about, almost certainly will alter the tobacco control landscape, as well as individual and population patterns of uptake, regular use, and even the existence of tobacco-derived nicotine. The implications for tobacco control, in general, and FDA regulation, in particular, are largely unknown and potentially profound. These trends and marketing practices must be carefully tracked, and their impact on the knowledge, attitudes, risk perceptions, and behavior of consumers must be rapidly ascertained to inform regulation and education using the public health standard mandated in the Family Smoking Prevention and Tobacco Control Act.[112-114]

Another consequence of these game changing moves will likely blur the previously clear boundary lines drawn over the last 50 years between: (a) tobacco control as total abstinence from all tobacco and nicotine products; (b) a new set of potentially appealing harm-reduction options for consumers that also can carry risks of harm escalation or of slowing the progress made towards eliminating the epidemic; and (c) the traditional and previously clear-cut use of medicinal nicotine replacement therapies (NRT), either for short term use to achieve complete desistence of all combustible tobacco products and nicotine or as a long-term alternative to continued use of the most harmful combustible tobacco products.

In the next 50 years, we may well be in an era of paradigm-shifting perspectives and a shortened timeframe. The viral spread and dramatic penetration of e-cigarettes in just 5 years, from 2007-2012, is a clear example.[115-117] As we look to the future through the lens of a new world of rapid digital communications and innovation of products to deliver nicotine in non-combustible form, whether for recreational use or as a therapeutic smoking cessation treatment, some informed speculation is inevitable as we conclude this chapter. The lessons of the last 50 years may no longer be the best predictor of the next 50 years, but lessons learned, better conceptual and analytic lenses, real-time data gathering capacity, and the FDA's regulatory authority all provide an evidence-based context to provide a sharper vision of what may be possible in the future.

In examining possible future directions for tobacco control, there are new uncertainties. In the future, we may not be able to rely on the past as a guide in developing recommendations due to the rapidly changing marketplace, the emphasis on digital and social media, tobacco product regulation by the FDA, and the industry's movement towards championing reduced-harm products and messages. We are entering a new era that, ironically, is fueled by similar dynamics that started the current epidemic a hundred years ago: new technology (spawning new modified risk/non-combustible alternative nicotine delivery products), digital communication (Internet, social media,

smart phone) with new forms of rapid and targeted marketing and the viral spread of messages and products, and the new leverage of an FDA regulatory authority that could also change the dynamics in the future tobacco control context. Future strategies and projections of impact must now consider a more complex and dynamic model that addresses short and long-term feedback loops at multiple levels of influence. New vectors and pathways may have unpredictable intended and unintended consequences.[118,119] Messages about new and existing tobacco products continue to spread virally and be maintained in the social networks of users in both developed and developing nations. Moving stealthily, just as a natural disaster might do, the vector continues to spread the message, damaging and "infecting" more and more humans with projections of over a billion lives lost in the 21st century. All of this new damage can be prevented or attenuated, while the already existing devastation can and should be cleaned up.

Implications for Research and Policy

Implications for Research

While much has been accomplished to date, there continues to be room for additional improvement in the current status of the field. As we have discussed in this chapter, there are significant if not groundbreaking changes occurring in the landscape of tobacco control at the present time. Evidence-based tobacco control interventions are unlikely to provide the same dramatic reductions in tobacco use over the next 50 years compared to the past 50 years. Novel endgame strategies that have a strong empirical basis may be needed to reduce the use of cigarettes and other forms of tobacco. While these new strategies may be able to minimize the harms of tobacco use over a shorter timeframe, there are likely to be political and social challenges to implementation. As proposed in a special supplement of the journal Tobacco Control in 2013, potential endgame strategies include reducing the nicotine content in cigarettes, creating a tobacco-free generation, creating limits on the amount of tobacco available for commercial sale ("sinking lid"), banning sales of cigarettes, and a range of other strategies.[14,120-138]

In addition to evaluating the outcomes and barriers to such endgame strategies, we also need to consider what constitutes a final endgame: is it a reduction in smoking prevalence or smoking-related deaths to a low rate, or is it the end of all smoking?[136] This approach needs to be compared to (or combined with) the two other main solutions, sticking to traditional evidence-based tobacco control measures and harm minimization.[140] Of all these strategies, perhaps the ones with the most potential for population benefit are those that reduce the appeal, addictiveness, and toxicity of cigarettes and increase their cost to consumers via taxation, while simultaneously providing reasonably satisfying but much less harmful forms of nicotine delivery such as prudently regulated and FDA-approved nicotine delivery products like e-cigarette or future innovative variants.

The convergence of three major trends challenges the status quo and what we know to date about effective and efficient tobacco-control interventions for prevention, treatment, and policy, which will need to be adjusted in light of new industry and regulatory developments. The two major trends in the morphing of the tobacco industry towards manufacturing a panoply of diverse products with claims for possible reduced harm and the digital media revolution are overshadowed by the regulatory authority given to the FDA Center for Tobacco Products in 2009. It is too early

to say how and to what extent the regulation of tobacco products by the Government will impact the landscape in the United States.[140] What is clear is that we need a new regulatory science that incorporates these emergent trends, embraces a transdisciplinary "systems thinking" and multi-level Social Ecological Model as one useful guiding framework,[110] and from the very outset of designing the science, considers how the science can inform policymaking.

System sciences likely will play an increasingly important role as a tool to inform research, practice, and policymaking in tobacco control. Behavioral and social sciences researchers will continue to play a critical role, but they must adopt systems thinking and modeling tools to add to their ability to rapidly collect data that can keep pace with the changing landscape and consumer behaviors. In systems thinking, a coherent system comprises relationships that unveil "emergent properties" that are only visible at the system level and not by examining them in isolation.[15] At work within the system are dynamic, non-linear feedback loops, stocks and flows, and time delays, including biological, organizational, social, and political forces.[15,18,19] Stakeholder engagement and transdisciplinary thinking are necessary to understand and manage these complex challenges, and various methodologies can be employed, such as system dynamics simulation, agent-based modeling, network analysis, or Markov modeling to address the heterogeneity of the system, especially when real-world experiments are not possible or ethical.[15,18,19]

The Institute of Medicine's report, A Blueprint for the Nation, and the U.S. Department of Health and Human Services' Strategic Action Plan recommend simulation modeling to project the impact of integrating tobacco control interventions and policies, given the dynamic nature of the tobacco epidemic and its drivers.[2,141,142] Figure 4 depicts a systems view of some of the major competing influences on patterns of tobacco use at a coarse level of granularity. In most cases, the tobacco industry strategies increase harmful patterns of tobacco use and have a health-damaging impact (red feedback loops) and the counter-tobacco strategies reduce tobacco use prevalence and have a health protective influence (blue feedback loops). The mechanisms of influence include tobacco industry anti-tobacco control strategies such as marketing, innovation of new products, litigation, and challenging of restrictive policies and higher taxes, whereas pro-tobacco control mechanisms include tobacco product regulation, restrictive tobacco control policies, mass media public education, State and local ordinances, tobacco control litigation, and prevention and cessation interventions. The systems diagram also provides a roadmap to examine the success of current tobacco control efforts and identify the policy and program levers that could dramatically reduce tobacco use and thus, tobacco-related morbidity and mortality and economic costs.

The problem then emerges from the reciprocal influences of multiple "systems within systems" that interact dynamically within and across many levels, ranging from molecular, neurobiological, and bio-behavioral systems within the individual to individual behavior and group-level reciprocal interactions that play out at cluster or aggregate levels, such as family, peer, neighborhood, community, and societal and global levels of influence.

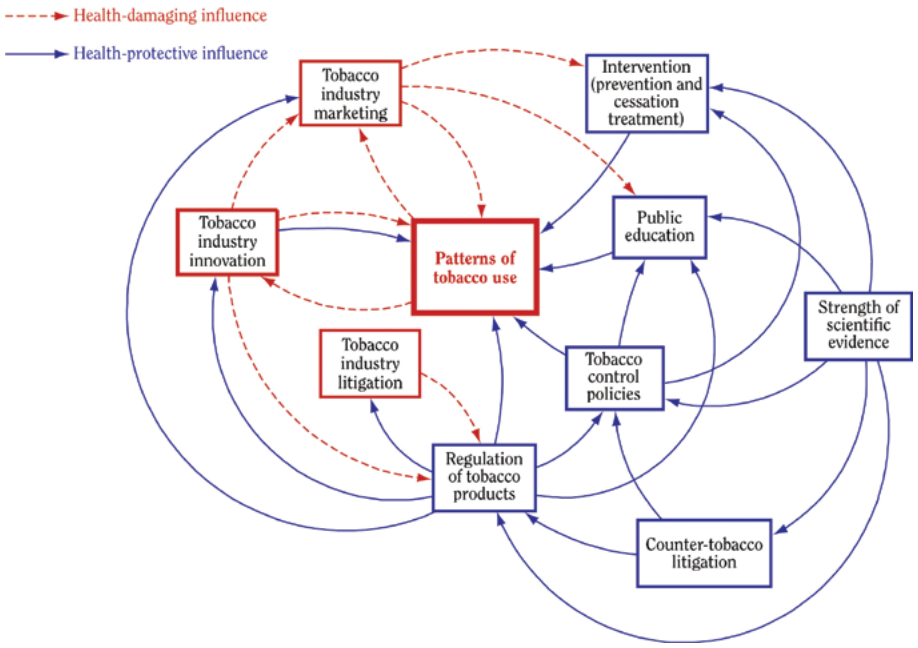


Figure 4. Simplified dynamic model of protobacco and antitobacco forces on patterns of tobacco use

While still a work in progress, advances in engineering such as control systems and simulation models have advanced in the last decade. This type of science calls for integrated and transdisciplinary synthesis by scientists, advocates, and legal, political, government, and non-government stakeholders.[110] This must be done within the context of a dramatically changing environment that includes the tobacco industry, marketing, and other stakeholders who are anticipating and reacting to the FDA's new authority. Now more than ever, science is needed to evaluate how industry efforts impact youth and adult tobacco use patterns, but it must be rapid, rigorous, and flexible to keep pace with the changing market. Research will be critical to help integrate new prevention and cessation interventions, and there is a pressing need to review and evaluate on an ongoing basis FDA CTP's performance to optimize its impact on reducing population harms from tobacco products. New methods of surveillance, including methods involving new technologies and shorter intervals between data collection, will be an essential complement to annual population-level surveys.

Given declining funding for tobacco prevention efforts, there is also a great need to optimize program effectiveness; this includes demonstrating the utility of new evaluation tools, potential products, and new channels and platforms of intervention delivery. Tobacco control interventions and policies must cover the rapidly growing digital environment and remain in-step with evolutions in social media.

The simultaneous opportunity and challenge for tobacco control researchers will be to conduct rigorous science that reduces the harms of tobacco use in the context of what we have referred to throughout this chapter as a series of "perfect storms," including the proliferation and ubiquity of digital and social media, a rapidly morphing tobacco industry with a panoply of new products and marketing tactics, and the dawning of the era of FDA-CTP regulatory authority over tobacco products. This perfect storm requires a strategic approach to science to determine the optimal leverage points for targeted intervention and policy changes.

Implications for Policy

Today's changing landscape includes vectors that can influence the status quo and where we go. For example, FDA can regulate existing and new products, and they can educate the public to reduce the death, disease, and economic burden associated with use of tobacco products. Thus, defining the goals of ending the tobacco problem at the outset is central to identifying the surveillance and intervention tools needed to reduce the death, disease, dollars, and disparities associated with tobacco use.

To date, Federal, State, and local policies shown to be effective in reducing tobacco use exhibit uneven implementation. In 2012, tobacco prevention spending was less than 2 percent of tobacco revenues across all States, State excise taxes on non-cigarette products remained markedly lower than cigarette excise taxes, and approximately two-thirds of Americans were not protected by State or local clean indoor air laws. It is our hope that health care policies like the Health Information Technology Economic and Clinical Health Act and the Affordable Care Act will be more effective than past policies, and that instead, they may help us to meet the Healthy People 2020 objectives of increasing tobacco use screening and cessation counseling in health care settings.

Development in 2010 of the Department of Health and Human Services Strategic Action Plan for Tobacco Control[142] provided a blueprint for coordinating a comprehensive tobacco control program in the United States across Federal, State, and local partners, including supporting FDA regulation of tobacco. Increased funding for these efforts is essential if we are to implement a coordinated tobacco control response at these multiple levels. Further guidance and action at the Federal and State levels are needed to ensure that tobacco-related benefits are consistently implemented across insurers.

There is an increasing interaction between industry, public health advocates, and the FDA that could blur the distinctions and the typical ways the public understands the intended and unintended consequences of the changing vectors, and especially the implications for how changes at individual levels may influence group-level and population-level impacts. Thus, it is unclear how the FDA standard of determining the population level benefits or harms will be operationalized and used, especially in the new marketplace of tobacco products and claims.

The FDA's bottom line is to employ science-based rule making by determining the impact of any regulation or educational program at the population level, taking into account both users and non-users. Its mission is to reduce the deaths and harms that result from use of tobacco products. The FDA's recent regulatory tools provide additional opportunities and strategies that are consistent

with the goal to end the tobacco problem or to change it into a less damaging problem without eliminating all tobacco product use per se. New strategies described as "endgame strategies" may offer a tremendous opportunity to transform the tobacco epidemic, though that will require strong scientific support and political will.

Conclusion

This chapter oversimplifies the achievements and challenges in tobacco control. The fact remains that to date, despite the wealth of evidence accumulated and the progress made in the 50 years now covered by the Surgeon General's reports, we as a society have not yet chosen to prohibit the manufacture, sale, and distribution of lethal tobacco products to adults, and we also have not been fully successful at keeping these products away from our children. The tobacco industry itself has not volunteered to take its lethal products off the market the way that the Firestone tire company and Ford Motor Company reacted when it became known that several hundred excess deaths were associated with the defective combination of Firestone tires on Ford Explorers.[48-50] We see in our ongoing story a variety of political, cultural, economic, and ideological reasons not to take defective tobacco products off the market. Even for the sake of both adult users and our vulnerable children who will make up our future generations, and despite the tremendous harms caused primarily by using combustible tobacco products (the lethal combination of burning tobacco producing an addictive pleasure within toxic inhaled smoke), it is unlikely or impossible to envision that America will take these dangerous combusting products off the market any time soon. Some might argue that this century-old American story of creating an unprecedented disease epidemic, followed by partially ameliorating it with a 54 percent reduction in prevalence, is among the most notable human engineered mass disasters and massive ongoing cleanups in America's recorded history. There is more work to be done as we review current status and future directions.

A renewed call to action through the lens of social justice is needed now more than ever.[143] This is not only for the sake of the current generation of adult tobacco users and for the current generation of vulnerable children, adolescents, and young adults who are potential users, but also especially for priority populations in whom cigarette use is highest (i.e., economically disadvantaged people and ethnic and racial minorities),[9,144] those with comorbid mental and substance abuse disorders,[145,146] and finally, for the generations to come.

There is great excitement at the prospect of the global eradication of smallpox, polio, and measles, and dramatic inroads and investments have been made into reducing the impact of malaria and HIV/AIDS, as well as increasing concerns about global warming and carbon footprints. Why not place the wholly preventable deaths and disease burdens of tobacco use behavior on the same priority list of scourges to be eradicated? We can plausibly imagine a world where our families and generations to come will all grow up free of the known preventable harms of using tobacco products, especially the lethal and addictive combustibles like cigarettes, cigars, and hookah.

Perhaps the world of the 75th Surgeon General's report in 2039 will look more like the world before 1900, when the lethal and most addictive cigarette was not widely available, was not mass produced, was not inexpensive, and was not mass marketed to create a consumer demand for an

addictive product that kills nearly half its users prematurely. A "back to the future" world where, as we saw prior to 1920, lung cancer was an extremely rare disease, and the other known major diseases caused by cigarette use, including cardiovascular and pulmonary diseases, were also much less prevalent because individuals were not as likely to inhale lethal smoke. A healthier world is possible, for the sake of our children and our children's children.

Acknowledgments

This work was supported by Legacy and the National Cancer Institute (P30CA051008). Parts of this chapter include material adapted from the authors' contributions to the 2014 publication, The Health Consequences of Smoking—50 Years of Progress: A Report of the Surgeon General, and commentaries and publications on e-cigarettes. The opinions expressed herein are those of the authors and may not necessarily reflect the position of the Agency for Healthcare Research and Quality, the National Institutes of Health, or the U.S. Department of Health and Human Services.

Authors' Affiliations

David B. Abrams, PhD, is Executive Director, the Schroeder Institute for Tobacco Research and Policy Studies, Legacy; Professor, Department of Health, Behavior and Society, The Johns Hopkins Bloomberg School of Public Health; and Professor of Oncology (adjunct), Georgetown University Medical Center, Lombardi Comprehensive Cancer Center. Allison Glasser, MPH, is Project Manager, the Schroeder Institute for Tobacco Research and Policy Studies, Legacy. Andrea C. Villanti, PhD, MPH, is Director for Regulatory Science and Policy at the Schroeder Institute for Tobacco Research and Policy Studies, Legacy. Raymond Niaura, PhD, is Director of Science and Director of Training, Schroeder Institute for Tobacco Research and Policy Studies, Legacy; and Professor (adjunct), Department of Health Behavior and Society, The Johns Hopkins Bloomberg School of Public Health.

Address correspondence to: David B. Abrams, Schroeder Institute for Tobacco Research and Policy Studies, Legacy, 1724 Massachusetts Avenue, NW, Washington, DC 20036;
email DAbrams@legacyforhealth.org

References

1. Smoking and health. Report of the Advisory Committee to the Surgeon General of the Public Health Service. Washington, DC: U.S. Department of Health, Education, and Welfare; 1964.
2. The health consequences of smoking—50 years of progress. A Report of the Surgeon General. Atlanta, GA: Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; 2014.
3. Carter BD, Abnet CC, Feskanich D, et al. Smoking and mortality--beyond established causes. N Engl J Med 2015;372:631-40.
4. Doll R, Peto R, Boreham J, et al. Mortality in relation to smoking: 50 years' observations on male British doctors. Br Med J 2004;328:1519.
5. Jha P, Ramasundarahettige C, Landsman V, et al. 21st-century hazards of smoking and benefits of cessation in the United States. N Engl J Med 2013;368:341-50.
6. Levy DT, Mabry PL, Graham AL, et al. Exploring scenarios to dramatically reduce smoking prevalence: a simulation model of the three-part cessation process. Am J Public Health 2010;100:1253-9.
7. Mendez D, Alshanqeety O, Warner KE. The potential impact of smoking control policies on future global smoking trends. Tob Control 2013;22:46-51.

8. Warner KE, Mendez D. Tobacco control policy in developed countries: yesterday, today, and tomorrow. Nicotine Tob Res 2010;12:876-87.
9. Fagan P, Moolchan ET, Lawrence D, et al. Identifying health disparities across the tobacco continuum. Addiction 2007;102(Suppl 2):5-29.
10. Schroeder SA, Warner KE. Don't forget tobacco. N Engl J Med 2010;363:201-4.
11. Abrams DB. Applying transdisciplinary research strategies to understanding and eliminating health disparities. Health Educ Behav 2006;33:515-31.
12. Proctor RN. Golden holocaust: origins of the cigarette catastrophe and the case for abolition. Oakland, CA: University of California Press; 2011.
13. Brandt AM. The cigarette century: the rise, fall, and deadly persistence of the product that defined America. New York: Basic Books; 2007.
14. Proctor RN. Why ban the sale of cigarettes? The case for abolition. Tob Control 2013;22(Suppl 1):i27-30.
15. Mabry PL, Olster DH, Morgan GD, et al. Interdisciplinarity and systems science to improve population health: a view from the NIH Office of Behavioral and Social Sciences Research. Am J Prev Med 2008;35:S211-24.
16. Abrams DB, Leslie F, Mermelstein R, et al. Transdisciplinary tobacco use research. Nicotine Tob Res 2003;5(Suppl 1):S5-10.
17. Rose G. Rose's strategy of preventive medicine. New York: Oxford University Press; 1992.
18. Leischow SJ, Milstein B. Systems thinking and modeling for public health practice. Am J Public Health 2006;96:403-5.
19. Sterman JD. Learning from evidence in a complex world. Am J Public Health 2006;96:505-14.
20. Abrams DB. Comprehensive smoking cessation policy for all smokers: systems integration to save lives and money. In Bonnie RJ, Stratton K, Wallace RB (Eds), Ending the tobacco problem: a blueprint for the nation. Washington, DC: National Academies Press; 2007.
21. Abrams DB. Nicotine addiction: paradigms for research in the 21st century. Nicotine Tob Res 1999;1(Suppl 2):S211-5.
22. Colby SM, Clark MA, Rogers ML, et al. Development and reliability of the lifetime interview on smoking trajectories. Nicotine Tob Res 2012;14:290-8.
23. Gilman SE, Rende R, Boergers J, et al. Parental smoking and adolescent smoking initiation: an intergenerational perspective on tobacco control. Pediatrics 2009;123:e274-81.
24. Vanyukov MM, Tarter RE, Kirillova GP, et al. Common liability to addiction and "gateway hypothesis": theoretical, empirical and evolutionary perspective. Drug Alcohol Depend 2012;123(Suppl 1):S3-17.
25. Chassin L, Presson CC, Sherman SJ, et al. Predicting the onset of cigarette smoking in adolescents: a longitudinal study 1. J Appl Soc Psychol 1984;14:224-43.
26. Chassin L, Presson C, Seo DC, et al. Multiple trajectories of cigarette smoking and the intergenerational transmission of smoking: a multigenerational, longitudinal study of a Midwestern community sample. Health Psychol 2008;27:819-28.
27. Chassin L, Presson CC, Pitts SC, et al. The natural history of cigarette smoking from adolescence to adulthood in a midwestern community sample: multiple trajectories and their psychosocial correlates. Health Psychol 2000;19:223-31.
28. U.S. v. Philip Morris USA, Inc., et al.: U.S. Dist. Ct., D.C.; Final Opinion, August 17, 2006.
29. Orleans CT, Mabry PL, Abrams DB. Increasing tobacco cessation in America: a consumer demand perspective. Am J Prev Med 2010;38:S303-6.
30. Abrams DB. Transdisciplinary concepts and measures of craving: commentary and future directions. Addiction 2000;95(Suppl 2):S237-46.
31. Cobb NK, Graham AL, Byron MJ, et al. Online social networks and smoking cessation: a scientific research agenda. J Med Internet Res 2011;13:e119.
32. Bonsack JA, inventor patent U.S. Patent No. 238,640; 1881.
33. Thun MJ, Jamal A. How much of the decrease in cancer death rates in the United States is attributable to reductions in tobacco smoking? Tob Control 2006;15:345-7.
34. Lee PN. Epidemiological evidence relating snus to health--an updated review based on recent publications. Harm Reduct J 2013;10:36.
35. Harris JE, Thun MJ, Mondul AM, et al. Cigarette tar yields in relation to mortality from lung cancer in the cancer prevention study II prospective cohort, 1982-8. Br Med J 2004;328:72.
36. Kozlowski LT, Edwards BQ. "Not safe" is not enough: smokers have a right to know more than there is no safe tobacco product. Tob Control 2005;14(Suppl 2):ii3-7.
37. Abrams DB. Promise and peril of e-cigarettes: can disruptive technology make cigarettes obsolete? JAMA 2014;311:135-6.
38. Cobb NK, Abrams DB. The FDA, e-cigarettes, and the demise of combusted tobacco. N Engl J Med 2014;371:1469-71.
39. Henley SJ, Thun MJ, Connell C, et al. Two large prospective studies of mortality among men who use snuff or chewing tobacco (United States). Cancer Causes Control 2005;16:347-58.
40. Torches of Freedom: How the world's first PR campaign came to be. your story, 2014. Available at http://yourstory.com/2014/08/torches-of-freedom/. Accessed April 23, 2015.
41. Torches of Freedom: Women and smoking propaganda. The Society Pages; February 27, 2012. Available at http://thesocietypages.org/socimages/2012/02/27/torches-of-freedom-women-and-smoking-propaganda/. Accessed April 23, 3015.

42. Warner KE. Selling smoke: cigarette advertising and public health. Washington, DC: American Public Health Association; 1986.
43. Donohue JM, Cevasco M, Rosenthal MB. A decade of direct-to-consumer advertising of prescription drugs. N Engl J Med 2007;357:673-81.
44. Chapman S. Cigarette advertising and smoking: a review of the evidence. In Smoking out the barons: the campaign against the tobacco industry. Chichester, England: John Wiley and Sons; 1986.
45. Niaura R. Nicotine. In: McCrady BS, Epstein EE (Eds), Addictions: a comprehensive guidebook (p. 269-84). New York, NY: Oxford University Press; 2013.
46. Russell MA. Low-tar medium-nicotine cigarettes: a new approach to safer smoking. Br Med J 1976;1:1430-3.
47. How tobacco smoke causes disease: the biology and behavioral basis for smoking-attributable disease. A report of the Surgeon General. Atlanta, GA: Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; 2010.
48. Firestone Tire Recall. 2001. Washington, DC: National Highway Traffic Safety Administration; 2001. Available at http://www nhtsa.gov/PR/FirestoneRecall. Accessed April 23, 2015.
49. Consumer Advisory: Owners of recalled GM vehicles with faulty ignition switches urged to bring cars in for free remedy. Washington, DC: National Highway Traffic Safety Administration; 2014. Available at http://www nhtsa.gov/About+NHTSA/Press+Releases/2014/owners-of-recalled-gm-cars-with-faulty-switches-urged-to-seek-free-remedy. Accessed April 23, 2015.
50. Timeline: A history of GM's ignition switch defect. NPR, 2014. Available at http://www npr.org/2014/03/31/297158876/timeline-a-history-of-gms-ignition-switch-defect. Accessed April 23, 2015.
51. Zeller M, Hatsukami D. The strategic dialogue on tobacco harm reduction: a vision and blueprint for action in the US. Tob Control 2009;18:324-32.
52. Sweanor D. Disruptive technology: a blessing and a curse. Rulebreaker Research 2014:54-7.
53. Trtchounian A, Williams M, Talbot P. Conventional and electronic cigarettes (e-cigarettes) have different smoking characteristics. Nicotine Tob Res 2010;12:905-12.
54. Williams M, Talbot P. Variability among electronic cigarettes in the pressure drop, airflow rate, and aerosol production. Nicotine Tob Res 2011;13:1276-83.
55. Dawkins L, Turner J, Roberts A, et al. 'Vaping' profiles and preferences: an online survey of electronic cigarette users. Addiction 2013;108:1115-25.
56. Etter JF, Bullen C. Electronic cigarette: users profile, utilization, satisfaction and perceived efficacy. Addiction 2011;106:2017-28.
57. Foulds J, Veldheer S, Berg A. Electronic cigarettes (e-cigs): views of aficionados and clinical/public health perspectives. Int J Clin Pract 2011;65:1037-42.
58. McQueen A, Tower S, Sumner W. Interviews with "vapers": implications for future research with electronic cigarettes. Nicotine Tob Res 2011;13:860-7.
59. Vansickel AR, Eissenberg T. Electronic cigarettes: effective nicotine delivery after acute administration. Nicotine Tob Res 2013;15:267-70.
60. Cobb NK, Byron MJ, Abrams DB, et al. Novel nicotine delivery systems and public health: the rise of the "e-cigarette." Am J Public Health 2010;100:2340-2.
61. Goniewicz ML, Kuma T, Gawron M, et al. Nicotine levels in electronic cigarettes. Nicotine Tob Res 2013;15:158-66.
62. Czogala J, Goniewicz ML, Fidelus B, et al. Secondhand exposure to vapors from electronic cigarettes. Nicotine Tob Res 2014;16:655-62.
63. Goniewicz ML, Knysak J, Gawron M, et al. Levels of selected carcinogens and toxicants in vapour from electronic cigarettes. Tob Control 2014;23:133-9.
64. Hutzler C, Paschke M, Kruschinski S, et al. Chemical hazards present in liquids and vapors of electronic cigarettes. Arch Toxicol 2014;88:1295-308.
65. Kim HJ, Shin HS. Determination of tobacco-specific nitrosamines in replacement liquids of electronic cigarettes by liquid chromatography-tandem mass spectrometry. J Chromatogr A 2013;1291:48-55.
66. McAuley TR, Hopke PK, Zhao J, et al. Comparison of the effects of e-cigarette vapor and cigarette smoke on indoor air quality. Inhal Toxicol 2012;24:850-7.
67. Pellegrino RM, Tinghino B, Mangiaracina G, et al. Electronic cigarettes: an evaluation of exposure to chemicals and fine particulate matter (PM). Annali di igiene : medicina preventiva e di comunita 2012;24:279-88.
68. Schripp T, Markewitz D, Uhde E, et al. Does e-cigarette consumption cause passive vaping? Indoor Air 2013;23:25-31.
69. Williams M, Villarreal A, Bozhilov K, et al. Metal and silicate particles including nanoparticles are present in electronic cigarette cartomizer fluid and aerosol. PLoS One 2013;8:e57987.
70. Kosmider L, Sobczak A, Fik M, et al. Carbonyl compounds in electronic cigarette vapors: effects of nicotine solvent and battery output voltage. Nicotine Tob Res 2014;16:1319-26.
71. Hecht SS, Carmella SG, Kotandeniya D, et al. Evaluation of toxicant and carcinogen metabolites in the urine of e-cigarette users versus cigarette smokers. Nicotine Tob Res 2014:e-pub.

72. Goniewicz ML, Lingas EO, Hajek P. Patterns of electronic cigarette use and user beliefs about their safety and benefits: an internet survey. Drug Alcohol Rev 2013;32:133-40.
73. Zhu SH, Gamst A, Lee M, et al. The use and perception of electronic cigarettes and snus among the U.S. population. PLoS One 2013;8:e79332.
74. Adkison SE, O'Connor RJ, Bansal-Travers M, et al. Electronic nicotine delivery systems: international tobacco control four-country survey. Am J Prev Med 2013;44:207-15.
75. Farsalinos KE, Romagna G, Tsiapras D, et al. Characteristics, perceived side effects and benefits of electronic cigarette use: a worldwide survey of more than 19,000 consumers. Int J Environ Res Public Health 2014;11:4356-73.
76. Kralikova E, Novak J, West O, et al. Do e-cigarettes have the potential to compete with conventional cigarettes? A survey of conventional cigarette smokers' experiences with e-cigarettes. Chest 2013;144:1609-14.
77. Richardson A, Pearson J, Xiao H, et al. Prevalence, harm perceptions, and reasons for using noncombustible tobacco products among current and former smokers. Am J Public Health 2014;104:1437-44.
78. Arrazola RA, Neff LJ, Kennedy SM, et al. Tobacco use among middle and high school students--United States, 2013. MMWR 2014;63:1021-6.
79. Tobacco product use among middle and high school students--United States, 2011 and 2012. MMWR 2013;62:893-7.
80. Corey CG, Wang B, Johnson SE, et al. Notes from the field: electronic cigarette use among middle and high school students - United States, 2011-2012. MMWR 2013;62:729-30.
81. King BA, Patel R, Nguyen KH, et al. Trends in awareness and use of electronic cigarettes among US adults, 2010-2013. Nicotine Tob Res 2015;17:219-27.
82. King BA, Alam S, Promoff G, et al. Awareness and ever-use of electronic cigarettes among U.S. adults, 2010-2011. Nicotine Tob Res 2013;15:1623-7.
83. Agaku IT, King BA, Husten CG, et al. Tobacco product use among adults--United States, 2012-2013. MMWR 2014;63:542-7.
84. Table 29. Use of cigarettes by e-cigarettes in the last 30 days. Ann Arbor, MI: University of Michigan, Monitoring the Future; 2014. Available at http://www.monitoringthefuture.org/data/14data/14tobtbl29.pdf. Accessed April 23, 2015.
85. Dutra LM, Glantz SA. Electronic cigarettes and conventional cigarette use among U.S. adolescents: a cross-sectional study. JAMA Pediatr 2014;168(7):610-7.
86. Abrams DB. Potential and pitfalls of e-cigarettes--reply. JAMA 2014;311:1922-3.
87. Niaura RS, Glynn TJ, Abrams DB. Youth experimentation with e-cigarettes: another interpretation of the data. JAMA 2014;312:641-2.
88. Richardson A, Williams V, Rath J, et al. The next generation of users: prevalence and longitudinal patterns of tobacco use among U.S. young adults. Am J Public Health 2014;104:1429-36.
89. Table 3. Trends in 30-day prevalence of use of various drugs in grades 8, 10, and 12. Ann Arbor, MI: University of Michigan, Michigan News; 2015. Available at http://www monitoringthefuture.org//pressreleases/14drugpr_complete.pdf. Accessed April 23, 2015.
90. Teen smoking continues to decline in 2013. Ann Arbor, MI: University of Michigan, Institute for Social Research; 2013. Available at http://home.isr.umich.edu/releases/teen-smoking-continues-to-decline-in-2013/. Accessed April 23, 2015.
91. Walton KM, Abrams DB, Bailey WC, et al. NIH electronic cigarette workshop: developing a research agenda. Nicotine Tob Res 2015;17:259-69.
92. Bullen C, Howe C, Laugesen M, et al. Electronic cigarettes for smoking cessation: a randomised controlled trial. Lancet 2013;382:1629-37.
93. Caponnetto P, Campagna D, Cibella F, et al. Efficiency and safety of an eLectronic cigAreTte (ECLAT) as tobacco cigarettes substitute: a prospective 12-month randomized control design study. PLoS One 2013;8:e66317.
94. Vickerman KA, Carpenter KM, Altman T, et al. Use of electronic cigarettes among state tobacco cessation quitline callers. Nicotine Tob Res 2013;15:1787-91.
95. Grana RA, Popova L, Ling PM. A longitudinal analysis of electronic cigarette use and smoking cessation. JAMA Intern Med 2014;174(5):812-3.
96. Pearson JL, Stanton CA, Cha S, et al. E-cigarettes and smoking cessation: insights and cautions from a secondary analysis of data from a study of online treatment-seeking smokers. Nicotine Tob Res 2014:e-pub.
97. Biener L, Hargraves JL. A longitudinal study of electronic cigarette use among a population-based sample of adult smokers: association with smoking cessation and motivation to quit. Nicotine Tob Res 2015;17:127-33.
98. Brown J, Beard E, Kotz D, et al. Real-world effectiveness of e-cigarettes when used to aid smoking cessation: a cross-sectional population study. Addiction 2014;109(9):1531-40.
99. Bullen C, McRobbie H, Thornley S, et al. Effect of an electronic nicotine delivery device (e cigarette) on desire to smoke and withdrawal, user preferences and nicotine delivery: randomised cross-over trial. Tob Control 2010;19:98-103.
100. Dawkins L, Corcoran O. Acute electronic cigarette use: nicotine delivery and subjective effects in regular users. Psychopharmacology (Berl) 2014;231:401-7.
101. Farsalinos KE, Spyrou A, Tsimopoulou K, et al. Nicotine absorption from electronic cigarette use:

comparison between first and new-generation devices. Sci Rep 2014;4:4133.
102. Vansickel AR, Weaver MF, Eissenberg T. Clinical laboratory assessment of the abuse liability of an electronic cigarette. Addiction 2012;107:1493-500.
103. Vansickel AR, Cobb CO, Weaver MF, et al. A clinical laboratory model for evaluating the acute effects of electronic "cigarettes": nicotine delivery profile and cardiovascular and subjective effects. Cancer Epidemiol Biomarkers Prev 2010;19:1945-
104. Talih S, Balhas Z, Eissenberg T, et al. Effects of user puff topography, device voltage, and liquid nicotine concentration on electronic cigarette nicotine yield: measurements and model predictions. Nicotine Tob Res 2015;17:150-7.
105. Spindle TR, Breland AB, Karaoghlanian NV, et al. Preliminary results of an examination of electronic cigarette user puff topography: the effect of a mouthpiece-based topography measurement device on plasma nicotine and subjective effects. Nicotine Tob Res 2015;17:142-9.
106. Dawkins L, Turner J, Crowe E. Nicotine derived from the electronic cigarette improves time-based prospective memory in abstinent smokers. Psychopharmacology (Berl) 2013;227:377-84.
107. Dawkins L, Turner J, Hasna S, et al. The electronic-cigarette: effects on desire to smoke, withdrawal symptoms and cognition. Addict Behav 2012;37:970-3.
108. Gladwell M. David and Goliath: Underdogs, misfits, and the art of battling giants. New York, NY: Little, Brown and Company; 2013.
109. Cobb CO, Villanti AC, Graham AL, et al. Markov modeling to estimate the population impact of emerging tobacco products: a proof-of-concept study. Tob Regul Sci 2015;1(2):121-41.
110. Villanti AC, Vargyas EJ, Niaura RS, et al. Food and Drug Administration regulation of tobacco: integrating science, law, policy, and advocacy. Am J Public Health 2011;101:1160-2.
111. McNeill A, Munafo MR. Reducing harm from tobacco use. J Psychopharmacol 2012:e-pub.
112. Southwell BG, Kim AE, Tessman GK, et al. The marketing of dissolvable tobacco: social science and public policy research needs. Am J Health Promot 2012;26:331-2.
113. McNeill A, Munafo MR. Reducing harm from tobacco use. J Psychopharmacol 2013;27:13-8.
114. Family Smoking Prevention and Tobacco Control Act, Public Law 111-31, U.S. Statutes at Large (2009):123.
115. Ayers JW, Ribisl KM, Brownstein JS. Tracking the rise in popularity of electronic nicotine delivery systems (electronic cigarettes) using search query surveillance. Am J Prev Med 2011;40:448-53.
116. Pearson JL, Richardson A, Niaura RS, et al. e-Cigarette awareness, use, and harm perceptions in U.S. adults. Am J Public Health 2012;102:1758-66.
117. Hersey JC, Niederdeppe J, Evans WD, et al. The effects of state counterindustry media campaigns on beliefs, attitudes, and smoking status among teens and young adults. Prev Med 2003;37:544-52.
118. About one in five U.S. adult cigarette smokers have tried an electronic cigarette (press release). Atlanta, GA: Centers for Disease Control and Prevention; 2013. Available at http://www.cdc.gov/media/releases/2013/p0228_electronic_cigarettes html. Accessed April 23, 2015.
119. Evans WD, Abroms LC, Poropatich R, et al. Mobile health evaluation methods: the Text4baby case study. J Health Commun 2012;17(Suppl 1):22-9.
120. Arnott D. There's no single endgame. Tob Control 2013;22(Suppl 1):i38-9.
121. Benowitz NL, Henningfield JE. Reducing the nicotine content to make cigarettes less addictive. Tob Control 2013;22(Suppl 1):i14-7.
122. Berrick AJ. The tobacco-free generation proposal. Tob Control 2013;22(Suppl 1):i22-6.
123. Borland R. Minimising the harm from nicotine use: finding the right regulatory framework. Tob Control 2013;22(Suppl 1):i6-9.
124. Callard CD, Collishaw NE. Supply-side options for an endgame for the tobacco industry. Tob Control 2013;22(Suppl 1):i10-3.
125. Chapman S, Wakefield MA. Large-scale unassisted smoking cessation over 50 years: lessons from history for endgame planning in tobacco control. Tob Control 2013;22(Suppl 1):i33-5.
126. Hatsukami DK. Ending tobacco-caused mortality and morbidity: the case for performance standards for tobacco products. Tob Control 2013;22(Suppl 1):i36-7.
127. Isett KR. In and across bureaucracy: structural and administrative issues for the tobacco endgame. Tob Control 2013;22(Suppl 1):i58-60.
128. Kozlowski LT. Ending versus controlling versus employing addiction in the tobacco-caused disease endgame: moral psychological perspectives. Tob Control 2013;22(Suppl 1):i31-2.
129. Malone RE. Tobacco endgames: what they are and are not, issues for tobacco control strategic planning and a possible US scenario. Tob Control 2013;22(Suppl 1):i42-4.
130. Myers ML. The FCTC's evidence-based policies remain a key to ending the tobacco epidemic. Tob Control 2013;22(Suppl 1):i45-6.

131. Rabe BG. Political impediments to a tobacco endgame. Tob Control 2013;22(Suppl 1):i52-4.
132. Reuter P. Can tobacco control endgame analysis learn anything from the U.S. experience with illegal drugs? Tob Control 2013;22(Suppl 1):i49-51.
133. Smith EA. Questions for a tobacco-free future. Tob Control 2013;22(Suppl 1):i1-2.
134. Tam J. Cultivating the next generation of tobacco endgame advocates. Tob Control 2013;22(Suppl 1):i47-8.
135. Thomas BP, Gostin LO. Tobacco endgame strategies: challenges in ethics and law. Tob Control 2013;22(Suppl 1):i55-7.
136. Warner KE. An endgame for tobacco? Tob Control 2013;22(Suppl 1):i3-5.
137. Wilson N, Thomson GW, Edwards R, et al. Potential advantages and disadvantages of an endgame strategy: a 'sinking lid' on tobacco supply. Tob Control 2013;22(Suppl 1):i18-21.
138. Zeller M. Reflections on the 'endgame' for tobacco control. Tob Control 2013;22(Suppl 1):i40-1.
139. Warner KE, Pollack HA. The nicotine fix. The Atlantic; 2014.
140. Zeller M. Three years later: an assessment of the implementation of the Family Smoking Prevention and Tobacco Control Act. Tob Control 2012;21:453-4.
141. Bonnie RJ, Stratton KR, Wallace RB (Eds). Ending the tobacco problem: a blueprint for the nation. Washington, DC: National Academies Press; 2007.
142. Ending the tobacco epidemic: a tobacco control strategic action plan for the U.S. Department of Health and Human Services. Washington, DC: U.S. Department of Health and Human Services, Office of the Assistant Secretary for Health; 2010.
143. Cheyne A, Dorfman L, Daynard RA, et al. The debate on regulating menthol cigarettes: closing a dangerous loophole vs freedom of choice. Am J Public Health 2014;104:e54-61.
144. Tobacco use among U.S. racial/ethnic minority groups—African Americans, American Indians and Alaska Natives, Asian Americans and Pacific Islanders, and Hispanics: a report of the Surgeon General. Atlanta, GA: Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; 1998.
145. Schroeder SA, Morris CD. Confronting a neglected epidemic: tobacco cessation for persons with mental illnesses and substance abuse problems. Annu Rev Public Health 2010;31:297-314 1 p following.
146. Hershey JC, Niederdeppe J, Evans WD, et al. The theory of "truth": how counterindustry campaigns affect smoking behavior among teens. Health Psychol 2005;24:22-31.

David B. Abrams, PhD, is Professor, Johns Hopkins Bloomberg School of Public Health; Professor of Oncology, Georgetown University Lombardi Cancer Center; and Executive Director, the Schroeder Institute for Tobacco Research and Policy Studies, Legacy. Previously, he was Director of the Office of Behavioral and Social Sciences Research, National Institutes of Health. He has served as President of the Society for Behavioral Medicine and was a recipient of their Distinguished Scientist and Research Mentorship awards. Dr. Abrams received the Joseph Cullen Memorial Award of the American Society for Preventive Oncology for lifetime contributions to tobacco control. He is a clinical psychologist specializing in health psychology, addictions, and tobacco use behavior.

Allison M. Glasser, MPH, is a Project Manager with the Schroeder Institute for Tobacco Research and Policy Studies at Legacy. She manages various projects related to Food and Drug Administration regulatory science, including conducting evidence reviews, preparing knowledge synthesis papers, and facilitating communication among scientists, advocates, and other stakeholders through regular meetings and strategic consensus conferences. Her previous work has involved advancing evidence-based practices and policies through promoting cancer screening, worksite wellness programs, harm reduction policy advocacy, and managing systematic evidence reviews.

Andrea C. Villanti, PhD, MPH, is Director for Regulatory Science and Policy at the Schroeder Institute for Tobacco Research and Policy Studies, Legacy, where she also serves as Principal Investigator on Legacy's Young Adult Cohort Study. Dr. Villanti holds an adjunct faculty appointment in the Department of Health, Behavior and Society at the Johns Hopkins Bloomberg School of Public Health. Her interests include translational research to improve tobacco control policy and program decisionmaking, particularly the Food and Drug Administration's regulation of tobacco. Her interests also include tracking tobacco use patterns, trajectories, and the impact of tobacco interventions in young adults.

Raymond Niaura, PhD, is Director of Science and Director of Training at the Schroeder Institute for Tobacco Research and Policy Studies, Legacy; Professor (adjunct) in the Department of Health Behavior and Society, Johns Hopkins Bloomberg School of Public Health; and Professor (adjunct) in the Department of Oncology, Lombardi Comprehensive Cancer Center, at the Georgetown University Medical Center. While at Brown University, he was Director of Transdisciplinary Research and Director of Postdoctoral Training in Behavioral Medicine in the Department of Psychiatry and Human Behavior. Dr. Niaura is the former President of the Society of Nicotine and Tobacco Research.

Physical Activity: Numerous Benefits and Effective Interventions

James F. Sallis and Jordan A. Carlson

Abstract

Physical activity has been required for obtaining food, work, and transportation throughout human history, until recently. Physical inactivity is now one of the leading causes of death, responsible for an estimated 5.3 million deaths per year worldwide, through its effects on cardiovascular diseases, diabetes, and some cancers. Physical activity's benefits for brain structure and function have emerged more recently. Prevalence rates of meeting physical activity guidelines vary dramatically, depending on whether self-report or objective measures are used. However, prevalence of meeting physical activity guidelines appears to be low among adults and youth in the United States and is decreasing worldwide. Thus, physical inactivity has been identified as a global pandemic. Numerous interventions to increase physical activity have been demonstrated to be effective in systematic reviews. Some interventions target individuals, such as educational programs delivered by in-person counseling, telephone, or computer. Other interventions target groups and organizations, such as social support groups, school physical education, and comprehensive worksite programs. Environmental interventions to improve access to and the quality of walkable neighborhoods, facilities for pedestrians and cyclists, and recreation facilities can also be effective. However, there is limited evidence that effective interventions are being widely implemented. Failure to act on the evidence of burden of disease, low prevalence, and effective interventions is costing millions of lives each year. Research priorities include implementation and dissemination of effective interventions, economic analyses of physical activity, and "natural experiments" to evaluate multi-level and policy interventions.

Introduction

Physical activity comprises a set of behaviors that appears to play a unique role in health and well-being due to a wide range of benefits. Physical activity is a strong protective factor from premature mortality, most of the leading chronic diseases, risk factors for chronic diseases, and common mental health problems, such as depression and anxiety and Alzheimer's disease.[1] Possibly because physical activity requires integrated and coordinated functioning of the whole body, it appears to benefit many biological systems. Throughout human history, physical activity has been required for obtaining food, working, getting from one place to another (transportation), and performing household chores. Dance and sports were developed in virtually every culture for pleasure and cultural expression.

Mechanization started during the Industrial Revolution and replaced many types of labor, both at work and at home. Activity is no longer required to obtain food. Automobiles have largely replaced walking for transportation in recent decades. Dance and leisure have become mostly spectator activities, while the ever-increasing options for electronic entertainment have become the dominant form of leisure activity. These long-term trends have helped produce inactive lifestyles in most of the world's populations, and the profoundly negative consequences for health have been extensively documented. Physical inactivity has become one of the biggest threats to worldwide health.[2]

Though many evidence-based interventions are available to increase physical activity, they have not been widely implemented; the prevalence of sufficient physical activity in the United States remains low, with few signs of improvement. Thus, there is a compelling need for increased attention to and investment in physical activity interventions.[3]

In this chapter, we summarize the numerous health effects of physical activity, with a special focus on less-familiar effects on brain and cognitive functioning. Prevalence, trends, and correlates of physical activity are briefly summarized. Substantial research has produced a diverse array of evidence-based interventions. Rationales for research, policy, and improved practice for physical activity are presented.

Breadth and Strength of Health Impacts

A recent analysis revealed that physical inactivity is responsible for over 5 million deaths annually worldwide, which is similar to the death toll of tobacco smoking.[4] The World Health Organization estimates that physical inactivity is the fourth leading cause of death globally; lower than hypertension and tobacco smoking, similar to obesity, and higher than dietary patterns and hyperglycemia. Most of the inactivity-related deaths are in low- and middle-income countries, so this is not just a concern in high-income countries.[2] Some consider physical inactivity to be a global pandemic.[3] Physical inactivity also is the fourth underlying cause of death in the United States, with an estimated 200,000 deaths per year. This is about half the deaths attributable to tobacco smoking but twice the deaths attributable to alcohol use and low intake of fruits and vegetables.[5]

Physical activity affects health and disease through many pathways and systems. The Report of the Physical Activity Guidelines Advisory Committee consisted of almost 700 pages of systematic literature reviews.[1] They found strong evidence that physical activity reduces risk of premature mortality, coronary heart disease, high blood pressure, stroke, metabolic syndrome, type 2 diabetes, breast cancer, colon cancer, depression, and falling, and it is associated with improved body composition, bone health, functional health, and cognitive health. It is likely that physical activity provides a broader range of documented health benefits than any other factor (e.g., behavior, medication, or medical procedure).

The strength of the effects of physical activity on leading chronic diseases is notable. In a recent review, I-Min Lee and colleagues[4] conservatively calculated that physical inactivity accounts for 5.8

percent of deaths from coronary heart disease, 7.2 percent from type 2 diabetes, 10.1 percent from breast cancer, and 10.4 percent from colon cancer. These calculations, while impressive, are almost certainly underestimated because they were adjusted for risk factors that are affected by physical activity such as obesity, lipids, and glucose. A U.S. study indicated that inactive adults would gain 1.3 to 3.7 years of life expectancy by becoming active at age 50.[6] This result compares favorably to 2.3 to 2.5 years gained among smokers who quit at age 50 and 0.5 to 0.7 years gained by all obese people becoming normal weight at age 50.[7] A recent study using national U.S. data estimated that 9-11 percent of aggregate U.S. health care expenditures were associated with physical inactivity.[8]

Brain and Cognitive Health

This section summarizes substantial recent evidence showing that the cognitive decline of aging can be ameliorated through physical activity. A review of three meta-analyses on physical activity interventions in older adults concluded that the beneficial effects on cognitive functioning were medium to large for executive functioning (e.g., working memory, reasoning, problem solving; d (effect size) = 0.58), with somewhat smaller effects for spatial (d = .36), controlled (d = .33), and speed tasks (reaction time; d = .19; see Figure 1).[9] Furthermore, a meta-analysis of 12 randomized trials in older adults with dementia found that physical activity ameliorated cognitive impairments (d = .57).[10]

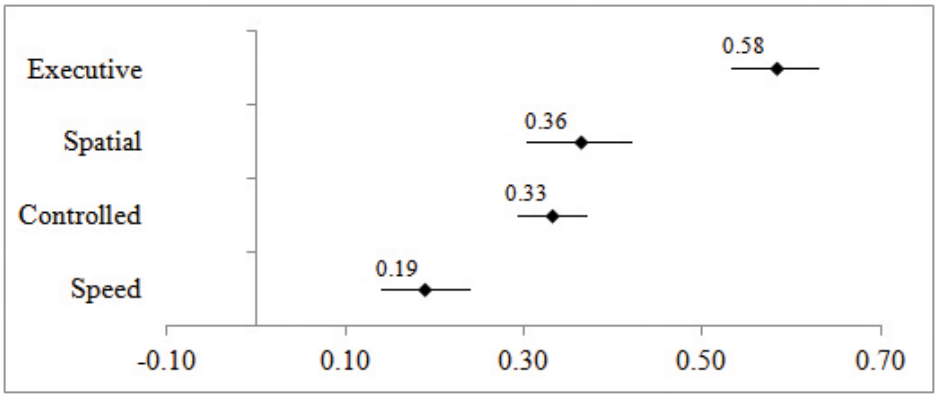


Figure 1. Effect sizes and 95% confidence intervals for effects of physical activity on cognitive functioning in older adults
Source: Kramer, Erickson, 2007.[9] Used with permission.

Physical activity also has benefits for children's cognitive functioning.[11] A meta-analysis of 44 studies in children found small to medium effect sizes for physical activity on various cognitive measures, with the largest effect for perceptual skills (d = .49; see Figure 2).[12] A single 20-minute bout of physical activity resulted in improved response accuracy, better performance on an academic achievement test, and more brain activity during cognitive tasks.[13]

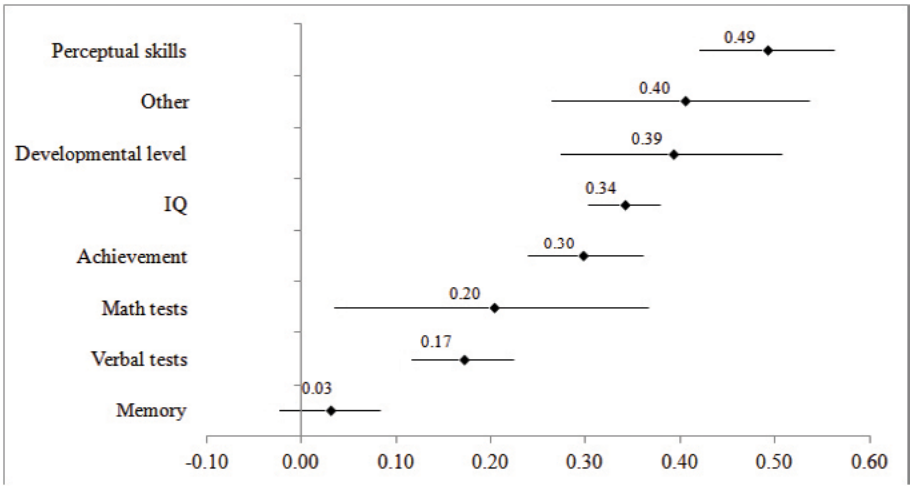


Figure 2. Effect sizes and 95% confidence intervals for effects of physical activity on cognitive functioning in youth

Source: Sibley, Etnier, 2003.[12] Used with permission.

While evidence supports the use of brain training games to improve cognitive functioning, the cognitive benefits from brain training typically do not transfer to cognitive tasks other than the task on which the training was based.[14] By contrast, the cognitive benefits of physical activity are both large and generalized, meaning the benefits are observed across multiple systems in the brain.[15,16]

Several mechanisms link physical activity to improved brain health and cognition. Physical activity stimulates brain-derived neurotropic factor (BDNF) and insulin-like growth factor 1 (IGF-1).[11] These proteins released in the brain trigger angiogenesis, the growth of new blood vessels in the brain, and neurogenesis, the growth of new neurons in the brain, which contribute to learning and memory.[17]

Prevalence and Trends of Meeting Physical Activity Guidelines

The overwhelming evidence about the numerous health benefits of physical activity was the basis for the first official physical activity recommendations in the United States[1] and from the World Health Organization.[18] The basic recommendation for adults is to accumulate at least 150 minutes of moderate intensity physical activity weekly, 75 minutes of vigorous intensity activity weekly, or some combination of the two. The main recommendation for youth is to accumulate at least 60 minutes per day of moderate to vigorous physical activity.

The prevalence of physical activity varies dramatically by measurement method, with higher prevalence estimates based on self-reports and lower estimates based on objective measures. However, by all methods the prevalence of meeting guidelines is low. Physical activity among adults in the United States is generally higher among younger people, males, people with higher education, and non-Hispanic whites. The groups with lower physical activity are generally at higher risk of

chronic diseases. Based on self-reports of leisure time physical activity, prevalence rates of meeting guidelines in 2005 ranged from 62 percent in men aged 18-24 to 44.5 percent in men over 65 and from 52.7 percent in women aged 18-24 to 36.3 percent in women over 65.[1] Comparisons across 122 countries using the International Physical Activity Questionnaire[a] yielded an estimate that 31 percent of adults worldwide were insufficiently active. The range was 4.7 percent in Bangladesh to 71.9 percent in Malta, with the United States having an inactivity prevalence of 33.5 percent for men and 47.4 percent for women.[19] Active transportation to work was compared across 16 high-income countries. The United States had one of the lowest rates of walking to work (3-4 percent); the rate of walking to work in China, Germany, and Sweden was 20 percent. The United States also had one of the lowest rates of cycling to work (0.5 - 3 percent), which was much lower than in China, Denmark, and the Netherlands (>20 percent).

Prevalence rates for adolescents were compared for 105 countries. About 80 percent of 13-15 year-olds did not meet the youth physical activity guideline of 60 minutes every day, based on self-report.[19]

Four countries (Norway, Portugal, Sweden, United States) have conducted national prevalence studies of adults using objective accelerometers that are typically worn for 1 week and provide minute-by-minute activity levels. Variation across countries was modest for men, with the United States having the lowest average of moderate-to-vigorous physical activity (33 minutes per day) and Portugal having the highest (37 minutes per day). There was more variation for women, ranging from 19 minutes per day in the United States to 45 minutes per day for Portugal.[19] Using a rigorous method of scoring the accelerometer data, prevalence of meeting guidelines among U.S. adults was less than 5 percent.[20] Samples of adolescents from 10 countries (N > 30,000) had accelerometer data. Minutes of physical activity per day ranged from 45.9 in the United States to 83.6 in Norway.[19]

Though physical activity prevalence estimates vary, it is most likely that at least a majority of U.S. adults and youth are exposed to the multiple, strong, and negative health consequences of physical inactivity. On all measures except self-reported physical activity of adults, people in the United States are low, if not the lowest, in international comparisons. Research to identify the modifiable factors that account for these international differences could point the way to promising interventions.

Physical Activity Trends

Though trends for tobacco consumption are available for about 100 years, and trends that allow estimates of food intake are available for many decades, national physical activity surveillance for adults was only initiated in the mid-1980s in the United States. The National Health Interview Survey (NHIS) assessed leisure time physical activity in a national sample of adults, and results from 1997-2006 show no change over time (see Figure 3).[1,21] Short-term trends for adolescent physical activity from the Youth Risk Behavior Surveillance System (YRBSS)[b] similarly showed no change from 1999 to 2006.[1]

[a] International Physical Activity Questionnaire, available at http://www.sdp.univ fvg.it/sites/default/files/IPAQ_English_self-admin_long.pdf.

[b] The Youth Risk Behavior Surveillance System is managed by the Centers for Disease Control and Prevention (CDC). See http://www.cdc.gov/HealthyYouth/yrbs/index htm for more information.

Long-term trends in active transportation can be derived from periodic travel studies conducted by the U.S. Department of Transportation. Over 40 percent of students walked or cycled to school in 1969, but by 2001 the prevalence had declined to about 14 percent.[22] Brownson and colleagues[23] found long-term evidence supporting conclusions that physical activity at home, at work, and for transportation have declined substantially among Americans in the past 50 years, though leisure time physical activity has remained relatively stable.

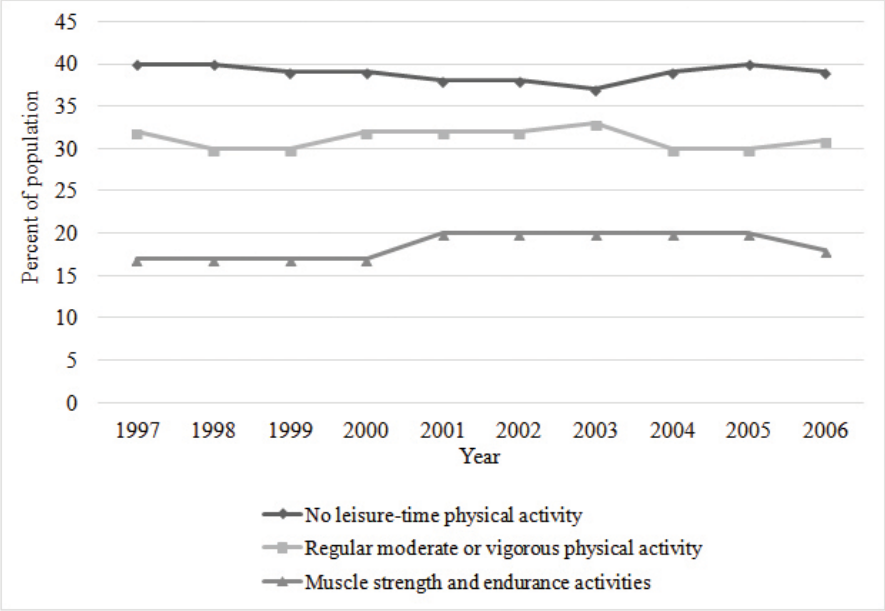


Figure 3. Reported physical activity by adults in the United States

Source: Physical Activity Guidelines Advisory Committee[1]; Healthy People 2020[21] Used with permission.

Etiology of Physical Activity: From Genes to Built Environments

Hundreds of studies have examined correlates (cross-sectional) and determinants (prospective) of physical activity. Such studies identify possible mechanisms of change that can be targeted in physical activity interventions, but the number and diversity of potential etiological variables demonstrate the challenges faced by those attempting to improve public health through increasing physical activity.

An analysis of reviews identified a wide range of variables consistently related to physical activity, including factors at biological, psychological, social, organizational, environmental, and policy levels of analysis.[24] Biology is working against efforts to increase physical activity. The age decline in physical activity documented in humans, rats, fish, and insects has been attributed to loss of dopamine receptors in the nucleus accumbens. This area of the brain links movement and reward centers, so as they age, people become less able to derive pleasure from movement.[25] It is clear that genes affect physical activity, based on heritability estimates. Genes could be involved in variations in experiencing reward and pain from physical activity, but few specific genes have been linked to physical activity.[24]

Understanding psychological and social correlates and determinants of physical activity can provide guidance to the design of interventions. Psychological correlates/determinants include intention to be active and self-efficacy, or confidence in one's ability to be active in specific situations. Stress may be a barrier to physical activity. Surprisingly, social support and social norms were not consistently identified as correlates/determinants by the reviews.[24]

Built environment correlates/determinants of physical activity is a relatively recent topic of research on modifiable factors that when properly designed could have long-term effects on large populations. Environmental factors are expected to have effects that are specific to the domain of physical activity, with walking being the most studied physical activity outcome.[26] For example, walking for transportation is expected to be higher among those living in neighborhoods with shops and services nearby (i.e., mixed use) and connected streets. The literature generally supports this hypothesis for youth but less so for adults.[24] Walking for leisure is expected to be related to the aesthetics of the area, presence of sidewalks, and design of street crossings. Leisure activity more generally is expected to be related to proximity of recreational facilities, such as parks. These hypotheses have generally been supported in the literature for youth and adults.[24] Total physical activity was consistently related to at least one variable in five categories of neighborhood environments, especially recreation facilities, transportation environments, and aesthetics. These results illustrate the important role that city planners, transportation engineers, and parks and recreation sectors need to play in creating activity-supportive environments.

Multiple Effective Interventions

The science of physical activity interventions is well advanced, based on more than 30 years of research, much of it supported by the National Institutes of Health (NIH). Many evaluated interventions are based on theories of behavior change, randomized controlled trials are common, and the literature is frequently reviewed.[27] This section highlights systematic reviews of interventions mainly targeted to non-clinical populations.

Interventions can be designed to change one or more of the levels of influence outlined in ecological models of behavior, including individuals, groups, organizations, communities, environments, and policies.[28] Most evaluated interventions target individuals and groups, and many randomized controlled trials have been reported. Except for interventions that only provide group exercise classes, most behavior change programs strive to enhance motivation and teach skills that allow individuals to modify their own behaviors. Modes of communication include individual counseling, group classes, printed materials, telephone counseling, and more recently, Web sites, text messaging, and mobile apps. There is a substantial literature targeting organizations and communities as a means of reaching large numbers of people with behavior change programs and creating supportive social environments. Studies of organizational interventions conducted in schools, worksites, and faith-based institutions are often randomized, but studies targeting whole communities are almost always quasi-experimental.

Recently, environmental and policy interventions and multi-level interventions have become more common, informed by ecological models of behavior.[29] Environmental interventions have ranged from signs promoting the choice of stairs rather than nearby escalators to construction or renovation of sidewalks, parks, and trails. These studies often are "natural experiments" or opportunistic

evaluations of interventions not controlled by investigators but implemented by local governments. Evaluations of environmental and policy changes on a larger scale—such as multi-year multi-component bicycle promotion, implementation of new zoning laws, or construction of public transit facilities—use uncontrolled pre-post evaluations, post-test-only comparisons with control areas, or historical data. Though multi-level interventions that include environmental and policy changes are expected to produce broad and long-lasting impacts, less rigorous study designs must be used than with interventions targeting individuals.

In the United States, the most definitive reviews for population health purposes are from the Web-based Guide to Community Preventive Services, sponsored by the Centers for Disease Control and Prevention (CDC).[c] The reviews of physical activity interventions were published in the early 2000s.[30] Types of interventions included information, behavioral and social, organizational, and environmental and policy; the results of the review are shown in Table 1. Several intervention types were identified as effective, while others had insufficient evidence. A subsequent analysis found good cost-effectiveness for virtually all of the interventions.[31] The most cost-effective approach was signage to encourage stair use, though these interventions had small effects. Individually-adapted behavior change interventions had the lowest cost-effectiveness, but they produced the strongest effects (35-43 percent of recommended daily physical activity).

Table 1. Physical activity intervention recommendations from the Community Guide to Preventive Services

Strategy	Level of Evidence
Point-of-decision prompts	Effective
Community-wide campaigns	Effective
Mass media campaigns	Insufficient
School-based PE	Effective
Health education	Insufficient
Family support	Insufficient
Social support	Insufficient
Individually-adapted program	Effective
Places for PA + outreach	Effective

Source: Kahn, Ramsay, Brownson, et al., 2002.[30] Used with permission.
Note: PA = physical activity; PE = physical education.

A further "Community Guide" review of built environment interventions was based mainly on cross-sectional studies (i.e., post-intervention comparisons).[32] Community-scale design and land use studies often compare walkable communities defined by mixed use (destinations within walking distance of homes), high residential density, grid-like pattern of connected streets, and sidewalks to suburban-style, automobile-oriented communities with separation of land uses, low residential density, and poorly connected streets. Residents of walkable communities had a median of 161 percent more physical activity, providing strong evidence of effectiveness. Street-scale urban design

[c] For more information, see http://www.thecommunityguide.org/

and land use interventions consisted of improved lighting, landscaping, traffic calming, improved sidewalks, and safer street crossings. These approaches were also found to be effective.[32]

Heath and colleagues[33] conducted a review of systematic reviews of physical activity interventions targeting individuals and groups; the results are shown in Figure 4. Though effect sizes varied substantially, interventions of all types and with all specific populations were significant, with two exceptions. The lowest significant effect sizes were with computer-tailored programs and counseling in health care. Among the strongest effect sizes were those for Web-based programs, after school programs, and programs including use of pedometers.

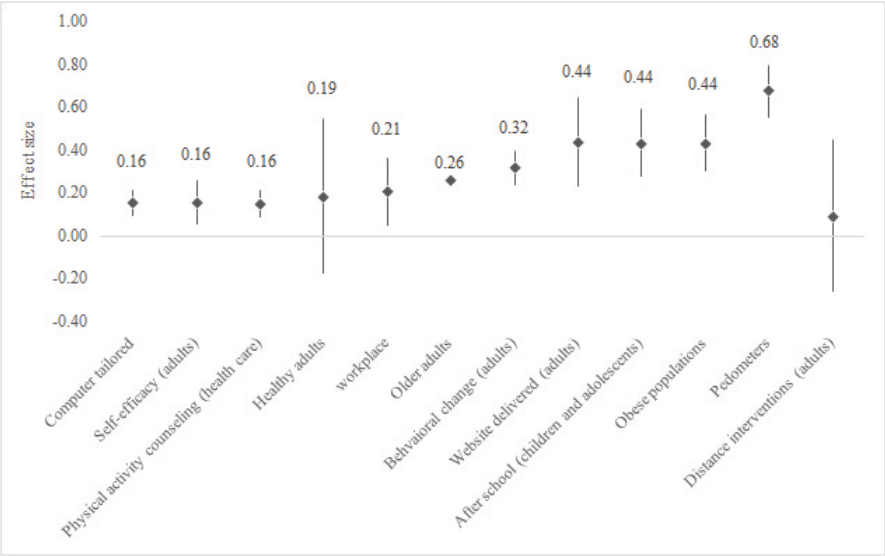


Figure 4. Meta-analysis of physical activity intervention effect sizes with 95% confidence intervals

Source: Heath, Parra-Perez, Sarmiento, et al., 2006.[33] Used with permission.

Mozaffarian and colleagues[34] conducted a systematic review of population-based physical activity interventions. Numerous intervention types were judged to have sufficient evidence to justify widespread implementation. Some findings confirmed results of "Community Guide" reviews: signs encouraging stair use, enhanced school physical education, and community-scale land use and walkability. Other recommended interventions reflected more recent evidence for schools (comprehensive multi-behavior programs, improved playgrounds and equipment, classroom activity breaks), worksites (comprehensive multi-behavior programs, scheduled time for physical activity at work, fitness centers at work), and communities (improved access to parks, improved safety of pedestrian facilities to support walking to school, improved traffic safety, improved aesthetics).

A report from the President's Council on Fitness, Sports and Nutrition[35] on strategies to increase physical activity in young people found that most of the interventions with strong evidence of effectiveness were in schools, such as using multiple strategies to increase physical activity throughout the school day and providing highly active physical education. Active commuting to school and classroom activity breaks had less support. Various interventions in preschool and

community built environment strategies had suggestive evidence. After school programs, family-based programs, and interventions in primary care had insufficient evidence (see Table 2).

Table 2. Conclusions about effectiveness of youth physical activity interventions

Strategy	Level of Evidence
Multi-component school programs	Sufficient
Physical education	Sufficient
Active transportation	Suggestive
Activity breaks in classroom	Suggestive
School physical environment	Insufficient
After school	Insufficient
Preschool and childcare settings	Suggestive
Built environment	Suggestive
Home & family	Insufficient
Primary care	Insufficient

Source: President's Council on Fitness, Sports, and Nutrition, 2013.[35] Used with permission.

Although there is some disagreement across reviews regarding support for specific intervention types, all reviews identified multiple interventions with strong evidence of increasing physical activity. Strategies targeting individuals, organizations, and community built environments were found to be effective. However, not all interventions were supported by the evidence. Thus, interventions need to be carefully selected for further implementation, and continued research is needed for understudied approaches and those with conflicting results.

Effective Strategies Not Implemented

Despite the availability of evidence-based and cost-effective interventions targeting individuals, groups, organizations, communities, and built environments, there is little indication that any of these interventions are being widely implemented in the United States. However, it is difficult to know with certainly because there is little surveillance of physical activity intervention implementation. There seems to be no systematic implementation of individually-tailored counseling or group behavior change programs. Even the most cost-effective intervention, signs encouraging stair use, is rarely applied. Almost all community-wide physical activity campaigns in the United States have been conducted in the context of research. Investment in pedestrian and bicycle facilities has remained less than 1 percent of Federal transportation funds, with the exception of a spike to 2 percent investment during the American Recovery and Reinvestment Act in 2009-2010.[36] Even though Web-based physical activity programs are plentiful, most are not evidence-based and have not been designed consistent with principles of effective behavior change.[37]

Physical education may be the largest societal investment in a physical activity intervention, and this is one of the few interventions for which some surveillance occurs. Though most States have physical education requirements, they vary widely and mainly address minutes of class time. Most physical education is not consistent with evidence-based practices of optimizing physical activity, and there is substantial evidence that on average, only about one-third of physical education class time is spent in physical activity.[38] However, in recent years 16 States have adopted laws or regulations requiring students to be active at least 50 percent of physical education class time or specifying that students must achieve a certain amount of physical activity throughout the school day, usually 30 minutes. Though this attention to physical activity in schools is encouraging, none of the regulations provide sufficient monitoring, consequences, or funding.[39] Thus, it is unlikely these regulations will have much effect.

There have been some recent major funding programs for environment, policy, and systems approaches to obesity prevention that included physical activity. "Stimulus" funding during the recession for Communities Putting Prevention to Work (CPPW; 2010-2012) and Affordable Care Act funding for Community Transformation Grants (CTG; 2012-2014) were administered by the CDC. However, the CTG program was cancelled before most interventions had begun. Both programs provided hundreds of millions of dollars for cities and States to support environment and policy interventions that were likely to be consistent with recent evidence,[34] so these programs may be among the largest commitments to physical activity interventions in U.S. history. Unfortunately, it is not clear how much of the funding was devoted to physical activity interventions, to what extent evidence-based interventions were supported, or what the effects were. Evaluation results are still pending.

Conclusions

The current status of the physical activity field is that the tremendous potential for public health benefit has not been achieved due to failure to widely implement the many evidence-based interventions. Physical inactivity is the fourth leading cause of death in the United States and worldwide,[2,5] and physical activity appears to be unique as a behavior that positively affects so many major diseases and conditions.[1] The need for improving physical activity is urgent, given the burden of disease, low prevalence rate, and flat (leisure time activities) or declining (active transportation) trends in both adults and adolescents.[21,22] The tools for increasing physical activity are available in the form of numerous evidence-based and cost-effective interventions. Thus, the physical activity research field has had dramatic successes in building evidence in several critical areas, though many questions remain. The biggest problem in the physical activity field is the failure to act on the evidence and make serious and well-funded efforts to implement evidence-based interventions.

Recommendations for Practice

Increased commitment to physical activity promotion is needed in the government, non-profit, and private sectors. Limited commitment can be seen at the Federal Government level, with CDC having a very small Physical Activity and Health Branch, the Department of Health and Human Services and NIH lacking offices for coordinating physical activity work, and the Department of Education having no person in charge of physical education. Most State public health departments

typically have one person devoted to physical activity, who may not be full time in that role and is often paid through a CDC grant. Increased funding for physical activity in all of these agencies is a prerequisite for making progress.

The lack of commitment also can be seen in the private sector, with most physical activity enterprises such as health clubs, dance studios, and Web sites not implementing evidence-based strategies. Insurance companies rarely pay for physical activity interventions, except for time-limited programs for people with specific diagnoses such as cardiac rehabilitation. These industries are encouraged to commit to broad implementation of evidence-based interventions. Partnerships with the public health sector and scientific community are encouraged to increase uptake of evidence-based interventions.

Recommendations for Research

Given that physical activity affects so many biological systems, it is reasonable to expect widespread effects on gene expression. Because of the possibility that inactive lifestyles could increase risk of disease in offspring, research on the effects of physical activity on gene expression and epigenetics should be increased.

Studies of the multiple economic and societal costs of physical inactivity and the co-benefits of physical activity should be prioritized. Economic data could be more effective than health outcome data to persuade decisionmakers to increase investments in physical activity interventions.

Evidence that physical activity, particularly active transportation, could contribute to reducing carbon emissions needs to be strengthened and quantified.[40] Exploring and quantifying diverse indirect and non-health impacts of active transportation and related built environment and policy interventions is a significant research need that would require collaborations between NIH and other agencies such as the Department of Transportation and the Environmental Protection Agency.

Reasons for the wide variation in physical activity across countries are not clear. A better understanding of effective strategies used around the world to promote physical activity should lead to better interventions that could be adapted and applied in the United States.

Measurement of physical activity is an ongoing challenge. Though substantial progress has been made in objective assessment using accelerometers and pedometers, further improvements in both electronic and self-report measures are needed. For example, integration of accelerometer measures with global positioning systems (GPS) and geographic information systems (GIS) allows assessment of the time and space dimensions of physical activity. This combination of methods can be used to improve assessment of both active and passive modes of transportation, as well as identify the times and places of active recreation. Mobile phones and commercial accelerometers could be used on a large scale for research and ongoing monitoring of physical activity in the population.[41] Another priority is to reduce the cost and simplify the use, data management, and data analysis of electronic measures such as accelerometers.

Self-report measures will continue to play an important role in research and public health surveillance.[42] One priority is to use electronic devices, such as smart phones and smart watches,

to improve self-reports by triggering queries based on time, place, or movement patterns. Another priority is to improve assessment of physical activity in the specific domains of leisure, transport, occupation, and household, because interventions need to be specific for each domain.

Enhanced electronic and self-report measures of physical activity could improve all facets of physical activity research, with the most critical contributions likely being (1) updating physical activity guidelines and (2) evaluating interventions. Almost all epidemiologic studies of physical activity and health outcomes used in creating physical activity guidelines were based on self-report measures, mainly of leisure time activity.[1] Thus, it is not known how the total dose of physical activity relates to health outcomes. It is important to apply objective and domain-specific self-report measures in epidemiological studies to refine understanding of dose-response relations that could lead to updated physical activity recommendations. Limited use of objective measures in intervention studies and excessive use of unvalidated self-reports likely lead to interpretation errors in intervention studies and underestimates of effects. Thus, better physical activity measures in intervention studies could enhance confidence in the results and perhaps reduce inconsistent findings across studies.

Though many effective interventions have been developed, there is a continuing need for stronger interventions as well as interventions tailored to high-risk populations, including demographic subgroups at high risk of inactivity-related diseases. A general recommendation for intervention research is to include cost-effectiveness analyses to assist practitioners and policymakers in selecting among intervention options. Research on individually-targeted interventions should focus on those that integrate technology, because Web- and mobile-based interventions have the potential for widespread impact at modest cost. Technology-based interventions should be assessed among people with low educational attainment, because graphics, photos, and videos may have special appeal for people with limited reading skills.

Multi-level interventions integrating efforts of multiple sectors (e.g., urban planning, transportation, parks, education) have the most potential to change population levels of physical activity over the long term. However, these interventions present several inherent design and methodological challenges, so special funding initiatives will be required to support progress, and more creative, rigorous, and long-term natural experiments are needed.

Because multi-level physical activity interventions require collaborations with decisionmakers in non-health sectors of society, research on effective communication of evidence-based strategies and their benefits to decisionmakers might accelerate the translation of research to practice and policy.

Since many evidence-based interventions for increasing population levels of physical activity exist, current research should focus on identifying how to more effectively translate these interventions into widespread practice. This requires testing implementation and dissemination strategies and adapting interventions for a new population or for scaling up activities.

Because of the profound health consequences, low prevalence rates, and unfavorable trends, a much higher priority on physical activity research is justified.

Authors' Affiliations

James F. Sallis, PhD, Department of Family and Preventive Medicine, University of California, San Diego. Jordan A. Carlson, PhD, Department of Family and Preventive Medicine, University of California, San Diego

Address correspondence to: James F. Sallis, PhD, Department of Family and Preventive Medicine, University of California, San Diego, 3900 Fifth Avenue, Suite 310, San Diego, CA 92103; email jsallis@ucsd.edu.

References

1. Physical Activity Guidelines Advisory Committee. Physical Activity Guidelines Advisory Committee report, 2008. Accessed September 8, 2014. Washington, DC: U.S. Department of Health and Human Services; 2008. Available at http://www health.gov/paguidelines/Report/pdf/CommitteeReport.pdf. Accessed October 6, 2014.
2. Physical inactivity: a global public health problem. Geneva: World Health Organization; 2011. Available at http://www.who.int/dietphysicalactivity/factsheet_inactivity/en/. Accessed October 6, 2014.
3. Kohl HW, Craig CL, Lambert EV, et al. The pandemic of physical activity: global action for public health. Lancet 2012;380:294-305.
4. Lee, I-M, Shiroma FJ, Lobelo F, et al. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. Lancet 2012;380:219-29.
5. Danaei G, Ding EL, Mozaffarian D, et al. The preventable causes of death in the United States: comparative risk assessment of dietary, lifestyle, and metabolic risk factors. PLoS Med 2009;6(4):epub.
6. Franco OH, de Laet C, Peeters A, et al. Effects of physical activity on life expectancy with cardiovascular disease. Arch Intern Med 2005;165:2355-60.
7. Crimmins EM, Preston SH, Cohen B (Eds). Explaining divergent levels of longevity in high-income countries. Washington, DC: National Academies Press; 2011.
8. Carlson SA, Fulton JE, Pratt M, et al. Inadequate physical activity and health care expenditures in the United States. Prog Cardiovasc Dis 2014; online first.
9. Kramer AF, Erickson KI. Effects of physical activity on cognition, well-being, and brain: human interventions. Alzheimers Dement 2007;3:S45-S51.
10. Heyn P, Abreu BC, Ottenbacher KJ. The effects of exercise training on elderly persons with cognitive impairment and dementia: a metaanalysis. Arch Phys Med Rehabil 2004;85:1694-704.
11. Voss MW, Nagamatsu LS, Liu-Ambrose T, et al. Exercise, brain, and cognition across the life span. J Appl Physiol 2011;111:1505-13.
12. Sibley B, Etnier J. The relationship between physical activity and cognition in children: a meta-analysis. Pediatr Exerc Sci 2003;15:243-56.
13. Hillman CH, Pontifex MB, Raine LB, et al. The effect of acute treadmill walking on cognitive control and academic achievement in preadolescent children. Neuroscience 2009;159:1044-54.
14. Owen AM, Hampshire A, Grahn JA, et al. Putting brain training to the test. Nature 2010;465:775-8.
15. Green CS, Bavelier D. Exercising your brain: a review of human brain plasticity and training-induced learning. Psychol Aging 2008;23:692-701.
16. Lustig C, Shah P, Seidler R, et al. Aging, training, and the brain: a review and future directions. Neuropsychol Rev 2009;19:504-22.
17. Deng W, Aimone JB, Gage FH. New neurons and new memories: how does adult hippocampal neurogenesis affect learning and memory? Nat Rev Neurosci 2010;11:339-50.
18. Global recommendations on physical activity for health. Geneva: World Health Organization; 2011. Available at http://www.who.int/dietphysicalactivity/factsheet_recommendations/en/. Accessed October 6, 2014.
19. Hallal PC, Andersen LB, Bull FC, et al. Global physical activity levels: surveillance progress, pitfalls, and prospects. Lancet 2012;380:247-57.
20. Troiano RP, Berrigan D, Dodd KW, et al. Physical activity in the United States measured by accelerometer. Med Sci Sports Exerc 2008;40:181-8.
21. Healthy People 2020. Washington, DC: U.S. Department of Health and Human Services; 2014. Available at http://www healthypeople.gov/2020/default.aspx. Accessed October 6, 2014.
22. McDonald NC. Active transportation to school: trends among U.S. schoolchildren, 1969–2001. Am J Prev Med 2007;32:509-16.

23. Brownson RC, Boehmer TK, Luke DA. Declining rates of physical activity in the United States: what are the contributors? Annu Rev Public Health 2005;26:421–43.
24. Bauman AE, Reis RS, Sallis JF, et al. Correlates of physical activity: why are some people physically active and others not? Lancet 2012;380:258-71.
25. Ingram DK. Age-related decline in physical activity: generalization to nonhumans. Med Sci Sports Exerc 2000;32:1623-9.
26. Sallis, J.F., Floyd, M.F., Rodriguez, D.A., and Saelens, B.E. (2012). The role of built environments in physical activity, obesity, and CVD. Circulation, 125, 729-737.
27. Marcus BH, Williams DM, Dubbert PM, et al. Physical activity intervention studies: what we know and what we need to know. Circulation 2006;114:2739-52.
28. Sallis JF, Owen N, Fisher EB. Ecological models of health behavior. In Glanz K, Rimer BK, Viswanath K (Eds), Health behavior and health education: theory, research, and practice. 4th ed. San Francisco: Jossey-Bass; 2008. p. 485.
29. Sallis JF, Cervero RB, Ascher W, et al An ecological approach to creating more physically active communities. Annu Rev Pub Health 2006;27:297-322.
30. Kahn EB, Ramsay LT, Brownson et al. The effectiveness of interventions to increase physical activity: a systematic review. Am J Prev Med 2002;22:S73-S107.
31. Wu S, Cohen D, Shi Y, et al. Economic analysis of physical activity interventions. Am J Prev Med 2011;40:149-58.
32. Heath GW, Brownson RC, Kruger J, et al. The effectiveness of urban design and land use and transport policies and practices to increase physical activity: a systematic review. J Phys Activ Health 2006;3:S55-S76.
33. Heath GW, Parra-Perez D, Sarmiento OL, et al. Evidence-based physical activity interventions: lessons from around the world. Lancet 2012;380:45-54.
34. Mozaffarian D, Afshin A, Benowitz NL, et al. Population approaches to improve diet, physical activity, and smoking habits: a scientific statement from the American Heart Association. Circulation 2013;126:1514-63.
35. President's Council on Fitness, Sports, and Nutrition. Physical activity guidelines for Americans mid-course report: strategies to increase physical activity among youth. Washington, DC: U.S. Department of Health and Human Services; 2013. Available at http://www.fitness.gov/be-active/physical-activity-guidelines-for-americans/. Accessed October 6, 2014.
36. Federal Highway Administration. Bicycle & Pedestrian. Washington, DC: U.S. Department of Transportation; 2014. Available at: http://www fhwa.dot.gov/environment/bicycle_pedestrian/funding/bipedfund.cfm. Accessed October 6, 2014.
37. Doshi A, Patrick K, Sallis JF, et al. Evaluation of physical activity web sites for use of behavior change theories. Ann Behav Med 2003;25:105-11.
38. Kohl HW, Cook HD (Eds). Educating the student body: taking physical activity and physical education to school. Washington, DC: National Academies Press; 2013.
39. Carlson JA, Sallis JF, Chriqui JF, et al. State policies about physical activity minutes in physical education or during the school day. J School Health 2013;83:150-6.
40. Frank LD, Greenwald MJ, Winkelman S, et al. Carbonless footprints: promoting health and climate stabilization through active transportation. Prev Med 2010;50:S99-S105.
41. Intille SS, Lester J, Sallis JF, et al. New horizons in sensor development. Med Sci Sports Exerc 2012;44, S24-S31.
42. Matthews CE, Moore SC, George SM, et al. Improving self-reports of active and sedentary behaviors in large epidemiologic studies. Exerc Sport Sci Rev 2012;40:118-6.

James F. Sallis, PhD, is Distinguished Professor of Family Medicine and Public Health at the University of California, San Diego and Director of Active Living Research, a program of the Robert Wood Johnson Foundation. His primary research interests are promoting physical activity and understanding policy and environmental influences on physical activity, nutrition, and obesity. Time Magazine has identified Dr. Sallis as an "obesity warrior." He has authored over 600 scientific publications and is one of the world's most cited authors in the social sciences. Dr. Sallis is President-elect of the Society of Behavioral Medicine.

Jordan A. Carlson, PhD, MA, is Director of Community-Engaged Health Research, Children's Mercy Hospital, and Assistant Professor of Pediatrics, University of Missouri-Kansas City School of Medicine. His research interests include active living, school-based physical activity, neighborhood walkability, improving uptake and implementation of physical activity interventions, and physical activity measurement technology.

How Health Impact Assessments Shape Interventions

Steven M. Teutsch, Katherine M. Butler,
Paul A. Simon, and Jonathan E. Fielding

Abstract

Despite paying more for medical care than any other nation, health in the United States lags other developed nations. This is not altogether surprising given that the largest contributors to overall health are the social and physical environments and health behaviors, areas in need of greater investment. To understand the health consequences of policy and program interventions in these areas requires a very different approach than the reductionist, biomedical model of disease. Complex studies are required to understand how education and income, family and other social structures, community resources, and the natural and built environments interact to shape health and well-being. Indeed, by their very nature these factors are contextual and change over time. Outcomes often occur long into the future, making longitudinal studies difficult and costly. However, we can harness and synthesize existing information to identify and understand the consequences of interventions that naturally occur. Health impact assessments (HIAs) are a practical way to identify and understand the key health benefits and harms that may result from an intervention or be of concern to stakeholders, estimate the importance of each, recommend steps to enhance benefits or reduce harms, communicate findings to key stakeholders, and track the impact of the assessment on the decisionmaking process. We present three case studies to show how HIAs can influence policy choices.

Introduction

It is now widely recognized that health in the United States lags other developed nations and, worse, is slipping further and further behind.[1] The single measure where the United States far exceeds competitor nations is the amount we spend on clinical care. The concomitant toll these factors take on the Nation's productivity is enormous. The excessive costs are functionally a tax on our international competitiveness, as ill health increases health care costs to business as well as lost productivity from absenteeism and presenteeism.[2] Our investment in basic and applied clinical research as well as clinical care has been second to none,[3] yet that investment fails to adequately address the primary drivers of health which lie outside the clinical care system. Health is commonly attributed 20 percent to clinical care, 30 percent to health behaviors, 40 percent to the social environment, and 10 percent to the built and natural environments.[4] Failure to invest in the underlying drivers of health contributes to excessive medical care costs. In embracing this phenomenon, the Triple Aim (better health, better care, lower cost)[5] recognizes that health will only be substantially improved if we attend to the underlying determinants of health. The larger problem will not be solved by more and better medical care alone.

To address behaviors, the social and physical environments require a very different approach than the reductionist, biomedical model of disease. Complex studies are required to understand how education and income, family structures and community resources, and the natural and built environments interact to shape health and well-being. Indeed, by their very nature these factors are contextual and change over time. Outcomes often occur long into the future making longitudinal studies difficult and costly. Most importantly, though, we can harness and synthesize existing information to identify and understand the consequences of interventions that naturally occur. Models are particularly important to assess long-term effects, since interventions to address them are challenging to conduct and generally not amenable to the most rigorous randomized study designs.[6]

The "Health in All Policies" (HIAP) framework recognizes the importance of policies and activities in other sectors in shaping health.[7] A transportation project can influence physical activity patterns or contribute to pollution and asthma. Living wage laws can reduce stress and free-up time for parenting. Yet decisions in other sectors have traditionally not considered the health consequences of those decisions or what can be done to ameliorate them. Growing recognition that these issues are critical to health as well as the salience of health in decisionmaking has fostered the growth of HIAP internationally and in the United States.[8]

Good decisions require good information. However, definitive studies about the consequences of many decisions are seldom available. This is due in part to the complexity of policies and the context in which they will be implemented, but it is also because the policies affect upstream social and environmental determinants which have many ramifications that affect health. To understand even the most important health aspects requires the synthesis of information from many sources. To make the task even more challenging, decisions are often made over relatively short periods of time that limit the opportunity to conduct new studies.

Assessing Interventions

While there is growing recognition of the impact of the social environment on health, the evidence on the effectiveness of interventions remains scant. However important the issues may be for society, there is no equivalent to a private biomedical industry ready to capitalize on the findings of basic research and develop profitable products; and few resources for research are available in the not-for-profit or government services sectors. But additionally, demonstrating the effectiveness of social sector policies and programs on health has intrinsic challenges. While social factors have broad and profound impacts on health, they generally are less specific than outcomes of typical clinical interventions. The complex interrelationships differ mightily from the more tidy pathophysiological pathways of specific diseases. Hence the effectiveness of social and environmental interventions on health is inherently more difficult to measure.

The problem is compounded by the many social, personal, and environmental factors that contribute to health outcomes but vary over time. These confounders are rarely fully understood or measured, despite the fact that their interaction effects can be large. Short-term intervention effects are usually about processes – was the intervention implemented with high fidelity? What was the participation rate? Intermediate and long term health outcomes are scarce. To overcome these challenges,

investigators pull information from multiple sources and synthesize it using techniques of systematic literature reviews, meta-analyses, and simulation modeling, which can be used to project longer-term health and economic outcomes.

From Application to Real-Time Decisionmaking

Programmatic and policy decisions need to be informed by the best available information. In fact, decisions are commonly made in the face of great uncertainty even when few data are available. In non-health sectors, health outcomes frequently are not considered at all, largely because they have neither appeared central to the decision process nor has anyone solicited or provided such information. Health impact assessments (HIAs) can fill this gap.

Health Impact Assessment

HIA is an important bridge between research and practice. It is a tool to identify and understand the key health benefits and harms that are of concern to stakeholders and those known to result from an intervention, estimate the importance of each, recommend steps to enhance benefits or reduce harms, communicate findings to key stakeholders, and track the impact of the assessment on the decisionmaking process. The National Research Council has described the process.[9] Figure 1 shows the steps in conducting an HIA. We present three case studies of how HIAs can influence policy choices. Each HIA is described more fully elsewhere.

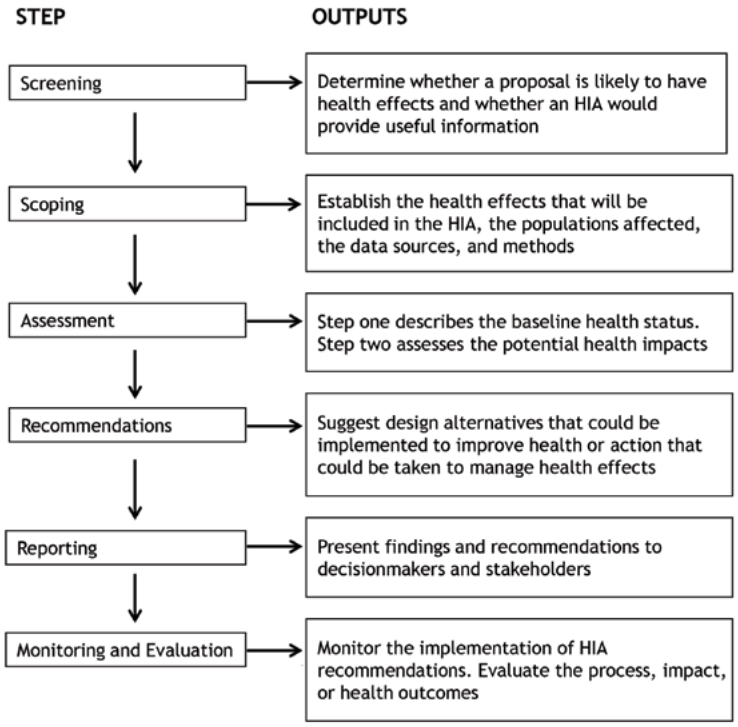


Figure 1. Framework for health impact assessment (HIA), summarizing steps and outputs

Source: Adapted from improving health in the United States: the role of health impact assessment. Washington, DC: National Academies Press; 2011. Used with permission.

Case Study 1: Menu Labeling

Laws that require menu labeling[a] at large chain food restaurants as a means of reducing the obesity epidemic began generating interest and legislative support as early as 2006 with the passage of a local ordinance in New York City. The rationale for this strategy is at least threefold. Studies suggest that the dramatic growth in per capita consumption of restaurant food contributes heavily to the U.S. obesity epidemic.[10] Restaurant super-sizing of food and beverage portions has become widespread and, unlike mandated calorie and nutrition information on packaged foods, such information is not readily apparent to customers at the point of purchase. In addition, studies show that most people, including nutritionists, greatly underestimate the caloric content of restaurant menu items.[11]

In California, a menu labeling bill (Senate Bill 120)[b] was introduced in 2007 to require posting calorie information on menus and menu boards at all chain restaurants with 15 or more in-State outlets. The public broadly supported the bill, but it was actively opposed by the California Restaurant Association and other trade organizations. After bruising political debate, it was passed by the State Assembly and Senate but vetoed by Governor Schwarzenegger.

A similar bill (Senate Bill 1420)[c] was introduced in 2008 but was limited to chain restaurants with 20 or more in-State venues. To inform the decisionmaking process on the bill, the Los Angeles County (LAC) Department of Public Health staff reviewed related research literature to assess the strategy's potential impact on obesity. Finding very limited direct information, the department initiated an HIA using small studies and assumptions to project the potential impact of a menu labeling law on obesity in LAC, all else being equal. In the base case, it assumed that in response to calorie postings, 10 percent of restaurant patrons ordered reduced-calorie meals, resulting in an average reduction of 100 calories per meal. This is not a large reduction; changing from a large to a medium soft drink would save 95 calories, from a large to medium order of French fries saves 163 calories, and from a double meat to a single meat burger would save 244 calories. The analysis (see Table 1) found that menu labeling would avert 41 percent of the 6.75 million pound average annual weight gain in the LAC population aged 5 years and older.[12] A sensitivity analysis (see Table 2) suggested that substantially larger impacts could be realized if more patrons ordered reduced-calorie meals or if average per-meal calorie reductions increased. On the other hand, fewer individuals may actually reduce their consumption, and those who do may not achieve 100 calorie reductions in the meals they actually order. More perversely, some might actually increase their purchase of high-calorie offerings believing they provide better value! In addition, the study does not consider that reductions may be attenuated over time.

---

[a] Adapted from Paul Simon and Suzanne Bogert, Tackling toxic food environments: A response to the obesity epidemic. In Fielding JE, Teutsch SM, eds. Public health practice: what works. Oxford University Press; 2013.

[b] See California Legislative Information, SB-120, Food Facilities: Nutritional Information. Available at http://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=200720080SB120.

[c] See California Nutrition Law: Information and Enforcement (SB-1420). Available at http://nutrition.levitas.com/.

Table 1. Projected impact of menu labeling

Item No.	Metric	Estimate	Basis
1	Total annual restaurant revenue, Los Angeles County (LAC)	$14,600,000,000	Statewide estimate from the National Restaurant Association pro-rated by LAC's percentage of the State population
2	Large chain restaurant market share, 15 or more stores in California	51%	Extrapolated from NPD Group, 2005*
3	Large chain restaurant revenue, LAC	$7,446,000,000	Calculated from items 1 and 2
4	Average price per meal in large chains (sit-down and fast food)	$7.48	Based on 1992 national meal price estimates, adjusted for inflation
5	Annual number of meals served, LAC	995,454,545	Calculated from items 3 and 4
6	Annual number of meals served, ages 0 to 5 years	36,500,000	LAC Health Survey (2005) **
7	Annual number of meals served, ages 5 and older	958,954,545	Calculated from items 5 and 6
8	Percentage of reduced calorie meals selected as a result of menu labeling	10%	Extrapolated from Burton et al. (2008) ***
9	Annual number of reduced-calorie meals	95,895,455	Calculated from items 7 and 8
10	Average calorie reduction per meal	100	Unpublished data, Bassett, et al. (2008) †
11	Total annual number of reduced calories attributed to menu labeling	9,589,545,455	Calculated from items 9 and 10
12	Calories per pound of weight	3,500	Duyff (2002) ‡
13	Total annual pounds of weight loss attributable to menu labeling	2,739,870	Calculated from items 11 and 12
14	Average annual weight gain, ages 18 and older	5,500,000	Calculated using data from the 1997 and 2005 LAC Health Survey**
15	Average annual weight gain, ages 5-17	1,250,000	Calculated using data from the 1999 and 2006 California Physical Fitness Testing Program§
16	Average annual weight gain, ages 5 and older	6,750,000	Calculated from items 14 and 15
17	Percentage of population weight gain averted due to menu labeling	40.6%	Calculated from items 13 and 16

Source: Adapted from Simon P, Jarosz CJ, Kuo T, et al. Menu labeling as a potential strategy for combating the obesity epidemic: a health impact assessment. Los Angeles County Department of Public Health, 2008. Available at http://www.publichealthadvocacy.org/printable/CCPHA_LAPHmlaspotentialstrategy.pdf. Used with permission.

*Cited in the U.S. District Court Declaration of Thomas R. Frieden, Commissioner of the New York City Department of Health and Mental Hygiene, July 5, 2007 (pg. 31).

**Los Angeles County Department of Public Health, Office of Health Assessment and Epidemiology, Health Assessment Unit, 2005 Los Angeles County Health Survey. Estimates are based on self-reported data by a random sample of 8,648 Los Angeles County adults and 6,032 parents/guardians of children 0-17 years, representative of the population of Los Angeles.

***Burton S, Cryer EH, Kees J, et al. Attacking the obesity epidemic: the potential health benefits of providing nutrition information in restaurants. Am J Publ Health 2006;96(9):1669-75.

†Bassett MT, Dumanovsky T, Huang C, et al. Purchasing behavior and calorie information at fast-food chains in New York City, 2007. Am J Publ Health 2008;98(8):1457-9.

‡Duyff RL. Complete Food and Nutrition Guide, 2nd ed. American Dietetic Association. Hoboken, NJ: John Wiley & Sons; 2002.

§California Department of Education. Physical Fitness Testing (PFT) program Web page. Available at: http://www.cde.ca.gov/ta/tg/pf/. Accessed February 1, 2008.

Table 2. Sensitivity analysis of menu labeling impact on percentage of population weight gain averted

Average Amount of Calorie Reduction	Percentage (%) of Patrons Who Purchase a Lower-Calorie Meal as a Result of Menu Labeling				
	10	20	30	40	50
25	10.1	20.3	30.4	40.6	50.7
50	20.3	40.6	60.9	81.2	101.5
75	30.4	60.9	91.3	121.8	152.2
100	40.6*	81.2	121.8	162.4	203.0
125	50.7	101.5	152.2	203.3	253.7
150	60.9	121.8	182.7	243.5	304.4
175	71.0	142.1	213.1	284.1	355.2
200	81.2	162.4	243.5	324.7	405.9

*Base Case

Source: Adapted from Simon P, Jarosz CJ, Kuo T, et al. Menu labeling as a potential strategy for combating the obesity epidemic: a health impact assessment. Los Angeles County Department of Public Health, 2008. Available at http://www.publichealthadvocacy.org/printable/CCPHA_LAPHmlaspotentialstrategy.pdf.

Despite the limitations noted above, the report influenced policy. The May 2008 study report came at the height of the public debate on the menu labeling bill, and it garnered extensive media coverage. Several department staff testified before State legislative committees on the HIA findings and on the toll of the obesity epidemic. One sponsor of the bill reported that the HIA was instrumental in negotiations with legislators and the governor. In response to the report, the County Board of Supervisors voted to implement a county menu labeling ordinance if the State bill was not enacted. In turn, the California Restaurant Association lowered its opposition, recognizing the advantages of a uniform statewide measure compared to the threat of this and similar action in other counties, enacting a patchwork of varying county ordinances. The bill was passed and signed into

law by the Governor. Two years later, a similar measure was approved at the national level as part of Federal health care reform, pre-empting the California law.

The Department of Public Health's Environmental Health Division staff has begun to assess compliance with posted calorie counts on menus at large chain restaurants as part of their routine restaurant inspections. The Environmental Health staff intends to collect food specimens from randomly sampled chain restaurants to check the accuracy of posted calorie information. Grants also will be sought to conduct pre- and post-implementation surveys of customers at chain restaurants to assess the law's impact on menu selections and calories consumed. The calorie content of restaurant offerings will be monitored before and after implementation of the law to determine if restaurant operators have modified recipes to reduce the calorie content of menu items. Finally, the department will track the trajectory of the obesity epidemic, though the independent effects of the menu labeling law may be difficult to ascertain in the context of many complementary obesity prevention efforts.

Studies of the impact of menu labeling in New York City and Seattle have produced mixed results, some showing modest reductions in the caloric content of food purchases, others showing no effect.[13-15] However, these short-term studies reflect consumer response relatively soon after menu labeling began, and it will be important to replicate them to see whether the impacts grow or diminish as customers become more familiar with the calorie information. Results may depend on the degree to which the intervention is accompanied by community education that promotes the use and interpretation of the calorie information. Lastly, menu labeling is but one component of a comprehensive healthy diet strategy.

Perhaps the most interesting early impact of menu labeling is the restaurant industry's response. Newspaper articles and industry trade publications and Web sites report that many chains are reformulating their menus, reducing the caloric content of standard items, or adding new low-calorie options. We are unaware of any formal assessment or quantification of this phenomenon, but the anecdotal reports suggest that the menu labeling policy may have additional benefits not measured in the department's HIA.

Case Study 2: Equal Employment Opportunity and Mental Health

In 2004, major U.S. cities began revising employment policies to prevent discrimination against people with criminal records, which has a disproportionate effect on minority communities.[16] This spurred a larger trend of looking into legislation at the State and Federal levels – an initiative widely known as "ban the box" to remove the box on employment applications that ask about criminal records. The ban-the-box movement and other efforts to revisit hiring laws coincided with an increased number of U.S. adults with arrest or conviction records and employers conducting background checks.[17]

In response to a proposed revision to employment policy in Illinois, the Adler School of Professional Psychology Institute on Social Exclusion initiated a Mental Health Impact Assessment (MHIA) in

2011 through work with Chicago's Englewood community to examine the impacts of using arrest information on mental health and well-being. The MHIA process introduced a "mental health lens" to an area of law typically evaluated in the context of civil rights or economic analyses. Peer-reviewed literature was available to describe the relationships between arrest, employment, and mental health; however, evidence was lacking to describe the interplay of important social determinants of health, such as social exclusion, income, and neighborhood conditions (Figure 2). To fill these data gaps, the MHIA employed a wide range of research activities such as community surveys, focus groups, and employer interviews in Englewood. Focus groups of job-seekers with arrest or other criminal records reported feeling "depressed," "hopeless," and "discouraged." Based on the synthesis of evidence from the literature review, surveys, and interviews, the MHIA predicted that the updated employment policy prohibiting blanket exclusions of people with criminal records would likely decrease the severity of depression and psychological distress.[18]

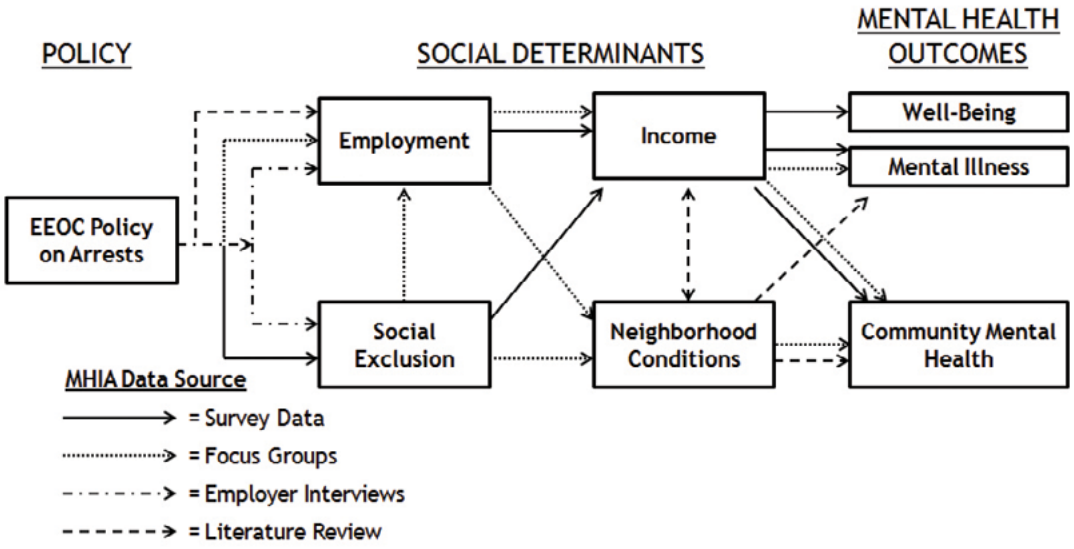


Figure 2. Assessing mental health impacts of EEOC policy revisions with multiple data sources

Source: Todman L, Taylor JS, McDowell T, et al; Adler School of Professional Psychology, Institute on Social Exclusion. U.S. Equal Employment Opportunity Commission policy guidance: a mental health impact assessment; 2013. Used with permission.

Note: EEOC = Equal Employment Opportunity Commission; MHIA = mental health impact assessment.

The efforts in Illinois to revise the employment policy were thwarted. However, a similar initiative to enforce anti-discrimination laws at the Federal level was developing. The Equal Employment Opportunity Commission (EEOC) was seeking to provide updated research to strengthen its existing employment policy and soliciting input from stakeholders through a series of public meetings in 2011 and 2012. Over the course of the public comment periods, the Adler School of Professional Psychology submitted recommendations to the EEOC to emphasize the importance of mental health considerations in the guidance. The EEOC Commissioner recognized the importance of the MHIA recommendations and commented that the guidance was updated to reflect additional health research and analysis provided by the Adler School, with examples of how arrest and criminal records can have disparate impacts on minority populations.[19] The 2012 final revision

of the EEOC policy guidance states that arrest records cannot be routinely used as a basis for exclusion, and employers must justify exclusion with job-related reasons.

Aside from informing decisionmakers of mental health considerations, possibly one of the most significant impacts of the MHIA was the process itself. The MHIA team involved community members in every step of the process, from creating the research objectives to developing policy recommendations. Local youth helped to design survey questions, and community members were trained to carry out research tasks. Throughout this process, the Englewood residents became more aware of employment and civil rights law and developed a better understanding of their community health conditions.

By evaluating the proposed revision to EEOC's policy and contributing to social science research content in the policy, HIA proved to be an important tool to understand mental health impacts of an issue in unchartered territory for public health practice – the intersection of employment, criminal justice and civil rights.

Case Study 3: Living Wage Ordinance

In 2006, San Francisco's legislative board was in the midst of considering adoption of a living wage of $11 per hour for municipal workers and commissioned an economic analysis to study the potential effects. Since poverty is one of the strongest determinants of poor health,[20] the San Francisco Department of Public Health (SFDPH) recognized the value of conducting an HIA on the impending living wage decision.

SFDPH was invited to join the city legislators during policy discussions, and staff provided testimony on the potential health benefits of increasing the minimum wage. The SFDPH pointed out five key considerations regarding the relationship between income and health.[21] Aside from low income wage earners not receiving health care benefits, low income is linked to poor health and disease. Low income neighborhoods lack access to basic needs, such as adequate food and housing. Difficult financial circumstances make it challenging to socialize with friends and family, even calling out of town relatives or inviting friends over to enjoy a meal together; this financial constraint may contribute to depression and social isolation. Low income neighborhoods tend to have fewer options for safe access to physical activity and public services, and therefore health-promoting activities are less common. Obesity, for example, is associated with socioeconomic conditions. And lastly, parents who are stressed financially may have to work more than one job, juggle irregular hours, and have less time to spend reading and conversing with their children. This may result in lower literacy rates and slower development of verbal skills among children.

In a quantitative approach to predict potential health benefits of adopting a living wage, SFDPH provided the San Francisco Board of Supervisors Budget and Finance Committee with estimates of decreased premature mortality risk, decreased sick days, and increased chances of children of workers completing high school (Table 3).[22] Overall, the analysis concluded that adoption of an $11 per hour living wage may have long-term positive benefits on individual and community health. The HIA contributed significant health research findings to the policy debate, and in 2003, San Francisco increased the minimum wage from $6.75 to $8.50 per hour for over 50,000 city workers. As of 2014,

the San Francisco minimum wage was increased to $10.74, approaching the $11 per hour examined in the HIA. An increase to $15 per hour is currently under consideration for the November 2014 ballot.

Table 3. Estimated health and educational effects on workers and their children resulting from adoption of a living wage for families with incomes of $20,000: San Francisco Bay Region, California, 1997-1999

Study/Outcome	Model	Effect Measure	Estimate for Full-time Workers (95% CI)	Estimate for Part-time Workers (95% CI)
Backlund (1996)*				
Mortality – male	Proportional hazards[a]	Hazard ratio	0.94 (0.92, 0.97)	0.97 (0.96, 0.98)
Mortality – female	Proportional hazards	Hazard ratio	0.96 (0.95, 0.98)	0.98 (0.97, 0.99)
Ettner (1996) †				
Health status	Ordered probit[b]	Relative risk	0.94 (0.93, 0.96)	0.97 (0.96, 0.98)
ADL limitations	Probit	Relative risk	0.94 (0.95, 0.98)	0.98 (0.97, 0.99)
Work limitations	Probit	Relative risk	0.94 (0.92, 0.96)	0.97 (0.95, 0.98)
CES-Depression scale	2-part[c]	Elasticity	-1.9%	-1.1%
Number of sick days	2-part	Elasticity	-5.8%	-3.2%
Alcohol consumption	2-part	Elasticity	+2.4%	+1.3%
Duncan (1998) ‡				
Completed schooling	OLS regression	Years of schooling	0.25 (0.20, 0.30)	0.15 (0.12, 0.17)
Completed high school	Logistic regression	Odds ratio	1.34 (1.20, 1.49)	1.18 (1.11, 1.26)
Nonmarital childbirth	Proportional hazards	Hazard ratio	0.78 (0.69, 0.86)	0.86 (0.81, 0.92)

Source: Bhatia R, Katz M. Estimates of health benefits from a local living wage ordinance. Am J Pub Health 2001;91(9):1398-1402. Used with permission.

Note: CI = confidence internval; ADL = activities of daily living; CES = Center for Epidemiologic Studies; OLS – ordinary least squares.

a = Effect measures for the 24- to 44-year age groups were used.

b = The probit models required specifying the values of all the model covariates; the values given above were calculated for a married 30-year White female with 2 children living in a metropolitan area.

c = The 2-part model used least squares regression on a log transformation of the dependent variable, with a conditional sample of subjects with positive values used for the outcome. The effect measure, elasticity, did not enable us to calculate confidence intervals.

*Backlund E, Sorlie P, Johnson N. The shape of the relationship between income and mortality in the United States. Evidence from the National Longitudinal Mortality Study. Ann Epidemiol 1996;6(1):12–20.

†Ettner SL. New evidence on the relationship between income and health. J Health Econ 1996;15(1):67–85.

‡Duncan GJ, Yeung W, Brooks-Gunn J, et al. How much does childhood poverty affect the life chances of children? Am Sociol Rev 1998;63(3):406–24.

Discussion

The most important determinants of health and the underlying causes of health disparities lie outside the traditional health sector, are woven into the very fabric of our daily lives, and shape our health behaviors. Social science research undergirds our understanding of how the social environment affects health. That knowledge compels the population health system to work with other sectors to assure that health consequences are regularly considered in programmatic and policy decisions. HIAs provide a systematic way to assess the potential health effects of interventions and communicate the results to decisionmakers, community organizations, and other stakeholders. By engaging stakeholders and identifying health consequences of concern, HIAs can assess the size and strength of those outcomes and provide guidance on how benefits can be enhanced or harms mitigated.

Implications for Research and Practice

Although decisions often are made before all the desired evidence is available, decisionmakers need the best available scientific information. HIAs provide baseline characteristics and health projections, then estimate the changes that would occur if policies or programs were implemented, using systematic searches of the scientific literature as well as additional studies conducted to fill in important gaps. Existing meta-analyses may be used or new ones conducted. Focus groups and interviews of key opinion leaders use well-established qualitative social science methods to shed light on local issues and context. These inquiries often elicit perspectives that need to be included in HIAs – in other words, they are a mechanism for empowering communities and assuring that their concerns are captured, substantiated, and communicated.

Surveys and analyses of existing data provide quantitative information that can be incorporated into models to predict health outcomes. Nevertheless, significant information gaps are commonly encountered, highlighting gaps that need to be informed by carefully performed research. In the absence of high-quality studies, expert opinion complemented by sensitivity analyses is commonly used. Although economic and financial analyses are not regularly included in HIAs, they may be of keen interest to decisionmakers. Understanding the financial costs and benefits of an action, and to whom they accrue, can help identify strategies for gain sharing or mechanisms to make decisions more palatable to stakeholders.

HIAs can help engage decisionmakers and broaden their understanding of the health consequences of their choices. Since decisionmakers and communities acknowledge that health is one of the most important considerations in many policies, HIAs can robustly meet that need. They provide an important addition to the policy-analytic arsenal and can facilitate improvements in the social and physical environments.

Acknowledgements

We acknowledge the assistance provided by Dr. Lynn Todman and Dr. Tiffany McDowell, who are primary authors of the MHIA on Equal Employment Opportunity and Mental Health. The opinions presented herein are those of the authors and may not necessarily represent the position of the Agency for Healthcare Research and Quality, the National Institutes of Health, the U.S. Department of Health and Human Services, or the Los Angeles County Department of Public Health.

Authors' Affiliations

Los Angeles County Department of Public Health (ST, KB, PS, JEF); University of California Los Angeles (UCLA) Fielding School of Public Health (ST, PS, JEF); and UCLA Geffen School of Medicine (JEF), Los Angeles, CA.

Address correspondence to: Steven Teutsch, 841 Moon Ave., Los Angeles, CA 90065.
email steventeutsch@gmail.com.

References

1. Organization for Economic Cooperation and Development. Health at a glance 2013: OECD indicators. OECD Publishing, 2013. DOI: 10.1787/health_glance-2013-en.
2. McCullough JM, Zimmerman FJ, Fielding JE, et al. A health dividend for America: the opportunity cost of excess healthcare expenditures in the United States. Am J Prev Med 2012;43(6):650-4. PMID: 23159261.
3. Chakma J, Sun GH, Steinberg JD, et al. Asia's ascent — global trends in biomedical R&D expenditures. N Engl J Med 2014;370(1):3-6. PMID: 24382062.
4. Booske BC, Athens JK, Kindig DA, et al. County health rankings working paper: different perspectives for assigning weights to determinants of health. Seattle, WA: University of Washington, Population Health Institute; 2010.
5. IHI Triple Aim Initiative. Cambridge, MA: Institute for Healthcare Improvement; 2013. Available at http://www.ihi.org/Engage/Initiatives/TripleAim/pages/default.aspx. Accessed September 24, 2014.
6. Institute of Medicine. For the public's health: revitalizing law and policy to meet new challenges. Washington, DC: National Academies Press; 2011.
7. Rudolph L, Caplan J, Mitchell C, et al. Health in all policies: improving health through intersectoral collaboration. Washington, DC: National Academies Press; 2013.
8. Kickbusch I, Buckett K (Eds). Implementing health in all policies: Adelaide 2010. Adelaide, Australia: Government of South Australia, Department of Health; 2010.
9. National Research Council. Improving health in the United States: the role of health impact assessment. Washington, DC: National Academies Press; 2011.
10. Kant AK, Graubard BI. Eating out in America, 1987–2000: trends and nutritional correlates. Prev Med 2004;38(2):243–9. PMID: 14715218.
11. Backstrand JR, Wootan MG, Young LR, et al. Fat chance: a survey of dietitians' knowledge of the calories and fat in restaurant meals. Washington, DC: Center for Science in the Public Interest; 1997.
12. Kuo T, Jarosz CJ, Simon P, et al. Menu labeling as a potential strategy for combating the obesity epidemic: a health impact assessment. Am J Public Health 2009;99(9):1680–6. PMID: 19608944.
13. Bassett MT, Dumanovsky T, Huang C, et al. Purchasing behavior and calorie information at fast-food chains in New York City, 2007. Am J Public Health 2008;98(8):1457–9. PMID: 18556597.
14. Elbel B, Kersh R, Brescoll VL, et al. Calorie labeling and food choices: a first look at the effects on low-income people in New York City. Health Aff 2009;28(6):w1110–21. PMID: 19808705.

15. Finkelstein EA, Strombotne KL, Chan NL, et al. Mandatory menu labeling in one fast-food chain in King County, Washington. Am J Prev Med 2011;40(2):122–7. PMID: 21238859.
16. National Employment Law Project. Ban the box: major U.S. cities and counties adopt fair hiring policies to remove unfair barriers to employment of people with criminal records. Resource Guide, 2014. Available at http://www nelp.org/publication/ban-the-box-fair-chance-hiring-state-and-local-guide/. Accessed September 24, 2014.
17. Survey of State criminal history information systems, 2012. Bureau of Justice statistics. Washington, DC: U.S. Department of Justice; 2014.
18. Todman L, Taylor JS, McDowell T, et al; Adler School of Professional Psychology, Institute on Social Exclusion. U.S. Equal Employment Opportunity Commission policy guidance: a mental health impact assessment. 2013.
19. Adler School of Professional Psychology, Institute on Social Exclusion. Arrest records as barriers to employment: Institute on Social Exclusion hosts Midwest Summit. Intersections 2013;8(1).
20. Wilkinson R, Marmot M (Eds). Social determinants of health: the solid facts. 2nd ed. Geneva: World Health Organization; 2003.
21. Katz M. Living wage and health. Testimony before the San Francisco Board of Supervisors, March 1, 2006.
22. Bhatia R, Katz M. Estimation of health benefits from a local living wage ordinance. Am J Publ Health 2011; 91(9):1398-1402. PMID: 11527770.

Steven M. Teutsch, MD, is an independent consultant, Adjunct Professor at the Fielding School of Public Health, UCLA, and Senior Fellow, Schaeffer Center, University of Southern California. Until 2014, he was the Chief Science Officer, Los Angeles County Public Health Department, where he continued his work on evidence-based public health and policy. Prior to that, he had been in the Outcomes Research and Management program at Merck. Before joining Merck, he was Director of the Division of Prevention Research and Analytic Methods at the Centers for Disease Control and Prevention. He has served as a member of the U.S. Preventive Services Task Force and chaired the U.S. Department of Health and Human Services Secretary's Advisory Committee on Genetics, Health, and Society. Dr. Teutsch is certified by the American Board of Internal Medicine and the American Board of Preventive Medicine and is a Fellow of the American College of Physicians and American College of Preventive Medicine.

Katherine M. Butler, MPH, is a health impact assessment analyst with the Health Impact Evaluation Center at the Los Angeles County Department of Public Health, where she designs and leads health assessment projects to examine the benefits and risks of policy and program decisions. Her experience includes collaborating with private and public organizations, including criminal justice, education, natural resources and energy, and housing. Ms. Butler previously worked for a private environmental health consulting firm, where she managed a team of health professionals for a variety of health assessment and epidemiology projects in the United States and abroad. In 2014, she served as President of the Southern California Society for Risk Analysis.

Paul Simon, MD, MPH, is the Director of the Division of Chronic Disease and Injury Prevention at the Los Angeles County Department of Public Health and an Adjunct Professor in the Department of Epidemiology at the UCLA School of Public Health. He oversees the Tobacco Control and Prevention Program, Nutrition and Physical Activity Program, Cardiovascular and School Health Program, Policies for Livable Active Communities and Environments (PLACE) Program, Injury and Violence Prevention Program, and Office of Senior Health. Dr. Simon also oversees a First 5 LA-funded early childhood obesity prevention project and a Centers for Disease Control and Prevention-funded Community Transformation Grant addressing obesity, tobacco use, and chronic disease prevention. Dr. Simon is board certified in pediatrics and preventive medicine.

Jonathan E. Fielding, MD, MPH, MBA, MA, has contributed to the field of public health for more than 40 years and has served in a variety of leadership positions. He led the public health activities for Los Angeles County as Director of Public Health and Health Officer for 16 years. He also served as Commissioner of the First 5 LA Commission, which provides over $100 million in funding annually for programs to improve the health and development of children aged 5 and younger. In September 2014, Dr. Fielding retired from county government service and is currently a Distinguished Professor of Public Health and Pediatrics at the University of California, Los Angeles, where he has been a tenured faculty member since 1979.

Behavioral and Social Science Aspects of the U.S. National HIV/AIDS Strategy (NHAS): 2010 to 2015 and Beyond

David R. Holtgrave, Cathy Maulsby, Chris Adkins, and Laura W. Fuentes

Abstract

In July 2010, President Barack Obama released the first ever comprehensive HIV/AIDS strategy for the United States. The U.S. National HIV/AIDS Strategy (NHAS) contained goals to be achieved by 2015 in four major areas: (1) reduction of HIV incidence; (2) improved linkage to HIV care, treatment, and housing services; (3) reduction of health disparities; and (4) enhanced service coordination. Here, we describe the contributions that the behavioral and social sciences have made in the past to these four broad domains, and we describe how the behavioral and social sciences can help monitor and achieve the 2015 NHAS goals, as well as help to plan for the next iteration of the NHAS for 2016 and beyond. It is clear from this review and analysis that behavioral and social science contributions are essential to achieving an AIDS-free generation in the near term.

Introduction

The human immunodeficiency virus (HIV) causes acquired immunodeficiency syndrome (AIDS) and has cumulatively infected over 75 million persons globally and 1.8 million persons in the United States since the late 1970s.[1-3] These infections have resulted in more than 39 million deaths worldwide and over 650,000 deaths in the United States alone.[2,3] HIV has been a part of the public discourse about health for over 30 years, and discussions about human sexual behavior, substance use, and social determinants of health have been forever altered by this epidemic.[4-6]

Despite the widespread impact of the epidemic, and the U.S. investment of hundreds of billions of dollars in addressing HIV/AIDS,[2,7] it was not until 2010 that the United States had a truly comprehensive National HIV/AIDS Strategy (NHAS).[8] Released in July 2010 by President Barack Obama, the NHAS sought to reduce the spread of HIV, help persons living with HIV obtain high quality medical care, decrease HIV/AIDS health disparities, and improve coordination of HIV-related activities in the public and private sectors.

In this chapter, we briefly describe the goals of the NHAS (which runs from 2010 through 2015) and then highlight the roles that the behavioral and social sciences can play in making the NHAS a success. To do so, we highlight key advances in behavioral epidemiology, prevention intervention development, social determinants and structural intervention development, and behavioral factors

in care and treatment. Next, we discuss how emerging biomedical advances have changed the landscape and provided new opportunities for behavioral and social science contributions to HIV prevention and care in synergy with biomedical breakthroughs. We then discuss key economic analyses related to HIV prevention and care and describe how such analyses can help inform the setting of the next generation of national strategy goals for 2016 and beyond. We conclude by highlighting some key implications for practice and future research.

Overview of the National HIV/AIDS Strategy

The NHAS has four major sections with goals stated for the time period 2010 through 2015;[8] all NHAS goals are quoted in Table 1. The first section calls for a reduction in new HIV infections, a decrease in the annual HIV transmission rate (defined as HIV incidence in a given year divided by HIV prevalence in that year, then multiplied by 100 for ease of exposition), and an increase in the level of awareness of seropositivity among persons living with HIV.

Table 1. National HIV/AIDS Strategy Goals: 2010 through 2015

Reducing New HIV Infections:

- Lower the annual number of new infections by 25 percent.
- Reduce the HIV transmission rate, which is a measure of annual transmissions in relation to the number of persons living with HIV, by 30 percent.
- Increase from 79 percent to 90 percent the percentage of people living with HIV who know their serostatus.

Increasing Access to Care and Improving Health Outcomes for People Living with HIV:

- Increase the proportion of newly diagnosed patients linked to clinical care within 3 months of their HIV diagnosis from 65 percent to 85 percent.
- Increase the proportion of Ryan White HIV/AIDS Program clients who are in continuous care (at least two visits for routine HIV medical care in 12 months at least 3 months apart) from 73 percent to 80 percent.
- Increase the percentage of Ryan White HIV/AIDS Program clients with permanent housing from 82 percent to 86 percent. (This serves as a measurable proxy of our efforts to expand access to HUD and other housing supports to all needy people living with HIV.)

Reducing HIV-Related Disparities and Health Inequities:

- Increase the proportion of HIV-diagnosed gay and bisexual men with undetectable viral load by 20 percent.
- Increase the proportion of HIV-diagnosed blacks with undetectable viral load by 20 percent.
- Increase the proportion of HIV-diagnosed Latinos with undetectable viral load by 20 percent.

Achieving a More Coordinated National Response to the HIV Epidemic:

- Increase the coordination of HIV programs across the Federal Government and between Federal agencies and State, territorial, local, and tribal governments.
- Develop improved mechanisms to monitor and report on progress toward achieving national goals.

Note: All headings and goals in this table are presented as they are stated in the Federal National HIV/AIDS Strategy issued in July 2010.

AIDS = acquired immunodeficiency syndrome; HIV = human immunodeficiency virus; HUD = Housing and Urban Development

The second section of Table 1 highlights the need for increased linkage to and retention in care and treatment among persons living with HIV (PLWH) and improvements in the provision of stable housing for homeless or unstably-housed PLWH. Figure 1 displays the remarkable drop-off in the number of persons living with HIV who are diagnosed, linked to care, retained in care, prescribed antiretroviral therapy, and achieve viral suppression.[9,10] Clearly, improvements in the provision of HIV care in the United States are urgently needed. The goals of this section largely are stated in terms of the population of clients receiving HIV care under the Ryan White Care Act programs (the Federal provider of last resort in the United States);[11,12] this is subtly but importantly different than stating goals for the entire population of persons living with HIV.

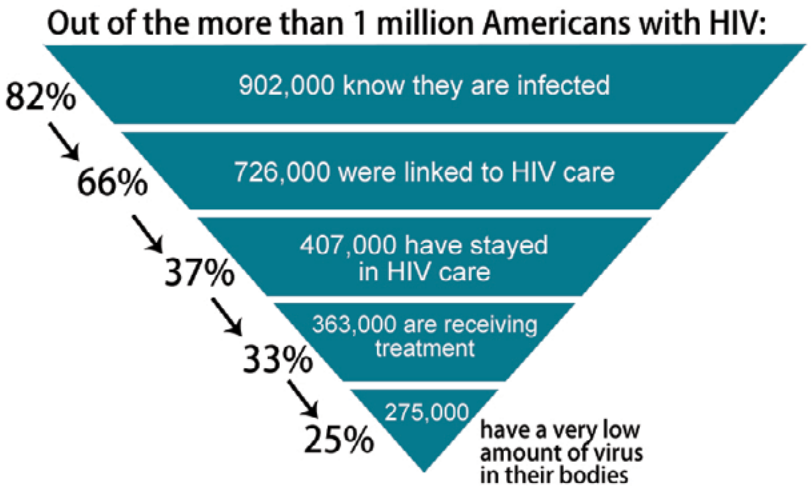


Figure 1: HIV Care Continuum for the United States

Source: Centers for Disease Control and Prevention; 2013. Available at http://www.cdc.gov/nchhstp/newsroom/docs/HIVFactSheets/TodaysEpidemic-508.pdf

The third section of the NHAS (see Table 1) discusses key health disparities in which communities are most heavily impacted by HIV. According to the Centers for Disease Control and Prevention (CDC), the 10 communities most disproportionately impacted by new HIV infections are white, black, and Hispanic men who have sex with men (MSM), black heterosexual women and men, white and Hispanic heterosexual women, black males and black females who inject drugs, and Hispanic heterosexual men).[10] Further, this section of the NHAS discusses in detail root causes of the HIV epidemic such as HIV-related stigma, discrimination, homophobia, poverty, and social inequality. While these social drivers of the epidemic are discussed in the NHAS, the goals of this section are all framed in terms of increases in viral suppression in three heavily and disproportionately impacted communities (see Table 1).

The fourth section of the NHAS notes the importance of the coordination of all HIV/AIDS related services. These goals are phrased as a pair of process goals rather than disease-specific outcome or impact goals (see Table 1).

Throughout the NHAS, it is emphasized that no one of these four sections can stand alone, but rather they must operate in a synergistic way to truly impact the course of the epidemic in the United States. This is especially important because the findings of the highly influential National Institutes of Health (NIH)-sponsored trial "HPTN 052" demonstrated that HIV treatment can also serve critical prevention purposes by suppressing viral load to a point that HIV transmission to HIV seronegative partners is indeed a very rare occurrence.[13,14] Hence, the lines between prevention and treatment have been heavily blurred in recent years, and one cannot truly discuss one section of the NHAS without the other.

Behavioral and Social Science Achievements Relevant to Four Sections of the NHAS

Modes of Transmission and Behavioral Epidemiology

As students of the behavioral and social sciences are certainly aware, these disciplines have seen research findings unfold over the past decades of direct relevance to the four sections of the NHAS. Here we describe a selected set of highlights of these findings (noting again that all parts of the NHAS are truly synergistic); in the following sections of this chapter, we describe future challenges for the behavioral and social sciences to best achieve the goals of the NHAS through 2015 and help establish goals for 2016 and beyond.

HIV is transmitted when there is unprotected risk behavior (generally sexual or drug injection risk behavior) in a serodiscordant partnership in which the partner living with HIV has unsuppressed virus.[15] Of course, this "partnership" maybe as brief as one sexual or drug-sharing encounter, or it may involve many encounters over a long period of time. This simple sentence has important implications; it means that if there are no transmission-relevant risk behaviors or if there is well-suppressed virus (or both), then transmission is indeed a very rare event (with the probability of transmission approaching zero). This implication provides two critical pathways for disrupting HIV transmission. First, transmission-relevant risk behaviors can be modified or eliminated, and second, viral levels in persons living with HIV can be suppressed.[15-17] The behavioral and social sciences have made important contributions to both main prevention pathways; we review each here in turn.

Regarding the identification and modification of HIV transmission-relevant risk behaviors, there is a rich history of success in the behavioral and social sciences. Behavioral and social scientists contributed in fundamental ways to the earliest cohort studies that helped elucidate the patterns of HIV transmission, the behaviors that transmitted the virus, and the consequences of transmission.[18-20]

Prevention Intervention Development

Almost as soon as the modes of transmission were identified and the disproportionate impact on some communities was known, behavioral and social scientists began to develop interventions that sought to provide information about HIV, shape attitudes toward protecting one's self and one's partners from becoming infected, and modify behaviors (such as increasing condom use

and elevating use of sterile injection equipment among persons who inject drugs). The earliest of these interventions were developed in the late 1980s and published in the early 1990s.[21,22] In the 1990s, the National Institute of Mental Health (NIMH) made a substantial investment in the development of individual-level and small-group-level interventions for a variety of heavily impacted communities, such as gay and bisexual men, persons who inject drugs, and heterosexual women (with communities of color being very disproportionately affected). Because intervention trials with HIV incidence as an outcome measure would be very expensive, NIH chose during this era to fund projects with psychometrically rigorous behavior change measures and non-HIV sexually transmitted infections (STIs) as outcome metrics;[23] this allowed more numerous, smaller scale studies to be done on a variety of intervention types for several communities at risk, rather than a very few large trials with HIV incidence as the outcome. This line of research led to the development of dozens of interventions capable of changing HIV-relevant risk behavior and a number with significant STI-change outcomes.[23-25]

Although the interventions described above varied in their content, of course, the prototypic intervention was a small group program in which peer opinion leaders would both participate in and lead the groups, and topics discussed included education on risk triggers and self-management, condom use, and peer support.[26-30] These types of interventions were highly successful at changing HIV risk behaviors (e.g., by improving correct and consistent condom use).[24,31] While also heavily reliant on peer opinion leaders, some interventions operated more at the community level, attempting to inform, motivate, and empower entire communities at the same time.[21,32]

CDC was also interested in developing behavioral interventions during this time period, most especially as behavior change counseling might be coupled with HIV testing. In 1998, CDC published the findings from Project RESPECT, which sought to determine the effectiveness of providing counseling in conjunction with HIV testing as a means to modifying HIV-related risk behaviors and reducing non-HIV STIs (as a proxy for reducing HIV incidence).[33] Project RESPECT found that pre- and post-test counseling could change HIV-related risk behaviors and incidence of non-HIV STIs among HIV seronegative persons attending sexually transmitted disease clinics in several U.S. cities.[33] Marrazzo and colleagues discuss the complex array of ensuing studies that examined the effects of counseling as it relates to HIV testing [25]

These are but a few examples of the types of behavioral interventions developed for HIV prevention. Other types of successful interventions have included couples counseling[34] and community empowerment interventions.[35] Also, over time, there was more attention paid to developing HIV prevention programs that served PLWH to avoid transmission to sex and drug injection partners, as well as to emphasize benefits from strong linkages to HIV care and ancillary services.[25,36] This reflects in part CDC's focus on HIV treatment as prevention, as well as their focus on prevention efforts to disrupt transmission from PLWH. In fact, CDC has labeled this approach as "high impact" HIV prevention.[37]

A crucial development in this line of behavioral intervention research was the development of CDC's Research Synthesis Project, which saw the methodological review of numerous studies, and, as of 2014, the identification of 51 "best" evidence and 33 "good" evidence interventions (where the adjectives in quotations are defined according to the level of empirical evidence available in the

literature for each intervention).[31] CDC then selected some interventions for further dissemination to the field and implemented the "DEBI" (for "diffusing evidence-based interventions") project. The DEBI project saw the packaging of key evidence based interventions into manual and training course form, so that the capacity of interested service delivery organizations could be built to deliver these interventions [37,38]

Social Determinants and Structural Interventions

Another important era in the development of behavioral and social interventions was the expanding recognition in the late 1990s and early 2000s[39,40] that social determinants of health (such as living in wartime, impoverished and/or socially inequitable situations) also were related to heightened HIV prevalence. CDC and HUD (the U.S. Department of Housing and Urban Development) were among the first Federal agencies to attempt to construct structural interventions to address these social determinants by fielding a major study of housing provision to homeless and unstably housed persons as an HIV prevention intervention.[41] In "as-treated" analyses, the study found the provision of housing to actually impact the level of viral load among clients living with HIV.[42] Further, the cost per quality adjusted life year (cost per QALY) saved by this intervention was $62,493 in 2005 year dollars, indicating a reasonably good level of cost-effectiveness for the provision of this resource-intensive program.[43] Most recently, CDC has issued a report examining the social determinants of health for PLWH in a number of jurisdictions in the United States.[44]

Another critical example of an evidence-based structural intervention is sterile syringe exchange programs in which persons who inject drugs (or someone acting for them) can exchange syringes that may contain HIV (and other infectious diseases) for sterile injection equipment.[45,46] Social scientists were involved in studying and establishing the evidence base for syringe exchange programs, most notably in studies done at Yale and the University of California, San Francisco in the 1990s with the collaboration of CDC.[47-52] In summary, they found that syringe exchange programs did lead to a decrease in HIV transmission among program participants and their partners and did so in a highly cost-effective manner (the economic argument of which was recently confirmed by Nyugen et al.).[53]

Behavioral Factors in Care and Treatment

The behavioral sciences can also play an important role in HIV care and treatment.[25] Relatively recent ground-breaking research suggested that widespread early initiation of antiretroviral therapy (ART) could substantially reduce the transmission of HIV.[13,54,55] This realization changed the landscape of HIV prevention and led to an increase in focus on access to HIV care and treatment services as a tool for reducing HIV transmission. There are many behavioral factors related to successfully navigating the treatment cascade in order to realize the optimal treatment and prevention effects of ART. Individuals must be tested for HIV and receive the result of their test, have an initial HIV care appointment, engage in ongoing HIV care, be prescribed and adhere to ART, and finally, achieve and maintain viral suppression.[56] Movement through these steps is not always linear, however, and many individuals cycle in and out of care.[57,58]

Behavioral and social science interventions make key contributions at each of these steps. CDC's National HIV Behavioral Surveillance system has been critical for understanding the level, locale,

and determinants of undiagnosed HIV seropositivity in a number of jurisdictions in the United States.[59,60] In order to promote the potential of HIV treatment as prevention, CDC has emphasized not only linkage and retention in care, but also HIV testing so that one can learn their HIV serostatus in the first place. This has largely taken the form of CDC advocating for universal, clinic-based HIV testing for 13 to 64-year-old persons, with persons in the highest risk communities testing at least once per year.[61] It is recognized however that not all testing will occur in clinical settings, as some members of some communities currently have limited or no access to clinical care. Therefore, community-based HIV testing programs (such as HIV testing provided in concert with syringe exchange programs, or venues such as "house balls" in transgender communities) can serve a critical function in increasing awareness of serostatus.[62] Especially in community-based testing, social networking approaches can be key in efforts to reach not just one person at risk of infection (or who is undiagnosed and living with HIV), but the entire network of sexual partners and/or drug injection partners of the index person living with HIV.[63]

Once a PLWH is aware of his or her status, he or she must attend HIV medical appointments. Behavioral intervention strategies such as motivational or strengths-based counseling, information and education about HIV care, peer navigation, accompanying clients to medical appointments, and appointment coordination have been found to increase linkage to and retention in care among PLWH who are newly diagnosed or out of care.[25,64,65]

Adherence to HIV medication is a health behavior that is vital to the success of HIV treatment for the health of the infected person and for prevention of onward transmission. A handful of evidence-based ART adherence interventions have successfully increased ART adherence among a variety of populations, including clinic patients, drug users, and men who have sex with men. These interventions share common components, such as skills-building around medical adherence, addressing barriers to adherence and problem solving, communicating with providers, and building social support.[66]

Behavioral scientists are building the evidence base for successful interventions that facilitate viral suppression, researching the process of translating these interventions into practice in new locations and with various populations, and testing new innovative intervention models. In addition, behavioral scientists are conducting important research to better understand the factors associated with successfully (or unsuccessfully) engaging in the steps along the spectrum of HIV care and treatment. For example, work by Lo and colleagues identified the critical role that ancillary services—such as receipt of mental health care, drug assistance, food/nutrition, and transportation—play in utilization and retention in care.[67] Keruly and colleagues found that factors such as age, race/ethnicity, sex, CD4 cell count, insurance status, and number of missed clinical visits were predictive of highly active antiretroviral therapy (HAART) use.[68] This line of research helps to identify the populations at greatest risk for not fully engaging in HIV care and treatment and suggests potential areas where interventions and policy could intervene.

The behavioral and social sciences are also directly relevant to the NHAS goal of reducing health disparities. Traditional methods of infectious disease epidemiology focus on patterns of HIV disease. These epidemiological investigations need the behavioral and social science information to make them optimally useful for the design of treatment and prevention strategies. CDC has made

a major priority of carrying out behavioral surveillance studies to complement more traditional epidemiologic investigations. Their National HIV Behavioral Surveillance System (NHBSS)[69,70] examines risk behaviors and epidemiologic contexts among men who have sex with men (defined in terms of behavior rather than in terms of gay or bisexual identity), persons who inject drugs, and heterosexual women at heightened risk of HIV infection (one population per year on a rolling 3-year basis).[69,70] This use of behavioral methods in surveillance studies has led to critical insights. For example, social scientists in Baltimore have used surveillance studies to better understand disparities among vulnerable populations, gather data on risk behaviors, and directly monitor and plan for HIV prevention.[59,60] Maulsby and colleagues used surveillance data to identify disparities in HIV testing-related behaviors and predictors of undiagnosed seropositivity among men who have sex with both men and women (MSMW).[59,60] They found that a range of partnership characteristics contributed to undiagnosed seropositivity among MSMW, highlighting the need for programs that address the unique partnership dynamics of this population. Behavioral scientists at Johns Hopkins have worked with administrators at the Maryland Department of Health and Mental Hygiene to determine alignment between surveillance indicators and local NHAS HIV prevention goals and to establish procedures to facilitate the future use of surveillance data in local HIV program planning and policymaking.[71]

In summary, the behavioral and social sciences have made critical contributions of direct relevance to achieving the goals of the NHAS, and we acknowledge that our brief review here portrays only some of the impact findings in this arena. Now, we turn to some emerging issues in the application of the behavioral and social sciences to the NHAS, and we discuss economic and implementation science aspects of these interventions, before describing how the behavioral and social sciences can help to inform future goals that may be part of the next NHAS.

Emerging Issues in Behavioral and Social Aspects of HIV

NIH HPTN trial 052 provided strong evidence that HIV treatment that successfully suppressed viral load actually served to substantially decrease the transmission of HIV to seronegative partners.[13,14] Because of this landmark finding, much subsequent attention has been paid to HIV "treatment as prevention." Initially, such biomedical findings led to coining of the term "combination prevention" in which clients were to receive the very best evidence-based behavioral and biomedical interventions for HIV prevention and care. However, over time, some have emphasized treatment as prevention to the exclusion of behavioral interventions (a phenomenon we elsewhere called "substitution prevention").[16,72] We have argued that what is truly needed is "synergistic prevention" in which biomedical and behavioral interventions are offered in a complementary manner so that clients receive the best mix of behavioral and biomedical interventions in a manner that truly results in a decrease in HIV incidence, with the whole being more than the sum of the parts.[16,72]

Other recent studies have shown that HIV drugs taken by HIV seronegative persons as "pre-exposure prophylaxis" (or PrEP) can substantially reduce the probability of someone contracting HIV from a person living with HIV.[25] PrEP has been controversial in that on the one hand the prevention potential of the intervention is quite strong, but on the other hand, PrEP effectiveness is very dependent on high levels of adherence, can be highly expensive, and may lead to the decrease

of other modes of prevention.[73] CDC, WHO (the World Health Organization), and other bodies have now recommended PrEP in relatively limited but important circumstances (illustrations of such circumstances include persons at heightened risk for HIV infection due to >2% background HIV seroprevalence, a recent sexually transmitted disease diagnosis, sharing of drug injection equipment, or needing to frequently utilize post-exposure prophylaxis).[25,74-76]

Economic and Implementation Research Issues

As noted previously in this chapter, a number of the behavioral and structural HIV prevention interventions described here have been subjected to economic evaluation and found to be cost saving or cost effective when compared with other interventions in medicine and public health.[43,77,78] One possible exception is that of PrEP, an intervention for which the cost-effectiveness analytic results vary widely with the variation driven largely by the HIV seroprevalence of the community under study.[79,80]

Not all jurisdictions have the resources necessary to deliver all useful, cost-effective interventions to scale to all who may benefit from them. Therefore, several recent studies have focused on determining the most impactful (in terms of HIV prevention) use of available resources. For example, two recent studies have examined how best to avert the maximum number of HIV infections given limited resources for the Baltimore-Towson metropolitan statistical area (MSA) and for the State of Iowa.[81,82] A common theme in these resource allocation studies is that because HIV prevention funds are so limited, very hard choices about maximizing impact with insufficient resources must be made.

Another type of economic analysis attempts to estimate the overall unmet HIV prevention, care, and/or housing needs among a particular population. For example, we recently conducted a study to determine the overall unmet HIV prevention, care, and housing service needs among black men who have sex with men in the United States.[83]

Interestingly, the NHAS makes no mention of the costs that would be incurred by complete implementation of the strategy or its return on investment. We made estimates of the costs and return on investment in 2010 and updated these estimates again in 2012.[84,85] Such analyses provide quantitative estimates of the level of investment that would be required to meet the NHAS 2015 goals and what might be expected at current resource levels. We found that while the programmatic costs of meeting all goals of the NHAS would be substantial, such an investment would not only change the course of the HIV epidemic in terms of reductions in HIV incidence and transmission rates and improvements in access to care, it would do so in a manner that would be considered cost effective by traditional economic evaluation standards.

Goals for 2016 and Beyond

Our previous analyses of the costs and consequences of the NHAS concluded that it was still possible to meet the goals of the NHAS if a substantial investment in new and/or redirected public

and/or private sector funds had been made. It would have been necessary to make this investment very rapidly,[84,85] since significant resources would be needed in fiscal years 2013 and 2014 (this is because investment in one fiscal year is unlikely to impact program delivery until later that year or the next year, with disease-specific impacts 1 or more years later). Unfortunately the FY2013 and FY2014 Federal appropriations processes did not yield the level of increased or redirected resources necessary to make the major move toward scale-up of services needed to achieve the NHAS 2015 goals.[86] Therefore, we recently published an article arguing that the full achievement of the NHAS goals is probably now out of sight at this point due to lack of sufficient investment (though data on some key indicators lags).[86] We argued further that even if the 2015 goals are seemingly out of reach at this point, earlier mathematical modeling could be used to productively inform possible NHAS goals for 2016 through 2020 (assuming there would be another NHAS for future years). These suggested goals are listed in Table 2.

Table 2. Proposed 2020 goals for a new National HIV/AIDS Strategy (NHAS) for the United States

Reducing New HIV Infections:

- Lower the annual number of new HIV infections by at least 45 percent (relative to baseline year 2010).
- Reduce the HIV transmission rate, which is a measure of annual transmissions in relation to the number of people living with HIV (PLWH), by at least 50 percent (relative to baseline year 2010).
- Increase to at least 90 percent the percentage of PLWH who know their serostatus (with an emphasis on identifying seropositivity as soon as possible after HIV infection).
- Further reduce the already low number of diagnosed PLWH who engage in unprotected, serodiscordant, transmission-relevant risk behavior by at least 50% (relative to baseline year 2010).

Measurement: All four quantitative constructs above are now estimated and published by CDC.

Increasing Access to Care and Improving Health Outcomes for People Living with HIV:

- Ensure that at least 85% of newly diagnosed PLWH are linked to clinical care within 3 months of their HIV diagnosis, and that at least 85% of all diagnosed PLWH are retained in care.
- Ensure that at least 81% of clients receiving HIV care achieve and maintain viral suppression.
- Ensure that at least 90% of PLWH in need of stable housing services receive and retain such services.

Measurement: The first and second goals in this section can be measured by existing CDC and HRSA systems. In the 2010 NHAS, the housing goal referenced Ryan White clients; we propose a broader measurement strategy reflecting more PLWH (such as an expansion of CDC's Medical Monitoring Project).

Reducing HIV-Related Disparities and Health Inequities:

- Ensure that all goals listed in the care section and the seropositivity awareness goal are achieved for PLWH of both sexes and for all racial/ethnic groups, sexual minorities, heightened risk groups (such as injection drug users), and age groups

- Among the nine subpopulations defined by CDC as totaling more than 85% of the HIV incidence in the United States in 2010, ensure that no subpopulation has more than 2,550 new HIV infections per year (and that incidence is level or decreasing for all subpopulations in the Nation in all years).
- Develop and annually report on measures designed to assess HIV-related stigma experienced by PLWH (especially, to gauge how such experiences serve as barriers to entry into, or retention in, HIV care).

Measurement: The first goal in the disparities section can be measured by seemingly feasible subgroup analyses in CDC and HRSA systems. The second goal can be measured from subgroup analyses conducted and published by CDC. The third disparities goal is a directly observable, nominal-scale process goal.

Achieving a More Coordinated National Response to the HIV Epidemic:

- Continue efforts to constantly refine and annually report on the coordination of HIV programs across the Federal Government and between Federal agencies and State, territorial, local, and tribal governments.
- Continually refine and annually report on metrics necessary to monitor and react to progress toward achieving all national goals.

Measurement: The goals in this section are directly-observable, nominal-scale process goals. The second coordination goal builds on existing efforts by the Department of Health and Human Services to have all HIV federally funded service delivery programs monitor a small, core set of NHAS-relevant core indicators.

Source: Reprinted from Holtgrave DR. Development of year 2020 goals for the National HIV/AIDS Strategy for the United States. AIDS Behav 2014;18(4):638-43. With kind permission from Springer Science and Business Media.

Notes: All four section headings are quoted exactly as presented in the Federal National HIV/AIDS Strategy issued in July 2010 (http://www.whitehouse.gov/sites/default/files/uploads/NHAS.pdf); the goals in each section are newly proposed or modified here.

CDC = Centers for Disease Control and Prevention; HIV = human immunodeficiency virus; HRSA = Health Resources and Services Administration; NHAS = National HIV AIDS Strategy; PLWH = people living with HIV.

What is immediately noticeable is that we do not attempt to "kick the can" down the road and simply repeat the 2015 goals. Rather, we crafted 2020 goals that would seem to be epidemiologically attainable with reasonable investment in needed services; in other words, these goals are feasible if the appropriate level of investment is made. These draft, possible goals are meant to inform discussion, and we believe behavioral and social scientists have a great deal to contribute to the crafting of the next generation of the NHAS.

Conclusions

The behavioral and social sciences have made significant contributions to the efforts to prevent and treat HIV in the United States. Now is a time of change in that new biomedical findings have caused a sea change in how HIV prevention and treatment are viewed in the United States; this challenges the behavioral and social sciences community to consider new target behaviors, modernize intervention content for relevance, further expand to include the addressing of social determinants

of HIV, and fully embrace methods of implementation research that include consideration of the costs, consequences, and setting of national HIV/AIDS goals. Despite the unfolding history of the HIV epidemic in the United States, the behavioral and social sciences community will have further key roles to play, so that one day soon we might actually see the elimination of HIV in the United States.[87-89]

Implications for Practice

Woven throughout the sections above, we have highlighted a number of ways in which the behavioral and social sciences have and can still impact HIV prevention and care practice. This is very well illustrated in the recent International AIDS Society-USA (IAS-USA) publication on combination HIV prevention in an era of treatment advances.[25] This IAS-USA panel comprised a highly interdisciplinary group of behavioral and social scientists, physicians, and epidemiologists. They sought to determine the optimal evidence-based combination of biomedical and behavioral interventions that clinicians could use to prevent HIV infection among their patients or onward transmission of HIV from their patients living with HIV. The complete recommendations can be found in the article online (free access at http://jama.jamanetwork.com/article.aspx?articleid=1889145); here we present some of the key behavioral and social elements of the recommendations:

- HIV risk screening and counseling that is evidence-based, effective, and meets client needs but is not so intrusive in the clinical practice as to be a barrier to routine HIV screening.
- Brief evidence-based risk reduction interventions suitable for use in the clinical setting for HIV seronegative clients at ongoing risk for HIV infection (including a complementary mix of behavioral risk reduction services and access to pre-exposure prophylaxis).
- Evidence-based HIV prevention services for patients living with HIV infection.
- Evidence-based supportive services to enhance linkage to, retention in, and adherence to HIV care and treatment.

A theme across these clinical practice recommendations for prevention is the synergistic combination of evidence-based behavioral and biomedical interventions. Of course these recommendations were focused on what clinicians can do, but many of the recommendations are salient at the population level as well and could be utilized by health departments and community-based service providers. Still, some structural interventions noted above are much more likely to be implemented at the population level, and these would include housing services, stigma prevention/reduction campaigns and modernization of HIV-related laws, and syringe exchange services. Further, the economic analyses described above also provide actionable policy recommendations in terms of resources needed to address unmet needs and what the likely epidemiologic and fiscal consequences of such investments would be. In particular, such modeling analyses give us a glimpse into the future as to whether or not the goals of the NHAS will be realized at the current levels of investment.

Implications for Research

Also embedded in the section above are many clear implications for needed future research, a few of which we underscore here. NIH frequently updates a comprehensive agenda for behavioral

and social science research focused on HIV. It is beyond the scope of this section to replicate such a comprehensive research agenda; instead, we highlight here just a few areas in need of rapid attention.

- Additional testing of prevention "bundles" of behavioral and biomedical interventions using state of the art implementation science methods (rather than testing one intervention at a time).
- Rapid turn-around analyses to inform the selection of input parameters for policy modeling exercises (for example, there is scant literature available on the costs and effects of partner notification services, but such services are important to consider in resource allocation analyses; additional, rapid, empirical guidance on such parameters would be directly and immediately useful).
- Rapid turn-around studies to bolster evidence-based interventions to improve HIV care linkage and retention (while linkage and retention in care are absolutely critical, the number of intervention studies in this arena is still relatively small).
- Methods to more quickly estimate key metrics needed to monitor NHAS progress (e.g., while the goals of the NHAS run through 2015, the most recent HIV incidence estimate for the United States from CDC is for the year 2010).
- Additional mathematical analyses to inform setting of NHAS 2020 goals (while we have published some such analyses, the mathematical modeling literature is always well served if there are multiple modeling exercises done by independent teams to determine if the overarching policy recommendations from different analyses are similar).
- Rapid turn-around mixed methods studies to help us understand what prevention and care services are needed to help stem the trend of new HIV infections among the populations most disproportionately impacted (with a special urgency on studies related to young black gay men).

A theme that runs throughout these research recommendations is that they tend to focus on matters of community need, scale up, implementation evaluation, and health services research; this is because we already have many evidence-based HIV prevention and care tools at our disposal, yet we may not always be employing them properly or with the right intensity. These are empirical matters that can be informed by rigorous yet urgent study.

Acknowledgments

The opinions presented in this chapter are those of the authors and may not necessarily represent the position of the Agency for Healthcare Research and Quality, the National Institutes of Health, or the U.S. Department of Health and Human Services.

Authors' Affiliations

David R. Holtgrave, PhD, Professor and Department Chair; Cathy Maulsby, PhD, Chris Adkins, MSW, Laura W. Fuentes, MM, Department of Health, Behavior & Society; Johns Hopkins Bloomberg School of Public Health.

Address correspondence to: David R. Holtgrave, PhD, Department of Health, Behavior & Society, Johns Hopkins Bloomberg School of Public Health, 624 N Broadway, Baltimore, MD 21205; email david.r.holtgrave@jhu.edu.

References

1. Blattner W, Gallo RC, Temin HM. HIV causes AIDS. Science 1988;241(4865):515-6.
2. The HIV epidemic in the United States: April 2014, fact sheet. Menlo Park, CA: Kaiser Family Foundation; April 7, 2014. Available at http://kff.org/hivaids/fact-sheet/the-hivaids-epidemic-in-the-united-states/. Accessed September 18, 2014.
3. Global fact sheet, 2014. Geneva, Switzerland: United Nations, UNAIDS; November 2012. Available at http://www.unaids.org/en/media/unaids/contentassets/documents/factsheet/2014/20140716_FactSheet_en.pdf. Accessed September 30, 2014.
4. Shilts R. And the band played on: politics, people and the AIDS epidemic. New York: St. Martin's Griffin; 1987.
5. Stine G (Ed). AIDS update 2013 (22 ed.). New York: McGraw-Hill/Dushkin; 2012.
6. A timeline of AIDS. Washington, DC: US Department of Health and Human Services; 2011. Available from http://aids.gov/hiv-aids-basics/hiv-aids-101/aids-timeline/. Accessed September 18, 2014.
7. Johnson JA. AIDS funding for Federal Government programs FY1981-FY2009. Washington, DC: Congressional Research Service; Apriil 2008. Available at http://www.law.umaryland.edu/marshall/crsreports/crsdocuments/96-293_SPR.pdf. Accessed September 18, 2014.
8. National HIV/AIDS strategy for the United States. Washington, DC: The White House; July 13, 2010. Available at http://www.whitehouse.gov/sites/default/files/uploads/NHAS.pdf. Accessed September 18, 2014.
9. HIV in the United States: the stages of care. CDC Factsheet. Atlanta, GA: Centers for Disease Control and Prevention; December 2013. Available at www.cdc.gov/nchhstp/newsroom/docs/HIV-Stages-of-Care-Factsheet-508.pdf. Accessed September 20, 2014.
10. Today's HIV/AIDS epidemic. CDC Factsheet. Atlanta GA: Centers for Disease Control and Prevention; December 2013. Available at http://www.cdc.gov/nchhstp/newsroom/docs/HIVFactSheets/TodaysEpidemic-508.pdf. Accessed September 20, 2014.
11. Remarks by the President at signing of the Ryan White HIV/AIDS Treatment Extension Act of 2009. Washington, DC: The White House; 2009. Available at http://www.whitehouse.gov/the-press-office/remarks-president-signing-ryan-white-hivaids-treatment-extension-act-2009. Accessed September 20, 2014.
12. Ryan White Comprehensive AIDS Resources Emergency Act of 1990. Public Law 101-381, 104 Statute 576 C.F.R. Washington, DC: U.S. Congress; 1990. Available at http://www.gpo.gov/fdsys/pkg/PLAW-111publ87/html/PLAW-111publ87 htm. Accessed September 20, 2014.
13. Cohen MS, Chen YQ, McCauley M, et al. Prevention of HIV-1 infection with early antiretroviral therapy. N Engl J Med 2011;365(6):493-505.
14. Cohen MS, McCauley M, Gamble TR. HIV treatment as prevention and HPTN 052. Curr Opin HIV AIDS 2012;7(2):99-105.
15. Hall HI, Holtgrave DR, Tang T, et al. HIV transmission in the United States: considerations of viral load, risk behavior, and health disparities. AIDS Behav 2013;17(5):1632-6.
16. Hall HI, Holtgrave DR, Maulsby C. HIV transmission rates from persons living with HIV who are aware and unaware of their infection. AIDS 2012;26(7):893-6.
17. Holtgrave DR, Hall HI, Des Jarlais DC, et al. Estimating number of diagnosed persons living with HIV in the United States engaged in unprotected serodiscordant risk behavior with unsuppressed viral load. J Acquir Immune Defic Syndr 2014;65(3), e125-8.
18. Coates TJ, Stall RD, Catania JA, et al. Behavioral factors in the spread of HIV infection. AIDS 1988;2(Suppl 1):S239-46.
19. Ehrhardt AA. Trends in sexual behavior and the HIV pandemic. Am J Public Health 1992;82(11):1459-61.
20. Vlahov D, Anthony JC, Munoz A, et al. The ALIVE study, a longitudinal study of HIV-1 infection in intravenous drug users: description of methods and characteristics of participants. NIDA Res Monogr 1991;109:75-100.
21. Kelly JA, St Lawrence JS, Diaz YE, et al. HIV risk behavior reduction following intervention with key opinion leaders of population: an experimental analysis. Am J Public Health 1991;81(2):168-71.
22. Valdiserri, RO, Lyter DW, Leviton LC, et al. AIDS prevention in homosexual and bisexual men: results of a randomized trial evaluating two risk reduction interventions. AIDS 1989;3(1):21-6.
23. Interventions to prevent HIV risk behaviors. NIH Consensus Statement Vol 15, No 2; February 11-13,1997. Available at http://consensus nih.gov/1997/1997PreventHIVRisk104PDF.pdf. Accessed September 20, 2014.
24. HIV/AIDS prevention research synthesis project. Atlanta, GA: Centers for Disease Control and Prevention; 2014. Available at http://www.cdc.gov/hiv/dhap/prb/prs/index html. Accessed September 19, 2014.
25. Marrazzo J, Del Rio C, Holtgrave D et al. HIV prevention in clinical care settings: 2014 recommendations of the International Antiviral Society - USA Panel. JAMA 2014;312(4):390-409.

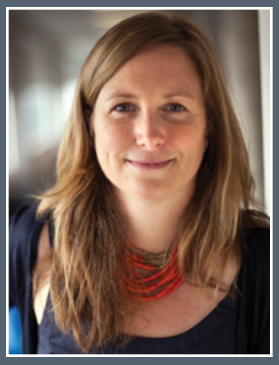
Available at http://jama.jamanetwork.com/article.aspx?articleid=1889145. Accessed September 24, 2014.
26. Kalichman SC, Sikkema KJ, Kelly JA, et al. Use of a brief behavioral skills intervention to prevent HIV infection among chronic mentally ill adults. Psychiatr Serv 1995;46(3):275-80.
27. Kelly JA, Murphy DA, Washington CD, et al. The effects of HIV/AIDS intervention groups for high-risk women in urban clinics. Am J Public Health 2994;84(12):1918-22.
28. Kelly JA, St Lawrence JS. The impact of community-based groups to help persons reduce HIV infection risk behaviours. AIDS Care 1990;2(1):25-36.
29. Kelly JA, St Lawrence JS, Betts R, et al. A skills-training group intervention model to assist persons in reducing risk behaviors for HIV infection. AIDS Educ Prev 1990;2(1):24-35.
30. Sikkema KJ, Winett RA, Lombard DN. Development and evaluation of an HIV-risk reduction program for female college-students. AIDS Educ Prev 1995;7(2):145-59.
31. Compendium of evidence-based interventions and best practicesfor HIV prevention. Atlanta, GA: Centers for Disease Control and Prevention; 2014. Available at http://www.cdc.gov/hiv/prevention/research/compendium/rr/complete html. Accessed September 22, 2014.
32. Kegeles SM, Hays RB, Coates TJ. The Mpowerment Project: a community-level HIV prevention intervention for young gay men. Am J Public Health 1996;86(8):1129-36.
33. Kamb ML, Fishbein M, Douglas JM Jr, et al. Efficacy of risk-reduction counseling to prevent human immunodeficiency virus and sexually transmitted diseases: a randomized controlled trial. Project RESPECT Study Group. JAMA 1998;280(13):1161-7.
34. Sullivan PS, Stephenson R, Grazter B, et al. Adaptation of the African couples HIV testing and counseling model for men who have sex with men in the United States: an application of the ADAPT-ITT framework. Springerplus 2014;3:249.
35. Kegeles SM, Rebchook G, Pollack L, et al. An intervention to help community-based organizations implement an evidence-based HIV prevention intervention: the Mpowerment Project technology exchange system. Am J Community Psychol 2012;49(1-2):182-98.
36. Kalichman SC. The other side of the healthy relationships intervention: mental health outcomes and correlates of sexual risk behavior change. AIDS Educ Prev 2005;17(1 Suppl A):66-75.
37. Effective interventions: what's new (High Impact HIV/AIDS Prevention Project (HIP), formerly DEBI). Silver Spring, MD: Danya International; 2012. Available at http://www.effectiveinterventions.org/en/WhatsNew.aspx. Accessed September 22, 2014.
38. Effective interventions: HIV prevention that works (High Impact HIV/AIDS Prevention Project (HIP), formerly DEBI). Silver Spring, MD: Danya International; 2012. Available at https://effectiveinterventions.cdc.gov/en/HighImpactPrevention.aspx. Accessed September 22, 2014.
39. Blankenship KM, Bray SJ, Merson MH. Structural interventions in public health. AIDS 2000;14(Suppl 1):S11-21.
40. Sweat MD, Denison JA. Reducing HIV incidence in developing countries with structural and environmental interventions. AIDS 1995;9(Suppl A):S251-7.
41. Kidder DP, Wolitski RJ, Royal S, et al. Access to housing as a structural intervention for homeless and unstably housed people living with HIV: rationale, methods, and implementation of the housing and health study. AIDS Behav 2007;11(6 Suppl):149-61.
42. Wolitski RJ, Kidder DP, Pals SL, et al. Randomized trial of the effects of housing assistance on the health and risk behaviors of homeless and unstably housed people living with HIV. AIDS Behav 2010;14(3):493-503.
43. Holtgrave DR, Wolitski RJ, Pals SL, et al. Cost-utility analysis of the housing and health intervention for homeless and unstably housed persons living with HIV. AIDS Behav 2013;17(5):1626-31.
44. Social determinants of health among adults with diagnosed HIV infection in 20 States, the District of Columbia, and Puerto Rico, 2010. HIV Surveillance Supplemental Report 2014;19(2). Revised edition. Atlanta, GA: Centers for Disease Control and Prevention; July 2014. Available at http://www.cdc.gov/hiv/pdf/surveillance_Report_vol_19_no_2.pdf. Accessed September 22, 2014.
45. Syringe exchange programs --- United States, 2008. MMWR 2010;59(45):1488-91. Available at http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5945a4.htm. Accessed September 22, 2014.
46. Hagan H, Des Jarlais DC, Purchase D, et al. The Tacoma syringe exchange. J Addict Dis 1991;10(4):81-8.
47. Heimer R, Kaplan EH, Khoshnood K, et al. Needle exchange decreases the prevalence of HIV-1 proviral DNA in returned syringes in New Haven, Connecticut. Am J Med 1993;95(2):214-20.
48. Jones S. NIDA/CSAT/CDC workshop on the use of bleach for the decontamination of drug injection equipment. Proceedings. Baltimore, MD, February 9-10, 1994. J Acquir Immune Defic Syndr 1994;7(7):741-76.

49. Kahn JG. Economic evaluation of primary HIV prevention in intravenous drug users. In Holtgrave D (Ed.), Handbook of economic evaluation of HIV prevention programs. New York: Plenum Press; 1998.
50. Kahn JG, Sanstad KC. The role of cost-effectiveness analysis in assessing HIV-prevention interventions. AIDS Public Policy J 1997;12(1):21-30.
51. Kaplan EH, Heimer R. A model-based estimate of HIV infectivity via needle sharing. J Acquir Immune Defic Syndr 1992;5(11):1116-8.
52. Kaplan EH, Heimer R. HIV incidence among needle exchange participants: estimates from syringe tracking and testing data. J Acquir Immune Defic Syndr 1994;7(2):182-9.
53. Nguyen TQ, Weir BW, Des Jarlais DC, et al. Syringe Exchange in the United States: A National Level Economic Evaluation of Hypothetical Increases in Investment. AIDS Behav 2014 May 15;epub.
54. Del Romero J, Castilla J, Hernando V, et al. Combined antiretroviral treatment and heterosexual transmission of HIV-1: cross sectional and prospective cohort study. Br Med J 2010;340:c2205.
55. Donnell D, Baeten JM, Kiarie J, et al. Heterosexual HIV-1 transmission after initiation of antiretroviral therapy: a prospective cohort analysis. Lancet, 2010;375(9731):2092-8.
56. Horstmann E, Brown J, Islam F, et al. Retaining HIV-infected patients in care: Where are we? Where do we go from here? Clin Infect Dis 2010;50(5):752-61.
57. Cabral HJ, Tobias C, Rajabiun S, et al. Outreach program contacts: do they increase the likelihood of engagement and retention in HIV primary care for hard-to-reach patients? AIDS Patient Care STDS 2007;21(Suppl 1):S59-67.
58. Mugavero MJ. Improving engagement in HIV care: what can we do? Top HIV Med 2008;16(5):156-61.
59. Maulsby C, Sifakis F, German D, et al. Differences and similarities in HIV testing among men who have sex with men and women (MSMW) and men who have sex with men only (MSMO). Open AIDS J 2012;6:53-9.
60. Maulsby C, Sifakis F, German D, et al. Partner characteristics and undiagnosed HIV seropositivity among men who have sex with men only (MSMO) and men who have sex with men and women (MSMW) in Baltimore. AIDS Behav 2012;16(3):543-53.
61. Branson BM, Handsfield HH, Lampe MA, et al. Revised recommendations for HIV testing of adults, adolescents, and pregnant women in health-care settings. MMWR Recomm Rep 2006;55(RR-14):1-17; quiz CE11-14.
62. Castillo M, Palmer BJ, Rudy BJ, et al. Creating partnerships for HIV prevention among YMSM: the Connect Protect(R) Project and House and Ball Community in Philadelphia. J Prev Interv Community 2012;40(2):165-75.
63. Latkin CA, Davey-Rothwell MA, Knowlton AR, et al. Social network approaches to recruitment, HIV prevention, medical care, and medication adherence. J Acquir Immune Defic Syndr 2013;63(Suppl 1):S54-8.
64. Higa DH, Marks G, Crepaz N, et al. Interventions to improve retention in HIV primary care: a systematic review of U.S. studies. Curr HIV/AIDS Rep 2012;9(4):313-25.
65. Liau A, Crepaz N, Lyles CM, et al. Interventions to promote linkage to and utilization of HIV medical care among HIV-diagnosed persons: a qualitative systematic review, 1996-2011. AIDS Behav 2013;17(6):1941-62.
66. Charania MR, Marshall KJ, Lyles CM, et al. Identification of evidence-based interventions for promoting HIV medication adherence: findings from a systematic review of U.S.-based studies, 1996-2011. AIDS Behav 2014;18(4):646-60.
67. Lo W, MacGovern T, Bradford, J. Association of ancillary services with primary care utilization and retention for patients with HIV/AIDS. AIDS Care 2002;14(Suppl 1):S45-57.
68. Keruly JC, Conviser R, Moore RD. Association of medical insurance and other factors with receipt of antiretroviral therapy. Am J Public Health 2002;92(5):852-7.
69. MMWR Weekly Reports. Atlanta, GA: Centers for Disease Control and Prevention; 2014. Available at http://www.cdc.gov/hiv/statistics/systems/nhbs/reports.html. Accessed September 23, 2014.
70. National HIV Behavioral Surveillance (NHBS). Atlanta, GA: Centers for Disease Control and Prevention; 2014. Available at http://www.cdc.gov/hiv/statistics/systems/nhbs/. Accessed September 23, 2014.
71. German D, Linton S, Cassidy-Stewart H, et al. Using Baltimore HIV behavioral surveillance data for local HIV prevention planning. AIDS Behav 2014;18(Suppl 3):359-69.
72. Holtgrave DR, Maulsby C, Wehrmeyer L, et al. Behavioral factors in assessing impact of HIV treatment as prevention. AIDS Behav 2012;16(5):1085-91.
73. Nadery S, Geerlings SE. Pre-exposure prophylaxis (PrEP) in HIV-uninfected individuals with high-risk behaviour. Neth J Med, 2013;71(6):295-9.
74. Interim guidance: preexposure prophylaxis for the prevention of HIV infection in men who have sex with men. MMWR 2011;60(3):65-8. Available at http://www.cdc.gov/mmwr/preview/mmwrhtml/mm6003a1.htm. Accessed September 23, 2014.
75. Preexposure prophylaxis for the prevention of HIV infection in the United States - 2014: a clincial practice guideline. Atlanta, GA: Centers for Disease Control and Prevention; 2014. Available at http://www.cdc.gov/hiv/pdf/guidelines/PrEPguidelines2014.pdf. Accessed October 1, 2014.

76. Consolidated guidelines on HIV prevention, diagnosis, treatment, and care for key populations. Geneva: World Health Organization; 2014. Available at http://apps.who.int/iris/bitstream/10665/128048/1/9789241507431_eng.pdf?ua=1&ua=1. Accessed September 23, 2014.
77. Holtgrave D (Ed). Handbook of economic evaluation of HIV prevention programs. New York: Plenum; 1998.
78. Pinkerton SD, Johnson-Masotti AP, Holtgrave DR, et al. Using cost-effectiveness league tables to compare interventions to prevent sexual transmission of HIV. AIDS 2001;15(7):917-28.
79. Gomez GB, Borquez A, Case KK, et al. The cost and impact of scaling up pre-exposure prophylaxis for HIV prevention: a systematic review of cost-effectiveness modelling studies. PLoS Med 2013;10(3):e1001401.
80. Paltiel AD, Freedberg KA, Scott CA, et al. HIV preexposure prophylaxis in the United States: impact on lifetime infection risk, clinical outcomes, and cost-effectiveness. Clin Infect Dis 2009;48(6):806-15.
81. Cassidy-Stewart H, Holtgrave D. Mathematical modeling and economic analysis to improve local HIV prevention and care planning: Baltimore/Towson. Paper presented at the 2011 National HIV Prevention Conference, Atlanta, GA.
82. Holtgrave DR, Young, PA, Mayer RR, et al. Employing resource allocation modeling to inform HIV prevention planning for the state of Iowa. AIDS Educ Prev 2013;25(5):423-9.
83. Holtgrave DR, Kim JJ, Adkins C, et al. Unmet HIV service needs among black men who have sex with men in the United States. AIDS Behav 2014;18(1):36-40.
84. Holtgrave DR, Hall HI, Wehrmeyer L, et al. Costs, consequences and feasibility of strategies for achieving the goals of the National HIV/AIDS strategy in the United States: a closing window for success? AIDS Behav 2012;16(6):1365-72.
85. Holtgrave D. On the epidemiologic and economic importance of the National AIDS Strategy for the United States. J Acquir Immune Defic Syndr 2010;55(2):139-42.
86. Holtgrave DR. Development of year 2020 goals for the National HIV/AIDS Strategy for the United States. AIDS Behav 2014;18(4):638-43.
87. Dowdle WR, Hopkins D. The eradication of infectious diseases. Hoboken, NJ: Wiley; 1998.
88. Hinman A. Eradication of vaccine-preventable diseases. Annu Rev Public Health 1999;20:211-29.
89. Smith G, Colfax G. (2012). Moving towards an AIDS-free generation (blog, November 28, 2012). Washington, DC: The White House, Office of National AIDS Policy. Available at http://www.whitehouse.gov/blog/2012/11/28/moving-towards-aids-free-generation. Accessed September 23, 2014.

David Holtgrave, PhD, is a professor and Chair of the Department of Health, Behavior and Society at the Johns Hopkins Bloomberg School of Public Health. He also holds joint appointments in the Johns Hopkins schools of medicine and nursing. Since 1991, Dr. Holtgrave's work in the field of HIV prevention has focused on the effectiveness and cost-effectiveness of HIV prevention interventions and the translation of those study findings to HIV prevention policy. He serves as vice-chair of the President's Advisory Council on HIV/AIDS. Previously, Dr. Holtgrave was a professor at the Rollins School of Public Health at Emory University. He also served as director of the Division of HIV/AIDS Prevention – Intervention Research and Support at the Centers for Disease Control and Prevention and was on the faculty at the Center for AIDS Intervention Research at the Medical College of Wisconsin.

Cathy Maulsby, PhD, is an Assistant Scientist in the Department of Health, Behavior and Society at the Johns Hopkins Bloomberg School of Public Health. Her research focuses on HIV prevention among underserved populations, including men who have sex with men and women and minority men who have sex with men. She currently is the principal investigator and director of evaluation for the national evaluation of Access to Care, an HIV prevention and care initiative funded by AIDS United that aims to improve the continuum of care for individuals who are out of care or in danger of falling out of care.



Chris Adkins, MSW, is a Senior Research Coordinator in the Department of Health, Behavior and Society at the Johns Hopkins Bloomberg School of Public Health. His work focuses on Lesbian, Gay, Bisexual, and Transgender (LGBT) health issues, particularly with regard to intersectionality and the life course perspective, as well as the facilitators and barriers to HIV medical care. He also serves on the Board of Directors of the Gay, Lesbian, Bisexual, and Transgender Community Center of Baltimore (GLCCB), where he works on HIV prevention and interventions with youth at-risk of homelessness.

Laura Fuentes, MM, is a key leader in the Lerner Center for Public Health Promotion in the Department of Health, Behavior and Society at the Johns Hopkins Bloomberg School of Public Health. Her research currently focuses on the evaluation of health promotion campaigns in the areas of smoking cessation and weight loss. She also serves as the program director for an ongoing collaboration between the Department of Health, Behavior and Society and the Monday Campaigns, a New York-based nonprofit organization committed to health promotion activities on the first day of the week.

Biological Influences on Development and Subsequent Well-Being

Epigenomics and the Unheralded Convergence of the Biological and Social Sciences

W. Thomas Boyce

Abstract

Given what we know, even now, about the interactive roles of genetic and environmental variation in pathogenesis, the advent of epigenomics, which illuminates the physical, molecular points of connection between genes and contexts, will almost certainly revolutionize theories of disease causation. Because social environments are among the most powerful determinants of health and illness, such theories will also increasingly blur conceptual distinctions between the social and biological sciences. Coming breakthroughs in epigenetic research are poised, as well, to elucidate fundamental enigmas of gene-environment interplay, such as the role of critical periods and developmental time, individual differences in susceptibility to environmental exposures, and the possible role of intergeneration transmission of risks. Solving such mysteries will profoundly influence the practice of medicine, the development of new preventive interventions, and the course of future research.

Introduction

Attention has been repeatedly called to the gaping conceptual and epistemic divisions between the biological and social sciences[1] or, more generally and historically, between the sciences and the humanities.[2] By contrast, this chapter presents an argument for the existence of a molecular convergence between the biological and social sciences—a convergence that will define and illuminate the common ground between nature and nurture. The nascent field of epigenetics is revealing a mechanism by which environmental variation influences gene expression through alterations in chromatin biology, thereby reconciling the organismic and contextual accounts for differences in health, disease, and human development.

The Human Genome Project advanced a revolutionary understanding of how human destiny is shaped by individual and population level differences in genetic sequences.[3,4] This sequencing of the human genome launched an exhilarating but often preoccupying debate over the roles played by common and rare genetic variants in the etiology of complex human diseases.[5,6] Concurrently, the burgeoning fields of social epidemiology, health psychology, and behavioral medicine have systematically revealed profound effects of socioemotional and cultural contexts on trajectories of development and risks of disease.[7-9] Genetically informed study designs have documented both genetic and contextual contributions to the origins of many, if not most, disease phenotypes, but the proportion of phenotypic variation attributable to differences in gene sequences remains, in most cases, substantially smaller than published heritability estimates.[10] A preemptive and catalytic

epigenomic science has now proposed that such "missing heritability" may be most readily discoverable in the epigenetic embedding of environmental exposures onto the background DNA sequence.[11,12] The obsolescent, genetic, and environmental forms of determinism spawned by each side of this scientific divide have given way to a new vision of how genetic biology and social environments physically intersect in the epigenetic events and processes shaping development and destiny.

Conrad Waddington first used the term "epigenetic" in relation to human development in a 1942 paper examining the processes and events by which genotypes are functionally linked to adult phenotypes.[13] The now flourishing science of epigenetics (from the Greek root *epi,* meaning upon or over) has been defined as the study of "structural adaptations of chromosomal regions so as to register, signal, or perpetuate altered [gene] activity states."[14] Epigenetic mechanisms change gene activity or expression by altering DNA organization without modifying the genetic code of the DNA.[11] As illustrated in Figure 1, chromatin is the physical packaging of DNA within chromosomes in the form of linear sequences of genes wrapped around groups of histone proteins, together resembling "beads on a string."

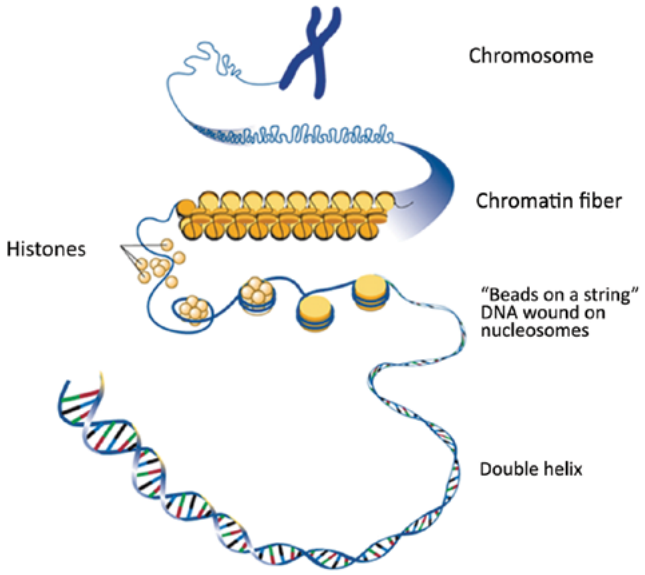


Figure 1: The chromatin packaging of DNA wound around histone protein octamers

Source: National Human Genome Research Institute; available at https://www.genome.gov/Glossary/index.cfm?id=32.

The multiple chemical tags—or epigenetic "marks"—placed on DNA (specifically, on the cytosine nucleotides within a cytosine-guanine, or CpG, sequence) or histone proteins as a consequence of environmental exposures and experiences, regulate the density of chromatin conformation, thereby modulating physical access of RNA polymerases, the decoding enzymes, to the DNA sequence. Epigenetic marks also determine the emergence of stable differences in cell types during embryogenesis, thus creating an "epigenetic paradox" in which histological stability over time and dynamic, moment-to-moment changes in response to experience are both encoded in the same

epigenetic patterns governing gene transcription.[15] The presence or absence of these chemical marks, such as DNA methylation and histone protein acetylation, methylation, ubiquitination, or phosphorylation, determines the organization of chromatin structure, which ultimately controls transcription. Transcription, in turn, governs the emergence of phenotypes through the influence of its protein products on, for example, the development of brain function and structure. Epigenetic marks are reversible through the activity of histone deacetylases (HDACs) and other demethylating enzymes, rendering chromatin restructuring processes potentially modifiable and accessible to pharmacological interventions that target such enzymes.[16] Patterns of epigenetic tagging of DNA and histone proteins thus constitute physical points of connection between genes and environments, not unlike the synapses that establish a physical nexus between neurons, allowing neuron-to-neuron communication and the formation of neural circuits.

Although patterns of epigenomic variation are clearly influenced by aspects of physical environmental exposure, such as diet and physical toxins,[17,18] it also has become clear that perturbations in socioemotional contexts can result in long-term changes in the human and animal epigenomes.[19-21] Epidemiologic research by Borghol et al.,[22] Essex et al.,[18]and Kobor and colleagues[23,24] has detected longitudinal associations between childhood disadvantage and genome-wide promoter methylation in mid-life, between parental stress in infancy and differential methylation in adolescence, and between early socioeconomic status (SES) and inflammatory gene expression in leukocyte transcriptomes. Essex and colleagues[18] found that systematic, long-term differences in DNA methylation in buccal epithelial cells were present in mid-adolescence among youth whose parents had reported high levels of stress and adversity in the youths' infancy and preschool years (see Figure 2). Further, the pattern of maternal and paternal differential methylation was commensurate with the known timing of parent-specific influences on development, i.e., greater maternal stressor effects in infancy, followed by an emergence of paternal effects in the preschool years. Such findings, of whole genome differences in DNA methylation status, are also consistent with the known cellular effects of stress hormones, such as glucocorticoids, which regulate expression of approximately 10 percent of the genome.[25] They are supported, as well, by reports of real time, dynamic changes in both promoter and exon region gene methylation among individuals exposed to laboratory-based, stressful challenges[26] and evidence of reduced expression of HDAC and proinflammatory genes under meditation conditions.[27] Epigenetic processes may also be implicated in the emergence of resilience[28] and in the beneficial developmental effects of enriched early environments.[29]

Chromatin modifications in gene-environment interplay

Modifications of chromatin also appear likely to underlie and, in some cases, mechanistically explain the gene x environment (GxE) interactions documented at an accelerating pace over the past decade among human samples. The reports of Caspi, Moffitt, and colleagues[30,31] from the Dunedin Multidisciplinary Health and Development Study revealed statistical interactions between early environmental conditions (e.g., child maltreatment and stressful life events) and functional gene polymorphisms (e.g., the *MAOA,* monoamine oxidase A, and *5HTT,* serotonin transporter, genes) in the prediction of antisocial behavior and depression/suicidality. Although legitimate concerns have been raised for the replicability of GxE interaction reports,[32] some such effects have been replicated

in independent samples,[33] and new evidence continues to accrue. GxE interactions have been observed and recorded in both human[34,35] and animal[36] species, leading some to assert that, in effect, virtually all disorders of health and development may be both genetic and environmental in origin, in the sense that all result from the mutually interactive influences of both.[37] The search for reliable GxE interactions may be abetted by the development of both empirical evidence that polygenic risk scores associated with developmental phenotypes can discern promising new single nucleotide polymorphism (SNP) targets[38] and computational models suggesting that genome-wide association study (GWAS) approaches to GxE discovery may be more promising than candidate SNP by SNP searches.[39] In fields such as psychiatric genomics, the way forward appears to lie in new knowledge of how multiple genes with additive or multiplicative effects, each with incremental influences, are assembled into functional genetic networks that interact with social environmental conditions to produce important phenotypic disorders.[40]

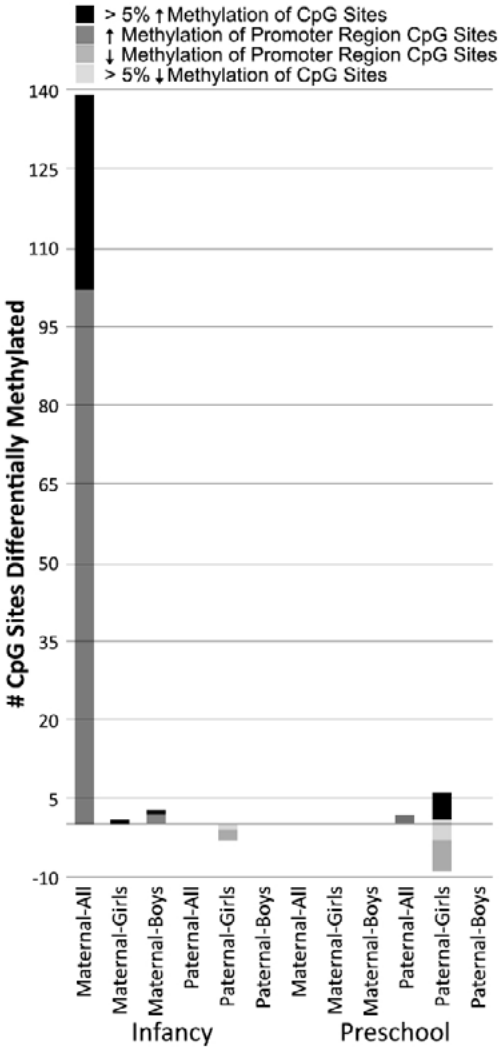


Figure 2: Differential methylation of buccal epithelial cell CpG sites in mid-adolescence, by parental reports of stress and adversity in children's infancy and preschool periods

Source: Essex MJ, Boyce WT, Hertzman C, et al. Epigenetic vestiges of early developmental adversity:

childhood stress exposure and DNA methylation in adolescence. Child Dev 2013;84(1):58-75. Used with permission.

Most recently, work by Binder and colleagues[41] has revealed a molecular process by which an epidemiologically observed interaction between a functional polymorphism in the gene coding for FK506 binding protein 5 (*FKBP5*, an intracellular regulator of the hypothalamic-pituitary-adrenocortical (HPA) stress response system) and childhood trauma predict symptoms of post-traumatic stress disorder (PTSD) in adulthood. Specifically, this work shows that the GxE interaction effect is mediated through demethylation of DNA in the glucocorticoid response elements of *FKBP5.* This observation is the first to show how chromatin modification and epigenetic marks may constitute the molecular mechanism for at least some GxE interactions. The work suggests that, rather than GxE interactions and epigenomic events being two distinctive forms of gene-environment interplay, epigenetic mechanisms may underlie most or all statistically documented GxE interactions.

Gene-environment interplay in the pathogenesis of disease and maladaptive development is important because the burden of such disorders is not randomly distributed within human populations. A disproportionate share of morbidities falls upon and afflicts a small subset of individuals—usually about 15-20 percent of both children and adults—who sustain well over half of the total illness and disability.[42] Much of this excessively burdened subpopulation is made up of those who live in conditions of poverty and, as a consequence, sustain far greater exposures to adversity and stress, as well as to environmental toxins, food insecurity and dietary constraints, and inadequate medical care. The work of Evans and colleagues[43] has shown how children from impoverished communities experience greater housing density, ambient noise, family chaos, and neighborhood violence than their counterparts from middle income homes. There is now strong evidence that such adversity— especially early in life—is associated with disturbances of childhood mental health, more disordered developmental trajectories, poorer educational achievements, and lifelong risks of chronic disorders of health and well-being.[41,44,45] Further, beyond childhood, the studies of Anda[46] and Felitti[47] have revealed linkages between adverse childhood events and long-term risks of disorders of both mental and physical health in adulthood. Striking social class differences in incident psychiatric and biomedical disorders are believed to be attributable to lasting differences in phenotype, stemming from adversity-related perturbations in epigenetic processes.[22]

Following upon the work of Meaney and colleagues demonstrating epigenetically mediated links between maternal deprivation in rat pups and lifelong stress reactivity,[11,48] human epidemiologic research by Borghol et al,[21] and Kobor and colleagues[22,23] has detected longitudinal associations between childhood disadvantage and genome-wide promoter methylation in mid-life and between early SES and inflammatory gene expression in leukocyte transcriptomes. Prenatal exposures to famine and adversity during the 1944-45 Dutch Hunger Winter have also been associated with differential methylation in a variety of other developmentally and immunologically active genes.[49,50] Similarly, institutionalized children from the former Russian Federation have been reported to show whole genome hypermethylation, compared to parent-reared controls.[51] Oberlander and colleagues[52] reported increased methylation of the glucocorticoid receptor (*GR*) gene in infants born to mothers with prenatal depression, and Radtke et al,[53] derived similar findings from the leukocytes of adolescents whose mothers were exposed to intimate partner violence during pregnancy. By

decreasing *GR* expression through epigenetic marks, the adrenocortical stress response system may be perennially up-regulated, resulting in maladaptive response to stress and trauma. Such findings, of both whole genome and gene-specific differences in DNA methylation status, are consistent with the cellular effects of stress hormones, such as glucocorticoids.[24,54] Taken together, these observations offer substantial evidence for an extensive interplay among genes, epigenomes, and social environments in the genesis of perturbed or problematic social and biological development.

Individual Differences in Biological Responses to Adversity

Biological and psychological responses to early life stressors are not uniform across individuals, and there is abundant and emerging evidence for substantial individual differences in susceptibility to environmental influences, both positive and negative. The collective work of Boyce and Ellis,[55,56] Belsky,[57] and van IJzendoorn and colleagues[58] has documented, at levels of analysis ranging from genetic variation and physiological stress reactivity to temperamental characteristics and individual behavior, the presence of childhood subgroups, usually comprising approximately one in five children, who sustain either the worst or best developmental and health outcomes, conditional upon the character of the social environments in which they are being reared. Children with high levels of laboratory-based, autonomic nervous system reactivity to challenge show either the highest or lowest levels of externalizing behavior problems, depending on their exposure to significant marital conflict within their families,[59] and Bush et al,[60] have recently shown that the *BDNF* Val66Met polymorphism confers an increased neuroendocrine sensitivity to socioeconomic context, with Met-carriers having the highest or lowest cortisol expression levels, depending on SES. This differential susceptibility to environmental influences has been demonstrated in both non-human primates[35,61] and human children[62] and detected at the levels of allelic variation,[59,63] autonomic and adrenocortical stress responsivity,[64,65] and individual differences in temperament and behavior.[56]

Differential susceptibility has also been shown to occur in relation to epigenetic marks and chromatin modifications. Beach et al,[66] in a sample of African American youth from working poor communities, found that cumulative socioeconomic adversity and the S- allele of the *5HTT,* serotonin transporter gene interactively predicted promoter region methylation within a group of 200+ depression-related genes. Youth with the S-allele had either the highest or lowest levels of depression-related gene methylation, depending on the intensity of exposures to poverty-related stress. Strunk and colleagues[67] have argued that the increased susceptibility of infants to infectious disease agents may be due to differential methylation of immune-regulating and other developmentally salient genes. Further, the Binder laboratory[40] has demonstrated a differential susceptibility of individuals bearing the AG/AA "risk" allele of the *FKBP5* gene (see Figure 3). Such individuals have either the highest or lowest rates of adult PTSD, conditional upon childhood exposures to sexual and/or physical abuse. Further, the molecular process by which this epidemiologically-observed interaction occurs is mediated through DNA demethylation in the glucocorticoid response elements of *FKBP5*. These observations are among the first to show how chromatin modification and epigenetic marks might constitute the actual molecular mechanisms for a differential susceptibility to environmental conditions.

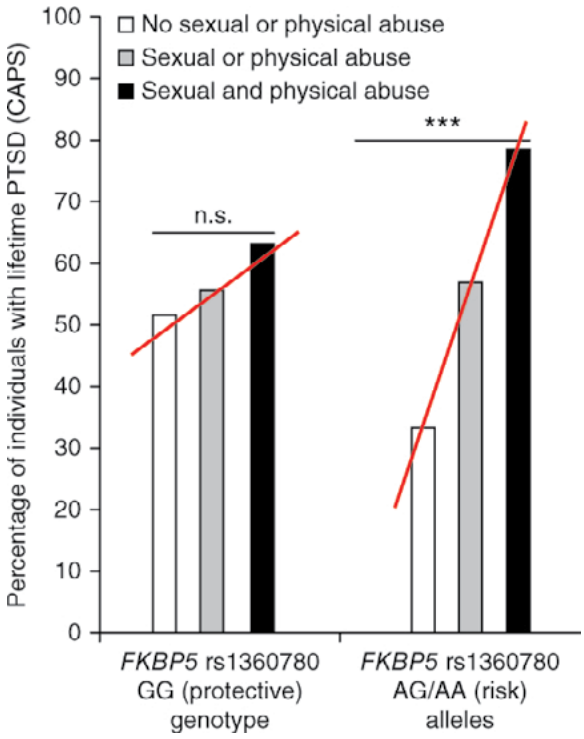


Figure 3: The interaction of early trauma exposure and FKBP5 risk allele predicting lifetime post-traumatic stress disorder (PTSD)

Source: Modified from Klengel T, Mehta D, Anacker C, et al. Allele-specific FKBP5 DNA demethylation mediates gene-childhood trauma interactions. Nat Neurosci 2013;16(1):33-41. Used with permission.

Note: Consistent with the differential susceptibility hypothesis, the "risk" allele is protective in the absence of trauma.

Critical Periods in the Epigenetic Regulation of Neurodevelopment

These epigenetically-mediated differences in sensitivity to early environments also emerge within a temporal framework. For example, unilateral blindness occurs if an occlusion of vision in one eye (due to strabismus, cataracts, or other causes) is not corrected during the early period of visual cortical development (i.e., birth to 6 or 7 years of age). This constitutes a so-called critical period, defined as a developmental interval in which the presence or absence of important experiences or exposures result in irreversible change in brain circuitry.[68] Critical periods have been described in the Bucharest Early Intervention Project, in which institutionalized Romanian children were randomized to either orphanage care-as-usual or placement into foster care.[69] As shown in Figure 4, teacher-rated social skills at 8 years of age were critically influenced by whether foster placement occurred before or after the age of 20 months; foster care children placed before 20 months had skills comparable to their never-institutionalized peers, while those placed after 20 months had the same deficits as children who remained in orphanages.[70]

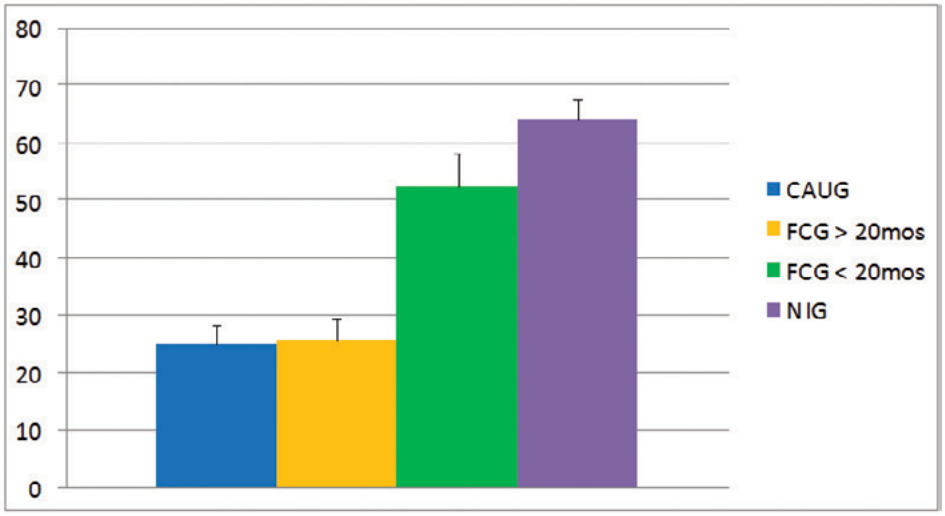


Figure 4: Teacher-rated social skills of Romanian orphans and never-institutionalized (NIG) Romanian children at 8 years of age, by assignment to a care as usual group (CAUG) versus a foster care group (FCG) before and later than 20 months of age.

Source: Almas A, Degnan K, Radulescu A, et al. The effects of early intervention and the moderating effects of brain activity on institutionalized children's social skills at age 8. Proc Nat Acad Sci U S A 2012;109(Suppl 2):17228-31. Used with permission.

There is new evidence that such critical periodicity in the acquisition of developmental skills may be mediated by epigenetic processes. Takesian and Hensch[71] have shown how molecular "triggers" and "brakes" initiate and constrain plasticity in the brain over time and how the molecular events underlying critical period onset and offset are epigenetic in origin. Critical period onset appears guided and timed by the maturation of excitatory-inhibitory (E-I) circuit balance, and epigenetic factors regulate the expression of the *GAD67* gene that codes for the GABA inhibitory neurotransmitter involved in E-I equilibration. Various pharmacological agents targeting epigenetic processes—drugs such as valproate, a histone deacetylase (HDAC) inhibitor—can shift the timing of critical period onset. The closure of the critical period for ocular dominance acquisition, on the other hand, involves the down-regulation of vision-dependent histone acetylation and phosphorylation. Valproate has been shown, moreover, to reopen the critical window for the acquisition of absolute pitch,[72] and E-I circuitry imbalance and critical period timing errors have been recognized within mouse models of autism spectrum disorder.[73] Thus, there is at least provisional evidence that much of the molecular machinery underlying critical period onset and offset is epigenetic in origin.[74]

Summary: What's Known and Unknown

Dramatic advancements in understanding and manipulating developmental events—those that catalyze the amplification or diminution of risks to mental or physical health—are often accompanied by a wholesale attribution of as yet unexplained biological enigmas to the newly discovered events. Thus, as knowledge of the epigenetic origins of key developmental processes has

come into view, chromatin modifications have become the mechanism *du jour* for explaining and exploiting molecular biology and gene-environment interplay. In such a scientific setting, it may therefore be useful to take stock of what we now actually know and what we do not yet understand.

What We Know

A reasonable summary of what we do know, partially or in full, would include the following points:

- Socially partitioned childhood adversities have potent and pervasive linkages to health and development, with some persisting over the lifetime of the individual; these effects are at least partially mediated by epigenetic processes guiding the adaptive differential expression of adversity-responsive genes.
- The effects of stress and adversity are also highly variable, both in terms of physiological and behavioral responses and in consequent physical and mental morbidities; a subset of childhood populations shows evidence of exquisite sensitivities to the character of both aversive and supportive social environments—that is, a differential susceptibility to socioemotional environmental influence.
- The resultant, high level of individual differences in health and development is thus an interactive product of biological susceptibility and environmental conditions; since genetic variation plays a role in the acquisition of environmental sensitivity, most disorders of maladaptive development will likely be attributable to the epigenetic co-operation of genetic and environmental variation.
- Epigenetic modifications of chromatin structure thus appear, at least to some degree, to mediate biological and psychological responses to adversity, to be the basis for differences in susceptibility to context, and to underlie the inception and termination of developmentally salient, critical periods.

Such discoveries hold great potential promise for understanding the striking individual differences in developmental trajectories and health within human populations. Nonetheless, this emerging field of developmental epigenetics presently remains, at best, in its infancy.

What We Don't Know

What is not yet known about chromatin modification, its origins and consequences, far outstrips that which we know. Among the research questions still incompletely explored or completely unexplored are the following:

- What are the molecular events by which GxE interactions occur? How are experiences and exposures molecularly transduced into the chemical marks and other processes that constitute the epigenome and regulate gene transcription?
- How can dynamic epigenetic change be measured against the backdrop of systemic, tissue-specific differences in epigenetic profiles? What (if anything) can be learned about early adversity-related epigenetic modifications in hippocampal neurons from those found in buccal epithelial cells (BECs)?

- What are the timeframes for dynamic epigenetic change? Are different forms of chromatin modification more or less rapid or slow, stable or evanescent? Can epigenetic interventions modify the windows of critical periodicity for psychosocial exposures, as well as those for physical exposures?
- What epigenetic basis may there be for the inter-generational character of severe adversity effects? How are risks for psychopathology transmitted from one generation to the next, sometimes despite an absence of exposure in the latter?
- Is there an epigenetic account for the remarkable heterogeneity in disease phenotypic effects of genetic (i.e., allelic) variation? How do assemblies of common and rare genetic variants operate conjointly with environmentally mediated epigenetic change to produce risks for complex disorders of human health?

Implications for Practice

In fact, this heterogeneity in disease phenotype, environmental exposures, genetic susceptibilities, and responsivity to treatment is among the most challenging dilemmas in the medical care of human populations. Addressing such complexity is both the promise and the profound difficulty of achieving a new "precision medicine," defined as "the tailoring of medical treatment to the individual characteristics of each patient."[75] Such a vision lies notably beyond the (relatively) facile challenges of pharmacogenetics, which would select medications based on the biogenomic characteristics of an individual patient. Precision medicine, by contrast, would discover, fabricate, and select medications based not only on diagnosis and pathogenic agent, but also on sensitivities known and unknown to the patient, host allelic variants with implications for drug metabolism and clearance, and genetic features of both agent and neoplasm. A full rendition of precision medicine would also demand and create a new taxonomy of human disease, involving the molecular disaggregation of old, organ-focused categories of morbidity and the use of vast new data sets to discover currently unknown cross-disease and cross-system commonalities in mechanism.

The science of epigenetics could play a key, catalytic role in the discernment of such a new disease taxonomy. Because most disorders will ultimately be shown to have both environmental and gene-based etiologic components, examination of the molecular interface between such interactive components, i.e., the epigenome, will almost certainly become a standard diagnostic procedure. Assessment of epigenetic profiles is indeed already a standard of high level care in oncology and psychiatry, where chromatin modifications serve as biomarkers of disease states and progression.[76]

The implications of using epigenetic science to effect a shift to a new disease taxonomy would be legion. The late 19th century identification of the tubercle bacillus as the pathogenic agent in tuberculosis (TB) led to an unanticipated recombination of scrofula (lymphatic TB), Pott disease (TB of the spine), pulmonary "consumption," and tuberculous meningitis—previously regarded as distinctive forms of morbidity—into a single disease category, i.e., the multiple forms and manifestations of infection with *Mycobacterium tuberculosis*. Similarly, insights into the epigenetic commonalities among once separate diagnoses, each with its own corresponding therapeutic approach, could shift, nearly overnight, the way that health care providers think about diagnostic categories, approach laboratory confirmation of a given diagnosis, and engage therapeutic strategies.

Implications for Research

Given the complexities that novel diagnostic and therapeutic practices would entail, the research enterprise that would guide and sustain such a shift in medical practice would require a new level of commitment to and investment in the emerging science of complexity and dynamic systems.[77] The computational models being constructed as the methodological platform of that science incorporate key features only modestly (and insufficiently) represented in current, conventional research and biostatistical methods: non-linearity of associations, multilevel causation, changes over developmental time, and interrelations among multiple, causal covariates. Given the scale and complexity of the epigenetic processes at work in human health and disease, the advent of new complex modeling approaches may be not only welcome but essential. Using such models to simulate experimental manipulations of causal factors, some of the epigenetic puzzles that might be addressable are the possible transgenerational inheritance of epigenetic marks, the origins and plasticity of critical and sensitive developmental periods, and the emergence, within populations, of individuals with either exceptional sensitivity or exceptional resilience to the privations of early life in impoverished, high adversity settings.

When viewed from the perspective of future generations of scientists and scholars, the most profound and momentous residual of this epigenetic era in the history of science may well be its dissolution of the boundary between the biological and social sciences. What began as two wholly different views of human nature has given way to a singular, new vision of how genetic and social environmental variation work together on risks for disease and maladaptive development. What began as two alternative or even antagonistic sets of methods, models, and theories—one biological and the other psychosocial in focus—have been forcibly and brilliantly reconciled by the scientific needs of each for the perspective of the other. It is almost certainly safe to claim that human biology and the social sciences will never be the same.

Acknowledgments

The research upon which this report is based was supported by the National Institute of Mental Health (grant award R24MH081797), the MacArthur Foundation Research Network on Psychopathology and Development, the Sunny Hill Health Centre/British Columbia Leadership Chair in Child Development at the University of British Columbia, and the Child and Brain Development Program of the Canadian Institute for Advanced Research. The opinions presented in this chapter are those of the author and do not necessarily represent the position of the Agency for Healthcare Research and Quality, the National Institutes of Health, or the U.S. Department of Health and Human Services.

Author's Affiliation

Departments of Pediatrics and Psychiatry, University of California, San Francisco; Canadian Institute for Advanced Research, Toronto, Canada.

Address correspondence to: W. Thomas Boyce, MD, UCSF Division of Developmental Medicine, Department of Pediatrics, 550 16th Street, 4th Floor, San Francisco, CA 94158;
email: tom.boyce@ucsf.edu.

References

1. Weingart P, Mitchell SD, Richerson PJ, et al. Human by nature: between biology and the social sciences. Mahwah, NJ: Lawrence Ehrlbaum; 2013.
2. Snow CP. The two cultures. Cambridge, UK: Cambridge University Press; 1959.
3. Bailey JN, Pericak-Vance MA, Haines JL. (2014). The impact of the human genome project on complex disease. Genes (Basel) 2014;5(3):518-35.
4. Lander ES. Initial impact of the sequencing of the human genome. Nature 2011;470(7333):187-97.
5. Gibson G. Hints of hidden heritability in GWAS. Nat Genet 2010;42(7):558-60.
6. McClellan J, King MC. Genomic analysis of mental illness: a changing landscape. JAMA 2010;303(24):2523-4.
7. Adler NE, Rehkopf DH. U.S. disparities in health: descriptions, causes, and mechanisms. Annu Rev Public Health 2008;29:235-252.
8. Berkman L, Epstein AM. Beyond health care – socioeconomic status and health. N Engl J Med 2008;358(23):2509-10.
9. Syme SL. Historical perspective: the social determinants of disease - some roots of the movement. Epidemiol Perspect Innov 2005;2(1):2.
10. Zaitlen N, Kraft P. Heritability in the genome-wide association era. Hum Genet 2012;131(10):1655-64.
11. Meaney, MJ. Epigenetics and the biological definition of gene x environment interactions. Child Dev 2010;81(1):41-79.
12. Nelson VR, Nadeau JH. Transgenerational genetic effects. Epigenomics 2010;2(6):797-806.
13. Waddington CH. The epidenotype. Endeavour 1942;1:18-20.
14. Bird A. Perceptions of epigenetics. Nature 2007;447(7143):396-8.
15. Sweatt JD, Meaney MJ, Nestler EJ, et al. Epigenetic regulation in the nervous system: basic mechanisms and clinical impact. London: Elsevier; 2013.
16. Allis CD, Jenuwein T, Reinberg D, et al. Epigenetics. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 2007.
17. Drummond EM, Gibney ER. Epigenetic regulation in obesity. Curr Opin Clin Nutr Metab Care 2013;16(4):392-7.
18. LaSalle JM. A genomic point-of-view on environmental factors influencing the human brain methylome. Epigenetics 2011;6(7):862-9.
19. Essex MJ, Boyce WT, Hertzman C, et al. Epigenetic vestiges of early developmental adversity: childhood stress exposure and DNA methylation in adolescence. Child Dev 2013;84(1):58-75.
20. Provencal N, Suderman MJ, Guillemin C, et al. The signature of maternal rearing in the methylome in rhesus macaque prefrontal cortex and T cells. J Neurosci 2012;32(44):15626-42.
21. Tung, J., Barreiro, L. B., Johnson, Z. P., et al. Social environment is associated with gene regulatory variation in the rhesus macaque immune system. Proc Natl Acad Sci U S A 2012;109(17):6490-5.
22. Borghol N, Suderman M, McArdle W, et al. Associations with early life socio-economic position in adult DNA methylation. Int J Epidemiol 2012;41(1):62-74.
23. Lam LL, Emberly E, Fraser HB, et al. Factors underlying variable DNA methylation in a human community cohort. Proc Natl Acad Sci U S A 2012;109(Suppl 2):17253-60.
24. Powell ND, Sloan EK, Bailey MT, et al. Social stress up-regulates inflammatory gene expression in the leukocyte transcriptome via beta-adrenergic induction of myelopoiesis. Proc Natl Acad Sci U S A 2013;110(41):16574-9.
25. Matthews SG, Phillips DI. Transgenerational inheritance of stress pathology. Exp Neurol 2012;233(1):95-101.
26. Unternaehrer E, Luers P, Mill J, et al. Dynamic changes in DNA methylation of stress-associated genes (OXTR, BDNF) after acute psychosocial stress. Transl Psychiatry 2012;2:e150.
27. Kaliman P, Alvarez-Lopez MJ, Cosin-Tomas M, et al. Rapid changes in histone deacetylases and inflammatory gene expression in expert meditators. Psychoneuroendocrinology 2014;40:96-107.
28. Zannas AS, West AE. Epigenetics and the regulation of stress vulnerability and resilience. Neuroscience 2014;264:157-70.
29. Sweatt JD. Experience-dependent epigenetic modifications in the central nervous system. Biol Psychiatry 2009;65(3):191-7.
30. Caspi A, McClay J, Moffitt TE, et al. Role of genotype in the cycle of violence in maltreated children. Science, 2002;297(5582):851-4.
31. Caspi A, Sugden K, Moffitt TE, et al. Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. Science 2003;301(5631):386-9.
32. Duncan LE, Keller MC. A critical review of the first 10 years of candidate gene-by- environment interaction research in psychiatry. Am J Psychiatry 2011;168(10):1041-9.
33. van Winkel M, Peeters F, van Winkel R, et al. Impact of variation in the BDNF gene on social stress sensitivity and the buffering impact of positive emotions: replication and extension of a gene-environment interaction. Eur Neuropsychopharmacol 2014;24(6):930-8.
34. Drury SS, Theall KP, Smyke AT, et al. Modification of depression by COMT val158met polymorphism

in children exposed to early severe psychosocial deprivation. Child Abuse Negl 2010;34(6):387-95.
35. Zohsel K, Buchmann AF, Blomeyer D, et al. Mothers' prenatal stress and their children's antisocial outcomes–a moderating role for the Dopamine D4 Receptor (DRD4) gene. J Child Psychol Psychiatry 2014;55(1):69-76.
36. Barr CS, Newman TK, Lindell S, et al. Interaction between serotonin transporter gene variation and rearing condition in alcohol preference and consumption in female primates. Arch Gen Psychiatry 2004;61(11):1146-52.
37. Rothman KJ, Greenland S. Causation and causal inference in epidemiology. Am J Public Health, 2005;95(Suppl 1):S144-50.
38. Rietveld CA, Medland SE, Derringer J, et al. GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. Science 2013;340(6139):1467-71.
39. Marigorta UM, Gibson G. A simulation study of gene-by-environment interactions in GWAS implies ample hidden effects. Front Genet 2014;5:225.
40. Gratten J, Wray NR, Keller MC, et al. Large-scale genomics unveils the genetic architecture of psychiatric disorders. Nat Neurosci 2014;17(6):782-90.
41. Klengel T, Mehta D, Anacker C, et al. Allele-specific FKBP5 DNA demethylation mediates gene-childhood trauma interactions. Nat Neurosci 2013;16(1):33-41.
42. Hertzman C, Boyce WT. (2010). How experience gets under the skin to create gradients in developmental health. Annu Rev Public Health 31:329-47 3p following 347.
43. Evans GW, Chen E, Miller G, et al. How poverty gets under the skin: a life- course perspective. In Maholmes V, King RB (Eds), The Oxford handbook of poverty and child development (pp. 13-36). Oxford: Oxford University Press; 2012.
44. Boyce WT, Sokolowski MB, Robinson GE. Toward a new biology of social adversity. Proc Natl Acad Sci U S A 2012;109(Suppl 2):17143-8.
45. Shonkoff JP, Boyce WT, McEwen BS. Neuroscience, molecular biology, and the childhood roots of health disparities: building a new framework for health promotion and disease prevention. JAMA 2009;301(21):2252-9.
46. Anda RF, Croft JB, Felitt, VJ, et al. Adverse childhood experiences and smoking during adolescence and adulthood. JAMA 1999;282(17):1652-8.
47. Felitti VJ, Anda RF, Nordenberg D, et al. Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults. The Adverse Childhood Experiences (ACE) Study. Am J Prev Med 1998;14(4):245-58.
48. Weaver IC, Meaney MJ, Szyf M. Maternal care effects on the hippocampal transcriptome and anxiety-mediated behaviors in the offspring that are reversible in adulthood. Proc Natl Acad Sci U S A 2006;103(9):3480-5.
49. Heijmans BT, Tobi EW, Stein AD, et al. Persistent epigenetic differences associated with prenatal exposure to famine in humans. Proc Natl Acad Sci U S A 2008;105(44):17046-9.
50. Tobi EW, Lumey LH, Talens RP, et al. DNA methylation differences after exposure to prenatal famine are common and timing- and sex-specific. Hum Mol Genet 2009;18(21):4046-53.
51. Naumova OY, Lee M, Koposov R, et al. Differential patterns of whole-genome DNA methylation in institutionalized children and children raised by their biological parents. Dev Psychopathol 2012;24(1):143-55.
52. Oberlander TF, Weinberg J, Papsdorf M, et al. Prenatal exposure to maternal depression, neonatal methylation of human glucocorticoid receptor gene (NR3C1) and infant cortisol stress response. Epigenetics 2008;3(2):97-106.
53. Radtke KM, Ruf M, Gunter HM, et al. Transgenerational impact of intimate partner violence on methylation in the promoter of the glucocorticoid receptor. Transl Psychiatry 2011;1:e21.
54. Hostinar CE, Sullivan RM, Gunnar MR. Psychobiological mechanisms underlying the social buffering of the hypothalamic-pituitary-adrenocortical axis: a review of animal models and human studies across development. Psychol Bull 2014;140(1):256-82.
55. Boyce WT, Ellis BJ. Biological sensitivity to context. I. An evolutionary-developmental theory of the origins and functions of stress reactivity. Dev Psychopathol 2005;17(2):271-301.
56. Ellis BJ, Essex MJ, Boyce WT. Biological sensitivity to context. II. Empirical explorations of an evolutionary-developmental hypothesis. Dev Psychopathol 2005;17(2):303-28.
57. Belsky J. Differential susceptibility to rearing influence: an evolutionary hypothesis and some evidence. In Ellis BJ, Bjorklund DF (Eds), Origins of the social mind: evolutionary psychology and child development (pp. 139-163). New York: Guilford; 2005.
58. Bakermans-Kranenburg MJ, van Ijzendoorn MH. (2011). Differential susceptibility to rearing environment depending on dopamine-related genes: new evidence and a meta-analysis. Dev Psychopathol 2011;23(01):39-52.
59. Obradovic J, Bush NR, Boyce WT. The interactive effect of marital conflict and stress reactivity on externalizing and internalizing symptoms: the role of laboratory stressors. Dev Psychopathol 2011;23(1):101-14.
60. Bush N, Guendelman M, Adler N, et al. BDNF allelic variants moderate social disparities in children's basal cortisol expression. 2014; submitted.
61. Boyce WT, O'Neill-Wagner P, Price CS, et al. Crowding stress and violent injuries among behaviorally inhibited rhesus macaques. Health Psychol 1998;17(3):285-9.
62. Ellis BJ, Boyce WT, Belsky J, et al. Differential susceptibility to the environment: an evolutionary–neurodevelopmental theory. Dev Psychopathol 2011;23(1):7-28.

63. Guerra S, Martinez FD. Asthma genetics: from linear to multifactoril approaches. Annu Rev Med 2008;59:327-41
64. Alkon A, Boyce WT, Davis NV, et al. Developmental changes in autonomic nervous system resting and reactivity measures in Latino children from 6 to 60 months of age. J Dev Behav Pediatr 2011;32(9):668-77.
65. Manuck SB, Craig AE, Flory JD, et al. Reported early family environment covaries with menarcheal age as a function of polymorphic variation in estrogen receptor-alpha. Dev Psychopathol 2011;23(1):69-83.
66. Beach SR, Brody GH, Lei MK, et al. Is serotonin transporter genotype associated with epigenetic susceptibility or vulnerability? Examination of the impact of socioeconomic status risk on African American youth. Dev Psychopathol 2014;26(2):289-304.
67. Strunk T, Jamieson SE, Burgner D. Genetic and epigenetic susceptibility to early life infection. Curr Opin Infect Dis 2013;26(3):241-7.
68. Fox SE, Levitt P, Nelson CA. How the timing and quality of early experiences influence the development of brain architecture. Child Dev 2010;81(1):28-40.
69. Nelson CA, Fox NA, Zeanah C. Romania's abandoned children. Cambridge, MA: Harvard University Press; 2014.
70. Almas A, Degnan K, Radulescu A, et al. The effects of early intervention and the moderating effects of brain activity on institutionalized children's social skills at age 8. Proc Nat Acad Sci U S A 2012;109(Suppl 2):17228-31.
71. Takesian AE, Hensch TK. Balancing plasticity/stability across brain development. Prog Brain Res 2013;207:3-34.
72. Gervain J, Vines BW, Chen LM, et al. Valproate reopens critical-period learning of absolute pitch. Front Syst Neurosci 2013;7:102.
73. Gogolla N, Leblanc JJ, Quast KB, et al. Common circuit defect of excitatory-inhibitory balance in mouse models of autism. J Neurodev Disord 2009;1(2):172-81.
74. Fagiolini M, Jensen CL, Champagne FA. Epigenetic influences on brain development and plasticity. Curr Opin Neurobiol 2009;19(2):207-12.
75. National Research Council, Committee on a Framework for Development a New Taxonomy of Disease. Toward precision medicine: building a knowledge network for biomedical research and a new taxonomy of disease. Washington, DC: National Academies Press; 2011.
76. Brookes E, Shi Y. Diverse epigenetic mechanisms of human disease. Annu Rev Genet 2014;48:237-68.
77. Galea S, Riddle M, Kaplan GA. Causal thinking and complex system approaches in epidemiology. Int J Epidemiol 2010;39(1):97-106.

W. Thomas Boyce, MD, is Distinguished Professor of Pediatrics and Psychiatry and heads the Division of Developmental Medicine in the Department of Pediatrics, University of California, San Francisco. Previously, he was Professor of Pediatrics and the Sunny Hill Health Centre-BC Leadership Chair in Child Development at the University of British Columbia, in the Human Early Learning Partnership, and at the Child and Family Research Institute of BC Children's Hospital. Dr. Boyce has served as a member of Harvard University's National Scientific Council on the Developing Child, UC Berkeley's Institute of Human Development, and as a founding co-Director of the Robert Wood Johnson Foundation Health & Society Scholars Program at Berkeley and UCSF. He co-directs the Child and Brain Development Program for the Canadian Institute for Advanced Research and serves on the Board on Children, Youth, and Families of the National Academy of Sciences. He was elected in 2011 to the Institute of Medicine.

The Brain on Stress: How Behavior and the Social Environment "Get Under the Skin"

Bruce S. McEwen

Abstract

The brain is the central organ of stress and adaptation. Brain circuits are remodeled by stress so as to change the ability to self-regulate anxiety and mood and to perform working and episodic memory, as well as executive function and decisionmaking. The brain regulates the body via the neuroendocrine, autonomic, immune, and metabolic systems, and the mediators of these systems and those within the brain and other organs activate epigenetic programs that alter expression of genetic information so as to change cellular and organ function. While the initial active response to stressors promotes adaptation ("allostasis"), there can be cumulative change (e.g., body fat, hypertension) from chronic stress and a resulting unhealthy lifestyle ("allostatic load"), which may lead to disease, e.g. diabetes, cardiovascular disease ("allostatic overload"). Besides early life experiences, the most potent of stressors are those arising from the social and physical environment, and these can affect both brain and body. Gradients of socioeconomic status generally reflect the cumulative burden of coping with limited resources, toxic environments, and negative life events, as well as health-damaging behaviors that result in chronic activation of physiological systems and lead to allostatic load and overload. Can we intervene to change this progression? After describing the new view of epigenetics that negates the old notion that "biology is destiny" and opens new avenues for collaboration between the biological and the behavioral and social sciences, I summarize some of the underlying cellular, molecular, and neuroendocrine mechanisms of stress effects on brain and body. I then discuss integrative or "top down" approaches involving behavioral interventions at the individual level that take advantage of the increasing ability to reactivate plasticity in the brain. At the societal level, virtually all policies of government and the private sector affect health directly or indirectly and must be redirected to allow people to make choices that improve their chances for a healthy life.

Introduction

Stress is a condition of the mind and a factor in the expression of disease that differs among individuals and reflects not only major life events, but also the conflicts and pressures of daily life that elevate physiological systems so as to cause a cumulative chronic stress burden on brain and body. This burden reflects not only the impact of life experiences, but also of genetic variations; individual health-related behaviors such as diet, exercise, sleep patterns, and substance abuse; and epigenetic modifications in development and throughout life that set life-long patterns of behavior and physiological reactivity through both biological embedding and cumulative change. Epigenetics

is the now popular way to describe gene x environment interactions via molecular mechanisms that do not change the genetic code but rather activate, repress, and modulate expression of the code.[1] Indeed, epigenetics denies the notion that "biology is destiny" and opens new opportunities for collaboration between the biological and behavioral and social sciences.

Acting epigenetically, hormones associated with stress protect the body in the short term and promote adaptation (allostasis), but in the long run, the burden of chronic stress causes changes in the brain and body that lead to cumulative change— such as accumulation of body fat (allostatic load)—or disease, such as diabetes or cardiovascular disease (allostatic overload). Brain circuits are plastic and appear to be continuously remodeled by stress, as well as by other experiences, so as to change the balance between anxiety and self-regulatory behaviors, including mood control and impulsivity, memory, and decisionmaking. Such changes may have adaptive value in danger, but their persistence and lack of reversibility in brains that are not resilient can be maladaptive.

Besides developmental influences associated with parent-infant interactions and the quality of early experiences, the most potent of stressors that influence adult life are those arising from the family, neighborhood, workplace, and exposure to local, national, and international events in the media that can affect both brain and body health and progression toward a variety of diseases. Social ordering in human society is associated with gradients of disease, with an increasing frequency of mortality and morbidity along a gradient of decreasing income and education (socioeconomic status, SES).[a] Although the causes of these gradients of health are very complex, they likely reflect, with increasing frequency going down the SES ladder, the cumulative burden of coping with limited resources, toxic and otherwise stressful living environments, and negative life events, as well as differences in health-related behaviors (aka "lifestyle") that result in chronic activation of physiological systems involved in adaptation leading to allostatic overload.[2]

Thus, the behavioral and social sciences have increasingly important roles in the evolution of our knowledge about brain-body interactions over the life course in the area known now as "social neuroscience." Because of the multiple levels of interaction from the physical and social environment down to individual health behaviors and the impact of all of these upon the physiology of the body and internal workings of the brain, interventions must occur at multiple levels. This topic will be discussed in relation to our increasing knowledge of the potential for brain plasticity as influenced by "top down," that is, integrative, behavioral, and societal interventions facilitated by activities like exercise that increase the potential for plasticity. This is particularly important in discussions about understanding the origins of diseases and improving health and health care in view of the "lifecourse health development" perspective that is now superseding the "biopsychosocial" and "germs, genes, and biomedical" models of the past.[3] Indeed, what happens at each stage of development, with particular potency early in life, has influences upon the trajectory that brain and body development take as the life course unfolds.

[a] See the University of California, San Francisco's MacArthur Research Network on Socioeconomic Status and Health. Available at http://www macses.ucsf.edu/.

Determinants of Physical and Mental Health

There are many aspects of life experiences that influence physical and mental health, and the brain is central to all of them (Figure 1).[4] Social stressors include trauma and abuse, major life events, and the daily experiences of work, family, neighborhood, and ongoing events in one's city, State, nation, and world. The brain processes all of this and determines the behavioral and physiological responses. Behavioral responses include quality and quantity of sleep and health-damaging behaviors, such as eating too much, smoking and substance abuse, including alcohol, as well as health promoting behaviors, such as regular physical activity and social integration and social support. Physiological responses that are normally adaptive ("allostasis" – see below) can lead to pathophysiology ("allostatic load and overload") when overused or dysregulated, as will be discussed below. Health behaviors feed into the network of allostasis and can lead to allostatic load (AL) and overload, sleep deprivation being a good example.[5] Socioeconomic status (SES), including both education and income, are reflected in AL and overload and gradients of disease.[b] Again, the brain with its influence on the rest of the body is key because subjective SES, reflecting perceived social position, is reflected in many aspects of physical and mental health.[6,7] The gradient of income in many societies is reflected not only in the frequency of diseases, but also in abnormal behavior, including depression, aggression, and violence and the degree of incarceration.[8,9] Next, we will consider in more detail the concepts of allostasis and allostatic load/overload and how experiences over the life course interact with the brain and body to cause disease.

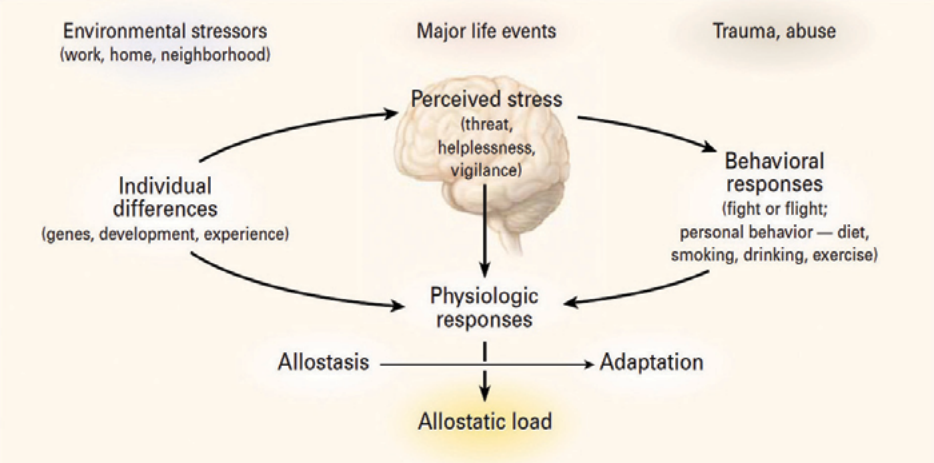


Figure 1. Central role of the brain in the protective and damaging effects of the mediators of stress and adaptation

Source: McEwen BS. Protective and damaging effects of stress mediators. NEJM 1998;338(3):171-9. Used with permission.

[b] See Reaching for a Healthier Life; Facts on Socioeconomic Status and Health in the U.S. Available at http://www macses.ucsf.edu/downloads/Reaching_for_a_Healthier_Life.pdf.

Cumulative Change: Stress, Lifestyle, and Allostatic Load/Overload

The brain regulates neuroendocrine, autonomic, immune, and metabolic systems and the mediators of these systems interact in a non-linear manner.[10] The concepts of allostasis and allostatic load/overload concern the protective, as well as potentially damaging, effects of the mediators of stress and adaptation, and they reflect a life course perspective that recognizes the power of the social environment and the health behaviors adopted by individuals. These concepts also include genetic contributions, and they recognize the central role of the brain and the importance of reciprocal brain-body interactions.[4,11]

The distinction between AL and overload is based on the severity of the outcome—allostatic load refers to cumulative change in biomarkers, whereas allostatic overload signifies cumulative change that has pathophysiological consequences.[2] In the natural world, bears putting on fat for the winter develop an AL in terms of body fat that they burn off during hibernation; a bear in a zoo, overfed and physically inactive, can develop an allostatic overload that involves diabetes and cardiovascular disease. Migrating salmon present another example of AL in the natural world; they die after spawning due, in part, to a massive over-secretion of glucocorticoids.[2]

Measuring allostasis and allostatic load/overload requires biomarkers that tap into multiple interactive systems and look at the brain, as well as systemic physiology.[12–14] These biomarkers tap into measures of the multiple interacting mediators that affect many body systems concurrently, including measures of blood pressure, metabolic parameters (glucose, insulin, lipid profiles, and waist circumference), markers of inflammation (interleukin-6, C-reactive protein, and fibrinogen), heart rate variability, sympathetic nervous system activity (12-hour urinary norepinephrine and epinephrine), and hypothalamic-pituitary-adrenal axis activity (diurnal salivary free cortisol).[14] Dehydroepiandrosterone (DHEA) and insulin-like growth factor 1 (IGF-1) have been used as positive markers of health. Choice of markers is limited by their cost and accessibility in studies involving large numbers of subjects. Telomeres and telomerase have been added as another endpoint of cumulative effect.[15,16]

Biomarkers for allostasis and allostatic load/overload fall into different classes: primary, secondary, tertiary.[17] Primary mediators include cortisol, sympathetic and parasympathetic activity, pro- and anti-inflammatory cytokines, metabolic hormones, and neurotransmitters and neuromodulators in the nervous system. Secondary mediators are those that reflect the cumulative actions of the primary mediators in a tissue/organ-specific manner, often reflecting the actions of more than one primary mediator, such as those described above: e.g. those that reflect abnormal metabolism and risk for cardiovascular disease, such as waste-hip ratio, blood pressure, glycosylated hemoglobin, cholesterol/high-density lipoprotein (HDL) ratio, and HDL cholesterol, as well as telomere length and telomerase activity.

In the realm of neuroimaging, secondary outcomes include functional activation of a set of brain regions that appear to define subtypes of anxiety disorder[18] and hypo- or hyperactivation of the insula region of the brain that defines antidepressant responsiveness.[19] In the case of telomeres and telomerase, which are secondary outcomes, oxidative stress and inflammatory processes are primary mediators that appear to cause these changes.[15]

Tertiary outcomes are actual diseases or disorders that are the result of AL that is predicted from the extreme values of the secondary outcomes and of the primary mediators.[17] Cardiovascular disease, decreased physical capacity, and severe cognitive decline have been used as outcomes in successful aging studies.[17] However, as noted by McEwen and Seeman,[17] "…cognitive function could be classified as a secondary outcome, although Alzheimer's disease or vascular dementia would be a tertiary outcome when there is clearly a serious and permanent disease. By the same token, cancer would be a tertiary outcome, whereas the common cold would be a secondary outcome and an indirect measure, in part, of immune system efficacy."

Both simple and more complex mathematical approaches have been used to analyze the biomarker data to create an "allostatic load battery."[12] Regardless of the analysis used, the AL battery has revealed relationships with behavioral and other outcomes, including predictions of outcomes over time.[13,14,20] In the MacArthur Successful Aging Study, for example, higher AL scores predicted increased mortality 7 years later.[21] Higher education was consistently associated with lower AL scores, while African Americans, in general, have higher AL scores and a flatter gradient across education. Neighborhood poverty has been found to be associated with higher AL scores among the residents.[22] As might be expected, higher AL is associated with increased social conflict, whereas positive social support and social connectedness is linked to lower AL scores.[13,23] Allostatic load concepts have also been applied to begin to elucidate the systemic consequences of psychiatric disorders such as bipolar illness, major depression, and schizophrenia[24,25]

In a recent review of the history of AL in relation to health disparities, Beckie concludes: "There is…. empirical substantiation for the relationships between AL and socioeconomic status, social relationships, workplace, lifestyle, race/ethnicity, gender, stress exposure, and genetic factors. The literature also demonstrated associations between AL and physical and mental health and all-cause mortality."[26] And, Beckie adds: "Targeting the antecedents of AL during key developmental periods is essential for improving public health. Priorities for future research include conducting prospective longitudinal studies, examining a broad range of antecedent allostatic challenges, and collecting reliable measures of multisystem dysregulation explicitly designed to assess AL, at multiple time points, in population-representative samples." Beckie, however, notes: "The results (of her systematic review) revealed considerable heterogeneity in the operationalization of AL and the measurement of AL biomarkers, making interpretations and comparisons across studies challenging, and therefore, future work should standardize the allostatic load battery along the lines used by Seeman and colleagues[14] so as to make easier comparisons across studies."[26]

Brain as Central Organ of Stress and Adaptation

The brain is the key organ of the response to stress because it determines what is threatening and therefore stressful, as well as the physiological and behavioral responses that can be either protective or damaging.[4] The brain is a target of stress, and the hippocampus was the first brain region besides the hypothalamus to be recognized as a target of glucocorticoids.[27] Stress and stress hormones produce both adaptive and maladaptive effects on the hippocampus throughout the life course. Early life events influence lifelong patterns of emotionality and stress responsiveness and alter the rate of brain and body aging. The amygdala is an important target of stress and is important in fear

and strong emotions, and the prefrontal cortex is involved in attention, executive function, and working memory.[28] The hippocampus and amygdala show an opposite response to repeated stress, involving remodeling of dendrites, whereas the prefrontal cortex shows both types of responses.[29] Hippocampal and medial prefrontal cortical neurons become shorter and less branched, and dentate gyrus neurogenesis is suppressed by repeated stress, whereas amygdala and orbitoprefrontal cortical neurons show signs of hypertrophy after repeated stress.[30,31] Repeated stress promotes impairment of hippocampal-dependent memory and enhances fear and aggression, as well as impairing attention set shifting, a form of executive function that indicates cognitive flexibility.[31] In the human brain, magnetic resonance imaging (MRI) has shown amygdala enlargement and overactivity and hippocampal and prefrontal cortical shrinkage in a number of mood disorders.[32] Hippocampal atrophy is also reported in Cushing's disease, post-traumatic stress disorder, recurrent depressive illness, and borderline personality disorder.[29,33] Knowledge of underlying anatomical changes and the mechanism of neuronal shrinkage or growth may help in developing treatment strategies to either reverse or prevent them, as will be discussed later in this chapter.

Early Life Adverse Experiences and Epigenetics

Early life events related to maternal care in animals, as well as parental care in humans, play a powerful role in later mental and physical health,[34,35] as demonstrated by the adverse childhood experiences (ACE) studies[36] and recent work that will be noted below. Animal models have contributed enormously to our understanding of how the brain and body are affected, starting with the "neonatal handling" studies of Levine and Denenberg[37] and the recent, elegant work of Meaney and Syzf.[38] Epigenetic, transgenerational effects transmitted by maternal care are central to these findings. Besides the amount of maternal care, the consistency over time of that care and the exposure to novelty are also very important, not only in rodents,[39,40] but also in monkey models.[41] Prenatal stress impairs hippocampal development in rats, as does stress in adolescence.[42] Abusive maternal care in rodents and the surprising attachment shown by infant rats to their abusive mothers appear to involve an immature amygdala,[43] activation of which by glucocorticoids causes an aversive conditioning response to emerge. Maternal anxiety in the variable foraging demand (VFD) model in rhesus monkeys leads to chronic anxiety in the offspring, as well as signs of metabolic syndrome.[44,45]

In studies of ACE in human populations, there are reports of increased inflammatory tone, not only in children, but also in young adults related to early life abuse, which includes family instability, use of chronic harsh language, and physical and sexual abuse.[46–48] Chaos in the home is associated with development of poor self-regulatory behaviors, as well as obesity.[49] It should be noted that the ACE study was carried out in a middle-income population,[50] highlighting that poverty is not the only source of early life stressors.

Nevertheless, low SES does increase the likelihood of stressors in the home and neighborhood, including toxic chemical agents such as lead and air pollution.[51] Without a determination of exact causes, it has been reported that low SES children are more likely than other children to be deficient in language skills and self-regulatory behaviors and also in certain types of memory that are likely to be reflections of impaired development of perisylvian gyrus language centers, prefrontal cortical

systems, and temporal lobe memory systems.[52,53] Low SES and family poverty are reported to correlate with smaller hippocampal volumes,[54] overall smaller gray matter volume,[55] and impaired development in children of reduced prefrontal control of amygdala activity resulting in impaired self-regulatory behavior.[29,31,56] Neglect is associated with impaired white matter development and integrity.[57] Lower subjective SES, an important index of objective SES, is associated with reduction in prefrontal cortical gray matter.[58] Moreover, individuals reared in a lower SES environment tend to show greater amygdala reactivity to angry and sad faces,[59] which, as noted above, may be a predisposing factor for early cardiovascular disease that is known to be more prevalent at lower SES levels.[60] Finally, depression is often associated with low SES, and children of depressed mothers, followed longitudinally, have shown increased amygdala volume, while hippocampal volume was not affected.[61]

Yet, on the positive side, there are the "reactive or context-sensitive alleles"[62] that, in nurturing environments, lead to beneficial outcomes and even better outcomes compared to less reactive alleles, even though those same alleles can enhance adverse outcomes in a stressful early life environment.[63-65] Regarding adverse outcomes and "good and bad environments," allostatic processes are adjusted via epigenetic influences to optimize the individual's adaptation to, and resulting fitness for, a particular environment, whether more or less threatening or nurturing.[66] Yet, there are "trade-offs" in terms of physical and mental health that, on the one hand, may increase the likelihood of passing on one's genes by improving coping with adversity and enhancing mental health and overall reproductive success, but on the other hand, may impair later health, e.g., by eating of "comfort foods" (see for example Jackson, et al).[67] Moreover, when an individual faces a new challenge, there is the question of resilience in terms of the ability to show experience-related adaptation, for example, when an individual from a safe environment is placed into a dangerous one or vice versa. This brings up the question of plasticity, particularly in brain architecture that is so fundamental to brain and body health, and there is both old and new evidence that glucocorticoids, often thought of in a negative sense in relation to stress effects, play an important role in the ability of the brain to adapt to new challenges and possibly also to remediate deficits associated with stress over the life course.

Role of Glucocorticoids and Other Mediators in Brain Plasticity

The discovery of receptors for glucocorticoids in the hippocampus has led to many investigations in animal models and translation to the human brain using modern imaging methods that show the degree to which the brain, on the one hand, may be damaged by excessive glucocorticoids, but on the other hand, the beneficial role that they play in adaptive plasticity together with other mediators.[30] Glucocorticoids thus provide insights into brain plasticity, as well as the more negative side related to AL and overload.

The most striking findings from animal models have identified structural plasticity in the hippocampus, consisting of ongoing neurogenesis in the dentate gyrus[68] and remodeling of dendrites and synapses in the major neurons of Ammon's horn.[17] The mediators of this plasticity include excitatory amino acids[69] and glucocorticoids, along with a growing list of other mediators, such as oxytocin, corticotrophin releasing factor, brain derived neurotrophic factor (BDNF), lipocalin-2, and tissue plasminogen activator (tPA).[30,33] Moreover, glucocorticoid actions involve both genomic and non-genomic mechanisms that implicate mineralocorticoid, as well as glucocorticoid, receptors

and their translocation to mitochondria, as well as to cell nuclei, and an as-yet unidentified G-protein coupled membrane receptor related to endocannabinoid production.[70–72]

Studies of the human hippocampus have demonstrated shrinkage of the hippocampus not only in mild cognitive impairment and Alzheimer's,[73] but also in Type 2 diabetes,[74] prolonged major depression,[75] Cushing's disease,[76] and posttraumatic stress disorder (PTSD).[77] Moreover, in non-disease conditions, such as chronic stress,[78] chronic inflammation,[79] lack of physical activity,[80] and jet lag,[81] smaller hippocampal or temporal lobe volumes have been reported.

These changes may not be due to neuron loss but rather to volume reduction in dentate gyrus due to inhibited neuronal replacement, as well as dendritic shrinkage and glial cell loss. Autopsy studies on depression-suicide have indicated loss of glial cells and smaller neuron soma size,[82] which is indicative of a smaller dendritic tree. With regard to type 2 diabetes, it should be emphasized that the hippocampus has receptors for, and the ability to take up and respond to, insulin, ghrelin, insulin-like growth factor-1 (IGF1), and leptin, and that IGF-1 mediates exercise-induced neurogenesis.[30] Thus, besides its response to glucocorticoids, the hippocampus is an important target of metabolic hormones that have a variety of adaptive actions in the healthy brain which is perturbed in metabolic disorders, such as diabetes.[30]

There is a positive side to glucocorticoid action that can be harnessed to promote plasticity and, through that, adaptation and resilience. Glucocorticoid actions involve both genomic and non-genomic mechanisms in mitochondria, as well as in synaptic terminals and dendrites and spines, and interactions with excitatory amino acids and endocannabinoids.[69,71] These actions mediate adaptive neuronal functions, including the ongoing turnover of spine synapses resulting from the ultradian fluctuation of glucocorticoids[83] that is involved in the diurnal fluctuations of behavior and is important for efficient motor learning.[84]

These glucocorticoid actions are likely to be involved in the remarkable reversal of amblyopia ("lazy eye" resulting from monocular deprivation during development) that was first shown to be reversed by patterned light exposure in adulthood facilitated by fluoxetine.[85] This is because, after showing that fluoxetine works, caloric restriction every other day was also shown to be effective.[86] Then, because caloric restriction elevates glucocorticoids, putting glucocorticoids in the drinking water every other day during visual stimulation was able to mimic the effects of both fluoxetine and food restriction.[86] Thus, glucocorticoids may play an important role in the re-establishment of a new window of plasticity.[87] The ultradian fluctuation of glucocorticoids has been shown to be essential for the ongoing turnover of spines on dendrites where excitatory synapses are formed,[83] and this turnover plays a role in motor learning and possibly other adaptive functions.[84]

A key mechanism in reactivating plasticity involves reducing inhibitory neuronal activity by GABAergic basket neurons.[88] One application of reactivation of plasticity is for depressive illness, which is more prevalent in individuals who have had adverse early life experiences.[50] BDNF may be a key feature of the depressive state, and elevation of BDNF by diverse treatments ranging from antidepressant drugs to regular physical activity may be a key feature of treatment.[89] Yet, there are other potential applications, such as the recently reported ability of fluoxetine to enhance recovery

from stroke.[90] In both examples, the reactivation of plasticity is accompanied by a behavioral intervention involving behavioral therapy for depression and physical therapy for stroke, since the reactivated plasticity must be directed towards a desired outcome by what I shall refer to as a "top down" or "integrative" intervention. This is to be distinguished from a "bottom up" intervention, such as the use of a pharmaceutical agent targeted at a molecular pathway

Interventions

What can be done to remediate the effects of chronic stress over the life course at both individual and societal levels? For the individual, the complexity of interacting, non-linear and biphasic actions of the mediators of stress and adaptation, as described above, emphasizes behavioral, or "top-down," interventions (i.e., interventions that involve integrated central nervous system [CNS] activity) that include cognitive-behavioral therapy, physical activity, and programs such as the Experience Corps that promote social support and integration and meaning and purpose in life.[12,30] In contrast, pharmacological agents, which are useful in many circumstances to redress chemical and molecular imbalances, nevertheless run the risk of dysregulating other adaptive pathways, i.e., no pharmaceutical is without side effects. It should also be noted that many interventions that are intended to promote plasticity and slow decline with age, such as physical activity and positive social interactions that give meaning and purpose, are also useful for promoting "positive health" and "eudamonia" [91-93] independently of any notable disorder and within the range of normal behavior and physiology.

A powerful "top down" therapy (i.e., an activity, usually voluntary, involving activation of integrated nervous system activity, as opposed to pharmacological therapy which has a more limited target) is regular physical activity, which has actions that improve prefrontal and parietal cortex blood flow and enhance executive function.[94] Moreover, regular physical activity, consisting of walking an hour a day, 5 out of 7 days a week, increases hippocampal volume in previously sedentary adults.[95] This finding complements work showing that fit individuals have larger hippocampal volumes than sedentary adults of the same age range.[80] It is also well known that regular physical activity is an effective antidepressant and protects against cardiovascular disease, diabetes, and dementia.[96-100] Moreover, intensive learning has also been shown to increase the volume of the human hippocampus.[101]

Social integration and support and finding meaning and purpose in life are known to be protective against AL and overload[102] and dementia,[103] and programs such as the Experience Corps that promote these along with increased physical activity have been shown to slow the decline of physical and mental health and to improve prefrontal cortical blood flow in a similar manner to regular physical activity.[104,105]

Depression and anxiety disorders are examples of a loss of resilience, in the sense that changes in brain circuitry and function—caused by the stressors that precipitate the disorder—become "locked" in a particular state and thus need external intervention. Indeed, prolonged depression is associated with shrinkage of the hippocampus[75,106] and prefrontal cortex.[107] While there appears to be no neuronal loss, there is evidence for glial cell loss and smaller neuronal cell nuclei,[82,108] which

is consistent with a shrinking of the dendritic tree described above after chronic stress. Indeed, a few studies indicate that pharmacological treatment may reverse the decreased hippocampal volume in unipolar[109] and bipolar[110] depression, but the possible influence of concurrent cognitive-behavioral therapy in these studies is unclear. And, yet, from the discussion of reversal of amblyopia (above), it is possible that a combination of a pharmaceutical or behavioral (e.g., exercise) intervention that opens up a "window of plasticity" might improve the efficacy of behavioral therapies.

In this connection, it is important to reiterate that successful behavioral therapy, which is tailored to individual needs, can produce volumetric changes in both prefrontal cortex in the case of chronic fatigue,[111] and in amygdala, in the case of chronic anxiety[112] as measured in the same subjects longitudinally. This reinforces the notion that plasticity-facilitating treatments should be given within the framework of a positive behavioral or physical therapy intervention. On the other hand, negative experiences during the window of enhanced plasticity may have undesirable consequences, such as a person going back into a bad family environment that may have precipitated anxiety or depression in the first place.[87] In that connection, it should be noted that BDNF, a plasticity enhancing class of molecules, also has the ability to promote pathophysiology, as in seizures.[113–115]

At the societal level, the most important top-down interventions are the policies of government and the private sector that not only improve education but also allow people to make choices that improve their chances for a healthy life.[12] This point was made by the Acheson report of the British Government in 1998,[116] which recognized that no public policy of virtually any kind should be enacted without considering the implications for the health of all citizens. Thus, basic education, housing, taxation, setting of a minimum wage, and addressing occupational health and safety and environmental pollution regulations are all likely to affect health via a myriad of mechanisms. At the same time, providing higher quality food and making it affordable and accessible in poor, as well as in affluent neighborhoods will be necessary for people to eat better, providing they also learn what types of food to eat.[117] Likewise, making neighborhoods safer and more congenial and supportive can improve opportunities for positive social interactions and increased recreational physical activity.[118,119] However, government policies are not the only way to reduce allostatic load. For example, businesses that encourage healthy lifestyle practices among their employees are likely to gain reduced health insurance costs and possibly a more loyal workforce.[120-122]

Conclusion

Biomedical science has progressed from the concept of identifying and treating single causes of disease, as in the germ theory, to the recognition of multiple risk factors, as in the biopsychosocial model, and now to "lifecourse health development" that recognizes multiple contributing factors from genes and experiences over the life course that epigenetically alter trajectories of health and disease.[3] At the same time, the behavioral and social sciences are recognizing that "biology is not destiny," and that social and behavioral influences continually affect the body through the brain. Together with recognition of the central role of the brain and behavior (Figure 1), the emerging science of "epigenetics" recognizes the continuing role of the physical and social environments, and of behavior, in getting "under the skin" to shape expression of inherited characteristics in

a continuous and evolving manner over an individual's lifespan. The complexities and degrees of freedom made possible by the many ways that "epigenetics" affects expression of the genetic blueprint make it essential that the social and behavioral sciences work closely with the biological sciences to achieve better understanding of brain and body function for the good of our own species and all living creatures on this earth.

Author's Affiliation

Bruce S. McEwen, PhD, is the Alfred E. Mirsky Professor of Neuroscience and runs the Harold and Margaret Milliken Hatch Laboratory of Neuroendocrinology at Rockefeller University.

Address correspondence to: Bruce S. McEwen, PhD, Laboratory of Neuroendocrinology, The Rockefeller University, 1230 York Avenue, New York, NY 10065; email: mcewen@rockefeller.edu

References

1. Mehler MF. Epigenetic principles and mechanisms underlying nervous system functions in health and disease. Prog Neurobiol 2008; 86:305-41.
2. McEwen BS, Wingfield JC. The concept of allostasis in biology and biomedicine. Horm & Behav 2003;43:2-15.
3. Halfon N, Larson K, Lu M, et al. Lifecourse health development: past, present and future. Matern Child Health J 2014; 18:344-65.
4. McEwen BS. Protective and damaging effects of stress mediators. N Engl J.Med 1998; 338:171-9.
5. Spiegel K, Leproult R, Van Cauter E. Impact of sleep debt on metabolic and endocrine function. Lancet 1999; 354:1435-9.
6. Adler NE, Boyce TW, Chesney MA, et al. Socioeconomic inequalities in health. JAMA 1993; 269:3140-5.
7. Goodman E, Huang B, Schafer-Kalkhoff T, et al. Perceived socioeconomic status: a new type of identity that influences adolescents' self-rated health. J Adolesc Health 2007; 41:479-87.
8. Marmot MG, Wilkinson RG. Social determinants of health. Oxford, New York: Oxford University Press; 2006.
9. Wilkinson RG, Pickett K. The spirit level: Why greater equality makes societies stronger. New York: Bloomsbury Press; 2010.
10. McEwen BS. Protective and damaging effects of stress mediators: central role of the brain. Dialogues Clin Neurosci 2006; 8(4):367-81.
11. McEwen BS, Stellar E. Stress and the individual. Mechanisms leading to disease. Arch Intern Med 1993; 153:2093-101.
12. Juster RP, McEwen BS, Lupien SJ. Allostatic load biomarkers of chronic stress and impact on health and cognition. Neurosci Biobehav Rev 2010; 35:2-16.
13. Seeman T, Epel E, Gruenewald T, et al. Socio-economic differentials in peripheral biology: cumulative allostatic load. Ann NY Acad Sci 2010; 1186:223-39.
14. Seeman T, Gruenewald T, Karlamangla A, et al. Modeling multisystem biological risk in young adults: the Coronary Artery Risk Development in Young Adults Study. Am J Hum Biol 2010; 22:463-72.
15. Blackburn EH, Epel ES. Telomeres and adversity: too toxic to ignore. Nature 2012; 490:169-71.
16. Epel ES, Blackburn EH, Lin J, et al. Accelerated telomere shortening in response to life stress. Proc Natl Acad Sci 2004; 101:17312-15.
17. McEwen BS, Seeman T. Protective and damaging effects of mediators of stress: elaborating and testing the concepts of allostasis and allostatic load. Ann NY Acad Sci 1999; 896:30-47.
18. Etkin A, Wager TD. Functional neuroimaging of anxiety: a meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia. Am J Psychiatry 2007; 164:1476-88.
19. McGrath CL, Kelley ME, Holtzheimer PE, et al. Toward a neuroimaging treatment selection biomarker for major depressive disorder. JAMA Psychiatry 2013; 70(8):821-9.
20. Seeman TE, Singer BH, Rowe JW, et al. Price of adaptation—allostatic load and its health consequences: MacArthur studies of successful aging. Arch Intern Med 1997; 157:2259-68.
21. Seeman TE, McEwen BS, Singer BH, et al. Increase in urinary cortisol excretion and memory declines:

MacArthur studies of successful aging. J Clin Endocrinol Metab 1997; 82:2458-65.
22. Schulz AJ, Mentz G, Lachance L, et al. Associations between socioeconomic status and allostatic load: effects of neighborhood poverty and tests of mediating pathways. Am J Public Health 2012; 102:1706-14.
23. Seeman TE, Singer BH, Ryff CD, et al. Social relationships, gender, and allostatic load across two age cohorts. Psychosom Med 2002; 64:395-406.
24. Bizik G, Picard M, Nijjar R. Allostatic load as a tool for monitoring physiological dysregulations and comorbidities in patients with severe mental illnesses. Harvard Rev Psychiatry 2013; 21:296-313.
25. Kapczinski F, Vieta E, Andreazza AC, et al. Allostatic load in bipolar disorder: implications for pathophysiology and treatment. Neurosci Biobehav Rev 2008; 32:675-92.
26. Beckie TM. A systematic review of allostatic load, health, and health disparities. Biol Res Nurs 2012; 14:311-46.
27. McEwen BS. Stress and hippocampal plasticity. Annu Rev Neurosci 1999; 22:105-22.
28. Vyas A, Mitra R, Rao BSS, et al. Chronic stress induces contrasting patterns of dendritic remodeling in hippocampal and amygdaloid neurons. J Neurosci 2002; 22:6810-8.
29. McEwen BS, Gianaros PJ. Stress- and allostasis-induced brain plasticity. Annu Rev Med 2011; 62:431-45.
30. McEwen BS. Physiology and neurobiology of stress and adaptation: central role of the brain. Physiol Rev 2007; 87:873-904.
31. McEwen BS, Morrison JH. The Brain on Stress: Vulnerability and Plasticity of the Prefrontal Cortex over the Life Course. Neuron 2013; 79:16-29.
32. Drevets WC, Videen TO, Price JL, et al. A functional anatomical study of unipolar depression. J Neurosci 1992; 3628-41.
33. McEwen BS. Stress, sex, and neural adaptation to a changing environment: mechanisms of neuronal remodeling. Ann N Y Acad Sci 2010; 1204(Suppl):E38-59.
34. Juster RP, Bizik G, Picard M, et al. A transdisciplinary perspective of chronic stress in relation to psychopathology throughout life span development. Dev Psychopathol 2011; 23:725-76.
35. Lupien SJ, McEwen BS, Gunnar MR, et al. Effects of stress throughout the lifespan on the brain, behaviour and cognition. Nature Rev Neurosci 2009; 10:434-45.
36. Felitti VJ, Anda RF, Nordenberg D, et al. Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults. The adverse childhood experiences (ACE) study. Am J Prev Med 1998; 14:245-58.
37. Levine S, Haltmeyer G, Kara G, et al. Physiological and behavioral effects of infantile stimulation. Physiol Behav 1967; 2:55-59.
38. Meaney MJ, Szyf M. Environmental programming of stress responses through DNA methylation: life at the interface between a dynamic environment and a fixed genome. Dialogues Clin Neurosci 2005; 7:103-23
39. Akers KG, Yang Z, DelVecchio DP, et al. Social competitiveness and plasticity of neuroendocrine function in old age: influence of neonatal novelty exposure and maternal care reliability. PLoS ONE 2008; 3(7):e2840.
40. Tang AC, Akers KG, Reeb BC, et al. Programming social, cognitive, and neuroendocrine development by early exposure to novelty. Proc Natl Acad Sci USA 2006; 103:15716-21.
41. Parker KJ, Buckmaster CL, Sundlass K, et al. Maternal mediation, stress inoculation, and the development of neuroendocrine stress resistance in primates. Proc Natl Acad Sci USA 2006; 103:3000-5.
42. Isgor C, Kabbaj M, Akil H, et al. Delayed effects of chronic variable stress during peripubertal-juvenile period on hippocampal morphology and on cognitive and stress axis functions in rats. Hippocampus 2004; 14:636-48.
43. Moriceau S, Sullivan R. Maternal presence serves as a switch between learning fear and attraction in infancy. Nature Neurosci 2006; 8:1004-6.
44. Coplan JD, Smith ELP, Altemus M, et al. Variable foraging demand rearing: sustained elevations in cisternal cerebrospinal fluid corticotropin-releasing factor concentrations in adult primates. Biol Psychiat 2001; 50:200-4.
45. Kaufman D, Smith ELP, Gohil BC, et al. Early appearance of the metabolic syndrome in socially reared bonnet macaques. J Clin Endocrin Metab 2005; 90:404-8.
46. Danese A, McEwen BS. Adverse childhood experiences, allostasis, allostatic load, and age-related disease. Physiol Behav 2012; 106:29-39.
47. Danese A, Moffitt TE, Harrington H, et al. Adverse childhood experiences and adult risk factors for age-related disease: depression, inflammation, and clustering of metabolic risk markers. Arch Pediatr Adolesc Med 2009; 163:1135-43.
48. Miller GE, Chen E. Harsh family climate in early life presages the emergence of a proinflammatory phenotype in adolescence. Psychol Sci 2010; 21:848-56.
49. Evans GW, Gonnella C, Marcynyszyn LA, et al. The role of chaos in poverty and children's socioemotional adjustment. Psychol Sci 2005; 16:560-5.
50. Anda RF, Butchart A, Felitti VJ, et al. Building a framework for global surveillance of the public health implications of adverse childhood experiences. Am J Prev Med 2010; 39:93-8.
51. McEwen BS, Tucker P. Critical biological pathways for chronic psychosocial stress and research opportunities to advance the consideration of stress in chemical risk

assessment. Am J Public Health 2011; 101(Suppl 1):S131-9.
52. Farah MJ, Shera DM, Savage JH, et al. Childhood poverty: specific associations with neurocognitive development. Brain Res 2006; 1110: 66-74.
53. Hart B, Risley, TR. 1995. Meaningful differences in the everyday experience of young American children. Baltimore, MD: Brookes Publishing Company; 1995. 304 p.
54. Hanson JL, Chandra A, Wolfe BL, et al. Association between income and the hippocampus. PLoS One 2011; 6:e18712.
55. Hanson JL, Hair N, Shen DG, et al. Family poverty affects the rate of human infant brain growth. PLoS One 2013; 8:e80954.
56. Kim P, Evans GW, Angstadt M, et al. Effects of childhood poverty and chronic stress on emotion regulatory brain function in adulthood. Proc Natl Acad Sci USA 2013; 110:18442-7.
57. Hanson JL, Adluru N, Chung MK, et al. Early neglect is associated with alterations in white matter integrity and cognitive functioning. Child Dev 2013; 84:1566-78.
58. Gianaros PJ, Horenstein JA, Cohen S, et al. Perigenual anterior cingulate morphology covaries with perceived social standing. Soc Cogn Affect Neurosci 2007; 2:161-73.
59. Gianaros PJ, Horenstein JA, Hariri AR, et al. Potential neural embedding of parental social standing. Soc Cogn Affect Neurosci 2008; 3:91-6.
60. Adler NE, Boyce TW, Chesney MA, et al. Socioeconomic inequalities in health. JAMA 1993; 269:3140-5.
61. Lupien SJ, Parent S, Evans AC, et al. Larger amygdala but no change in hippocampal volume in 10-year-old children exposed to maternal depressive symptomatology since birth. Proc Natl Acad Sci USA 2011; 108:14324-9.
62. Obradovic J, Bush NR, Stamperdahl J, et al. Biological sensitivity to context: the interactive effects of stress reactivity and family adversity on socioemotional behavior and school readiness. Child Dev 2010; 81:270-89.
63. Boyce WT, Ellis BJ. Biological sensitivity to context: I. An evolutionary-developmental theory of the origins and functions of stress reactivity. Dev Psychopathol 2005; 17:271-301.
64. Caspi A, Sugden K, Moffitt TE, et al. Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. Science 2003; 301:386-9.
65. Suomi SJ. Risk, resilience, and gene x environment interactions in rhesus monkeys. Ann NY Acad Sci 2006; 1094:52-62.
66. Del Giudice M, Ellis BJ, Shirtcliff EA. The Adaptive Calibration Model of stress responsivity. Neurosci Biobehav Rev 2011; 35:1562-92.
67. Jackson JS, Knight KM, Rafferty JA. Race and unhealthy behaviors: chronic stress, the HPA axis, and physical and mental health disparities over the life course. Am J Public Health 2010; 100:933-9.
68. Cameron HA, Gould E. The control of neuronal birth and survival. In Shaw C (ed), Receptor Dynamics in Neural Development, pp. 141-57. Boca Raton, FL: CRC Press.
69. Popoli M, Yan Z, McEwen BS, et al. The stressed synapse: the impact of stress and glucocorticoids on glutamate transmission. Nat Rev Neurosci 2012; 13:22-37.
70. Du J, McEwen BS, Manji HK. Glucocorticoid receptors modulate mitochondrial function. Commun Integr Biol 2009; 2:1-3.
71. Hill MN, McEwen BS. Involvement of the endocannabinoid system in the neurobehavioural effects of stress and glucocorticoids. Prog Neuropsychopharmacol Biol Psychiatry 2010; 34:791-7.
72. Tasker JG, Di S, Malcher-Lopes R. Minireview: rapid glucocorticoid signaling via membrane-associated receptors. Endocrinology 2006; 147:5549-56.
73. de Leon MJ, George AE, Golomb J, et al. Frequency of hippocampus atrophy in normal elderly and Alzheimer's disease patients. Neurobiol Aging 1997; 18:1-11.
74. Gold SM, Dziobek I, Sweat V, et al. Hippocampal damage and memory impairments as possible early brain complications of type 2 diabetes. Diabetologia 2007; 50:711-9.
75. Sheline YI. Neuroimaging studies of mood disorder effects on the brain. Biol Psychiat 2003; 54:338-52.
76. Starkman MN, Giordani B, Gebrski SS, et al. Decrease in cortisol reverses human hippocampal atrophy following treatment of Cushing's disease. Biol Psychiat 1999; 46:1595-602.
77. Gurvits TV, Shenton ME, Hokama H, et al. Magnetic resonance imaging study of hippocampal volume in chronic, combat-related posttraumatic stress disorder. Biol Psychiatry 1996; 40:1091-9.
78. Gianaros PJ, Jennings JR, Sheu LK, et al. Prospective reports of chronic life stress predict decreased grey matter volume in the hippocampus. NeuroImage 2007; 35:795-803.
79. Marsland AL, Gianaros PJ, Abramowitch SM, et al. Interleukin-6 covaries inversely with hippocampal grey matter volume in middle-aged adults. Biol Psychiat 2008; 64:484-90.
80. Erickson KI, Prakash RS, Voss MW, et al. Aerobic fitness is associated with hippocampal volume in elderly humans. Hippocampus 2009; 19:1030-9.
81. Cho K. Chronic 'jet lag' produces temporal lobe atrophy and spatial cognitive deficits. Nature Neurosci 2001; 4:567-8.

82. Stockmeier CA, Mahajan GJ, Konick LC, et al. Cellular changes in the postmortem hippocampus in major depression. Biol Psychiat 2004; 56:640-50.
83. Liston C, Gan WB. Glucocorticoids are critical regulators of dendritic spine development and plasticity in vivo. Proc Natl Acad Sci USA 2011; 108:16074-9.
84. Liston C, Cichon JM, Jeanneteau F, et al. Circadian glucocorticoid oscillations promote learning-dependent synapse formation and maintenance. Nat Neurosci 2013; 16:698-705.
85. Vetencourt JFM, Sale A, Viegi A, et al. The antidepressant fluoxetine restores plasticity in the adult visual cortex. Science 2008; 320:385-8.
86. Spolidoro M, Baroncelli L, Putignano E, et al. Food restriction enhances visual cortex plasticity in adulthood. Nat Commun 2011; 2:320.
87. Castren E, Rantamaki T. The role of BDNF and its receptors in depression and antidepressant drug action: reactivation of developmental plasticity. Dev Neurobiol 2010; 70:289-97.
88. Bavelier D, Levi DM, Li RW, et al. Removing brakes on adult brain plasticity: from molecular to behavioral interventions. J Neurosci 2010; 30:14964-71.
89. Duman RS, Monteggia LM. A neurotrophic model for stress-related mood disorders. Biol Psychiat 2006; 59:1116-27.
90. Chollet F, Tardy J, Albucher JF, et al. Fluoxetine for motor recovery after acute ischaemic stroke (FLAME): a randomised placebo-controlled trial. Lancet Neurol 2011; 10:123-30.
91. Fredrickson BL, Grewen KM, Coffey KA, et al. A functional genomic perspective on human well-being. Proc Natl Acad Sci USA 2013; 110:13684-9.
92. Ryff CD, Singer B. The contours of positive human health. Psychol Inq 1998; 9:1-28.
93. Singer B, Friedman E, Seeman T, et al. Protective environments and health status: cross-talk between human and animal studies. Neurobiol Aging 2005; 26(Suppl):S113-8.
94. Colcombe SJ, Kramer AF, Erickson KI, et al. Cardiovascular fitness, cortical plasticity, and aging. Proc Natl Acad Sci USA 2004; 101:3316-21.
95. Erickson KI, Voss MW, Prakash RS, et al. Exercise training increases size of hippocampus and improves memory. Proc Natl Acad Sci USA 2011; 108:3017-22.
96. Babyak M, Blumenthal JA, Herman S, et al. Exercise treatment for major depression: maintenance of therapeutic benefit at 10 months. Psychosom Med 2000; 62:633-8.
97. Bonen A. Benefits of exercise for Type II diabetics: convergence of epidemiologic, physiologic, and molecular evidence. Can J Appl Physiol 1997; 20: 261-79.
98. Kahle EB, Zipf WB, Lamb DR, et al. Association between mild, routine exercise and improved insulin dynamics and glucose control in obese adolescents. Int J Sports Med 1996; 17:1-6.
99. Larson EB, Wang L, Bowen JD, et al. Exercise is associated with reduced risk for incident dementia among persons 65 years of age or older. Ann Intern Med 2006; 144:73-81.
100. Rovio S, Kareholt I, Helkala EL, et al. Leisure-time physical activity at midlife and the risk of dementia and Alzheimer's disease. Lancet Neurol 2005; 4:705-11.
101. Draganski B, Gaser C, Kempermann G, et al. Temporal and spatial dynamics of brain structure changes during extensive learning. J Neurosci 2006; 26:6314-7.
102. Seeman TE, Singer BH, Ryff CD, et al. Social relationships, gender, and allostatic load across two age cohorts. Psychosom Med 2002; 64:395-406.
103. Boyle PA, Buchman AS, Barnes LL, et al. Effect of a purpose in life on risk of incident Alzheimer disease and mild cognitive impairment in community-dwelling older persons. Arch Gen Psychiatry 2010; 67:304-10.
104. Carlson MC, Erickson KI, Kramer AF, et al. Evidence for neurocognitive plasticity in at-risk older adults: the experience corps program. J Gerontol A Biol Sci Med Sci 2009; 64:1275-82.
105. Fried LP, Carlson MC, Freedman M, et al. A social model for health promotion for an aging population: initial evidence on the experience corps model. J Urban Health Bull NY Acad Med 2004; 81:64-78.
106. Sheline YI. Hippocampal atrophy in major depression: a result of depression-induced neurotoxicity? Mol Psychiatry 1996; 1:298-9.
107. Drevets WC, Price JL, Simpson Jr JR, Todd RD, Reich T, et al. Subgenual prefrontal cortex abnormalities in mood disorders. Nature 1997; 386:824-27.
108. Rajkowska G. Postmortem studies in mood disorders indicate altered numbers of neurons and glial cells. Biol Psychiat 2000; 48:766-77.
109. Vythilingam M, Vermetten E, Anderson GM, et al. Hippocampal volume, memory, and cortisol status in major depressive disorder: effects of treatment. Biol Psychiat 2004; 56:101-12.
110. Moore GJ, Bebehuk JM, Wilds IB, et al. Lithium-induced increase in human brain grey matter. Lancet 2000; 356:1241-2.
111. de Lange FP, Koers A, Kalkman JS, et al. Increase in prefrontal cortical volume following cognitive behavioural therapy in patients with chronic fatigue syndrome. Brain 2008; 131:2172-80.
112. Holzel BK, Carmody J, Evans KC, et al. Stress reduction correlates with structural changes in the amygdala. Soc Cogn Affect Neurosci 2010; 5:11-7.
113. Heinrich C, Lahteinen S, Suzuki F, et al. Increase in BDNF-mediated TrkB signaling promotes epileptogenesis in a mouse model of mesial temporal lobe epilepsy. Neurobiol Dis 2011; 42:35-47.

114. Kokaia M, Ernfors P, Kokaia Z, et al. Suppressed epileptogenesis in BDNF mutant mice. Exp Neurol 1995; 133:215-24.
115. Scharfman HE. Hyperexcitability in combined entorhinal/hippocampal slices of adult rat after exposure to brain-derived neurotrophic factor. J Neurophysiol 1997; 78:1082-95.
116. Acheson SD. Independent inquiry into inequalities in health report. London: The Stationary Office; 1998.
117. Drewnowski A, Specter SE. Poverty and obesity: the role of energy density and energy costs. Am J Clin Nutr 2004; 79:6-16.
118. Kawachi I, Kennedy BP, Lochner K, et al. Social capital, income inequality, and mortality. Am J Public Health 1997; 87:1491-8.
119. Sampson RJ, Raudenbush SW, Earls F. Neighborhoods and violent crime: a multilevel study of collective effects. Science 1997; 277:918-24.
120. Aldana SG. Financial impact of health promotion programs: a comprehensive review of the literature. Am J Health Promot 2001; 15:296-320.
121. Pelletier KR. A review and analysis of the clinical- and cost-effectiveness studies of comprehensive health promotion and disease management programs at the worksite: 1998-2000 update. Am J Health Promot 2001; 16:107-15.
122. Whitmer RW, Pelletier KR, Anderson DR, et al. A wake-up call for corporate America. J Occup Environ Med 2003; 45:916-25.

Bruce S. McEwen, PhD, is the Alfred E. Mirsky Professor of Neuroscience and heads the Harold and Margaret Milliken Hatch Laboratory of Neuroendocrinology at Rockefeller University. As a neuroscientist and neuroendocrinologist, he studies environmentally-regulated, variable gene expression in the brain. His laboratory discovered adrenal steroid receptors in the hippocampus in 1968. He is a member of the National Academy of Sciences, the Institute of Medicine, the American Academy of Arts and Sciences, and the National Council on the Developing Child. Dr. McEwen served as President of the Society for Neuroscience in 1997-98.

Value of Investing in Health Care

The Science of Making Better Decisions About Health: Cost-Effectiveness and Cost-Benefit Analysis

Louise B. Russell

Abstract

Despite spending far more on medical care, Americans live shorter lives than the citizens of other high-income countries. The situation has been getting worse for at least three decades. This chapter describes the main scientific methods for guiding the allocation of resources to health: cost-effectiveness analysis (CEA) and cost-benefit analysis (CBA), sketches their methodological progress over the last several decades, and presents examples of how medical practice in other high-income countries, where people live longer, follows the priorities indicated by cost-effectiveness analysis. CEA and CBA support democratic decisionmaking processes, which have themselves benefited from scientific inquiry; these are touched on at the end of the chapter.

Introduction

Medical scientists are at work around the world producing new knowledge about health and disease and new technologies to combat disease and improve health. They have had many successes: from the epidemiological investigations around London's Broad Street pump that confirmed cholera's transmission by water, to the identification of the agents of infectious diseases themselves—bacteria and viruses—to the development of vaccines, antibiotics, and other interventions that have brought so many diseases under control.

Those discoveries are not the end of the story, nor are they the end of the need for good science. New knowledge and new technologies present human beings with choices. How those choices are made is as important to their ultimate effect on health as the knowledge and technologies themselves. Because of their importance, how those choices are made is equally in need of a scientific approach – an approach based on asking searching questions, seeking empirical evidence to address those questions, reasoning carefully about the evidence, and presenting the evidence and reasoning transparently so that others can understand, attempt to replicate, and agree or disagree with the conclusions.

Cost-effectiveness analysis (CEA) and cost-benefit analysis (CBA) are two important components of the science of decisionmaking for health. Great strides have been made in both CEA and CBA in recent decades, not only in their scientific methods but also in their use for real-world decisions. I describe them in this chapter, sketch some of the methodological progress, and present examples to show how thinking about the cost-effectiveness of health interventions has contributed to

better health in high-income countries. CEA and CBA are not the only components of scientific decisionmaking, and I mention others, such as systematic reviews and meta-analysis. All these components support democratic decisionmaking processes, which have themselves benefited from scientific inquiry, and which I discuss briefly at the end of the chapter.

The Decision Context: An International Perspective

The United States spends a much larger share of its national income on medical care than any other country, 18 cents of every dollar in 2011 (Table 1); that is 50 percent more than the next highest spender and almost double the amount spent by still other high-income countries. By rights then, the United States should have the best health in the world, including the longest life expectancy. Instead U.S. life expectancy at birth is 1 to 4 years shorter than in the other high-income countries in Table 1. The deficit continues across the lifespan, with U.S. life expectancy at age 65 also less than that in most of the other countries and no better than equal to the rest. A 2013 Institute of Medicine report shows that the situation has only gotten worse in recent years, as U.S. life expectancy has risen more slowly than that of other high-income countries.[1]

Table 1. Health spending (as a percent of national income) and life expectancy at birth and at age 65, selected high-income countries

	Health spending as a % of GDP, 2011	Life expectancy at birth, 2011	Life expectancy at 65, 2009	
			Men	Women
United States*	17.9	79	17	20
Netherlands	12.0	81	17	21
France	11.6	82	18	23
Canada*	11.2	81	18	21
Denmark	11.2	80	17	20
Germany	11.1	81	18	21
Switzerland	10.9	83	19	22
Greece	10.8	81	18	20
Austria	10.6	81	18	21
Belgium	10.6	80	18	21
Italy	9.5	82	18	22
Spain	9.4	82	18	22
Sweden	9.4	82	18	21
Japan**	9.3	83	19	24
United Kingdom*	9.3	81	18	21
Norway	9.1	81	18	21
Australia	9.0	82	19	22
Finland	8.9	80	17	22

Sources: Percent of gross domestic product (GDP) and life expectancy at birth from the World Bank, World Development Indicators 2013. Life expectancy at 65 from Health at a Glance 2011: OECD Indicators.[2]

Notes: * These three countries are compared later in this chapter as to their policies about adult tetanus boosters and cervical cancer screening; note how they compare in terms of life expectancy. **Japan has the longest life expectancy of any high-income country.

The differences do not arise from medical science. All these countries have access to the same scientific evidence and the same medical technologies as the United States. The same medical equipment, medications, and surgical procedures are available to all. The science base in medicine is so international that all countries, including the United States, routinely rely on randomized controlled trials and epidemiological studies conducted in the populations of other countries, not just their own. More than half of the clinical trials registered on clinicaltrials.gov are located outside the United States. The Cochrane Collaboration, with branches around the world, bases its systematic reviews of what works in medicine on all the evidence, not just evidence for one country or a few countries.

The central thesis of this chapter is that differences in health across these countries arise in large part from differences in the way medical choices are made at the national level. Those choices involve how the national health sector is structured, the level of resources provided through public sources, and the framework created by national decisions for health plans, insurers, providers, and patients. Such decisions can involve more specific recommendations about the use of particular interventions, such as vaccines. They can be made by a variety of groups: governments; non-government groups such as medical professional societies, which develop guidelines for care; or organizations like the Institute of Medicine, which evaluate policies and propose changes.

To appreciate the difficulty of choosing well, consider the enormous range of possibilities facing decisionmakers. The medical sector alone offers many kinds of preventive care, acute care, chronic disease management, rehabilitative care, and palliative care. Public health programs may fall in the medical sector or cross sectors. An even wider range of activities outside of medicine are important for health and life expectancy: water and sewage systems; safe handling of food and related inspections of food imports, food manufacturing plants, and restaurants; building codes and construction methods; safety regulation and safe manufacture of consumer products; safe design and maintenance of roads, bridges, train tracks, airports; regulation of environmental pollutants; and many others.

How can decisionmakers select the best ways to improve health from this vast array of choices? That is the question addressed by the science of decisionmaking in health. It may not be coincidental that the countries that do better than the United States – longer life expectancies despite spending less of their national income on medical care – have been quicker to use those methods to help set priorities. Their better health does not derive from better scientific knowledge or better medical technologies, but from better choices.

Social scientists study how people make choices, how they can be misled or misconstrue choices, and how they can make choices that best serve their goals. Decision scientists have developed methods to help people make choices. The goal of these methods is to provide a formal process for evaluating decisions in order to help decisionmakers identify the best choices – the choices that serve their values and the values of those they represent – in situations where it is difficult for them to do so unaided.

CEA, which grew out of operations research and then began to blend with CBA, is a central decision-support method used in the medical sector. I describe it next, sketch its scientific evolution, and provide examples showing how it aligns with decisions about medical care in the United

States, Canada, and the United Kingdom. CEA is less useful for the wider range of choices that affect health, those that lie outside the medical sector. So I next discuss CBA, which comes from economics, and provide examples of its use.

This is perhaps a good place to explain why I say "aligns with." CEA, which is relatively recent (CBA has a longer history), formalizes a way of thinking about social decisions that has probably always existed. The basic idea is simple: think about, even try to estimate, what a decision will cost and what its effects (benefits) will be. Then compare those with the costs and effects of competing possibilities to decide which is best. People have evaluated decisions in terms of costs and effects for eons, and some of the decisions I describe were made before CEA had developed into a formal method of evaluation.

It is important to remember that every high-income country, including the United States, provides medical services, and many other services crucial to good health, through systems that are largely financed by people other than those who receive the services or who make decisions about them. These decisions thus involve public goods, or what Nobelist Elinor Ostrom terms "social dilemmas"—terms that point to the fact that they affect groups, not just individuals—and are often made by public groups of one sort or another.[3] That being the case, I conclude the chapter by turning briefly to the processes of which CEA and CBA are a part – that is, the processes for making decisions about the social dilemmas of health in representative democracies. Because, as everyone agrees, CEA and CBA are aids to decisionmaking, not the entire decisionmaking process, and that process is as important to the quality and validity of decisions as are the CEAs and CBAs that inform it, perhaps more so. It is through that process that important social values, such as fairness and helping those worst off, are incorporated into decisions.

Cost-Effectiveness Analysis: A Guide to Better Decisions

CEA begins with the objective of maximizing health subject to a budget constraint. In other words, if $3 trillion is available for medical care for the U.S. population, CEA asks: What services should be purchased with that money to produce the best possible health for people living in the United States?

The next step is conceptually simple. Evaluate the choices – the alternatives facing decisionmakers – in the terms established by the objective (health) and the constraint (costs). Make sure to include whatever is already being done, the status quo, among the alternatives. Based on the scientific evidence, estimate how much each alternative contributes to health and how much it costs. Then rank the alternatives in terms of their contributions to health, starting with the one that contributes the least, and compare them.

Table 2 presents a particularly straightforward CEA of tetanus boosters for adults. The purpose of the analysis was to evaluate the U.S. Centers for Disease Control and Prevention's recommendation that adults receive routine boosters at intervals of 10 years after the completion, at age 6, of a full schedule of childhood vaccinations. The analysis estimated the costs and years of life of three alternatives: no routine boosters after age 6; a single booster at age 65; and the recommended

boosters every 10 years. All three alternatives included boosters for deep wound injuries. (The authors of the analysis, Balestra and Littenberg,[4] identified a single booster at age 65 as worth considering based on evidence that the rate of tetanus in the United States, although low, rises sharply with age.) Projecting outcomes for a cohort of 3.6 million 6-year-olds over 80 years, the authors estimated that boosters at age 65 would prevent about 300 cases of tetanus, compared with no routine adult boosters. Boosters every 10 years were estimated to prevent an additional 500 cases. (Although tetanus toxoid is typically delivered in a combination vaccine that includes diphtheria and pertussis, the analysis considered only tetanus toxoid and assumed that it was delivered without the other two.)

Table 2. Cost-effectiveness analysis of adult tetanus boosters

	Undiscounted		Discounted at 5%/year			
	Cost*	Life expectancy	Cost*	Life expectancy	ICER*	ICER, 2012
No routine booster	$0.832	68.315319	$0.060	19.464526		
Booster at age 65	$0.996	68.315417	$0.074	19.464529	$4,500	$15,800
Booster every 10 yrs	$4.135	68.315464	$0.919	19.464532	$281,700	$986,100

Source: Adapted from Balestra DJ, Littenberg B. J Gen Intern Med. 1993;8:405-412.[4]

*Note: Incremental cost-effectiveness ratio, 1986 dollars, rounded to the nearest hundred. All three alternatives include a booster in the event of a deep wound.

The alternatives are arranged in Table 2 in order of effectiveness, starting with no routine boosters, the least effective. If a more effective alternative costs less, the decision is easy. The less effective alternative can be eliminated from the choice set.

In this case, however, as commonly happens, greater effectiveness comes at higher cost, making the decision more difficult. So the next step is to compare alternatives two at a time, calculating a cost-effectiveness ratio to summarize the comparison. A cost-effectiveness ratio, often called an incremental cost-effectiveness ratio (ICER), is the difference in health between two alternatives, divided by the difference in their costs. For a single booster at age 65 compared with no routine boosters, the cost-effectiveness ratio is ($0.074 - $0.060)/(19.464529 - 19.464526), which yields a cost-effectiveness ratio of $4,500 for each year of life gained, in 1986 dollars. Scaled up to 2012 price levels, a single booster at age 65 would cost an additional $15,800 for every year of life gained, compared with no routine boosters after age 6. This is very reasonable compared with many interventions routinely offered in high-income countries.

The comparison between boosters every 10 years and a single booster at age 65 is strikingly different. Boosters every 10 years cost an additional $281,700 in 1986 dollars for each year of life gained – almost $1 million in 2012 dollars.

That comparison, health for dollars, is the central contribution of CEA. CEA does not ask how much an intervention costs per person – in this case, not very much as Table 2 shows. Nor does it ask how much the vaccine costs per dose; the study used a price of 43 cents per dose (and assumed

no additional cost for administering it). Even today, a single dose costs a dollar or two. By these measures, it would hardly seem worthwhile to take the trouble to evaluate such an inexpensive intervention.

CEA asks: How much does it cost to contribute to the objective – good health? What is the cost in relation to what we really care about, health? Cost per dose, or cost per person, is not the point, and looking at those numbers can be misleading. It turns out that giving boosters every 10 years is a very expensive way to get better health.

That information by itself, however, is not enough to rule out boosters every 10 years. The next question is: What else is out there? What else can be done to improve health, and how does that compare with boosters every 10 years?

To illustrate that next step, Table 3 presents cost-effectiveness ratios for a larger set of choices. The first column shows the alternatives compared and the second column the cost-effectiveness ratio for that comparison from the original study. The third column adjusts the original ratios to 2012 dollars to permit comparisons across studies. The final column flips each cost-effectiveness ratio on its head to bring attention back to the goal, better health: it shows the number of additional healthy years that can be produced by that intervention, compared with its alternative, for an expenditure of $1 million.

The final column shows that there are many possibilities for improving health that make better use of health resources. Aspirin for men at high risk of heart disease is cost saving, an easy choice.[5] Smoking cessation programs are next most productive, bringing 161 years for each $1 million; the number reflects the cost-effectiveness ratios of a mixture of programs with different approaches, weighted by enrollment.[6] Aspirin for lower-risk men,[5] total knee arthroplasty for people at low risk of perioperative complications,[7] and tetanus boosters at 65[4] are also productive choices.

Clearly tetanus boosters every 10 years would not be at the top of the list, but neither would they automatically be rejected. The guideline for using CEA is to select the most productive choices first – the ones that contribute the most to health – and keep selecting as long as the resources hold out. If sufficient resources are available to provide all the interventions, all would be provided. If the resources are used up by interventions that provide more health, boosters every 10 years would not be provided — or should not be provided. The title of the table includes the economic term “opportunity cost.” The opportunity cost of a choice is the benefit lost because of that choice. Good decisions bring the most benefit – whatever was not chosen would have brought less. Table 3, last column, illustrates the concept. If the choice is tetanus boosters every 10 years, and smoking cessation programs are not chosen, each year gained from tetanus boosters has an opportunity cost of 161 years, the years lost because smoking cessation programs were not provided.

This may be a good point at which to make clear that CEA is designed to set priorities, given that health is the objective. It does not set the budget. That comes from elsewhere. To be more explicit, CEA is not a tool for controlling costs. It is a tool for making the best possible use of

Table 3. Opportunity costs: healthy years for $1 million, selected interventions

	Cost per healthy year (original year $)	Cost per healthy year (2012 $)	Healthy years per $1 million (2012 $)
Tetanus boosters			
Booster at age 65	4,527 (1986)	15,845	63
Booster every 10 years	281,748 (1986)	986,118	1
Total knee arthroplasty (TKA) for advanced knee osteoarthritis			
TKA, low risk of perioperative complications	9,700 (2006)	11,971	84
TKA, medium risk of complications	18,700 (2006)	23,077	43
TKA, high risk of complications	28,100 (2006)	34,678	29
Aspirin to prevent heart disease			
men 45, 10-year risk 2.5%	9,800 (2003)	13,686	73
men 45, 10-year risk 5.0% or higher	cost-saving	cost-saving	cost-saving
Mammography for breast cancer			
women 50-79, screened every 2 years	25,021 (2002)	36,349	28
Screening for diabetes			
age 55/high blood pressure vs no screening	34,375 (1997)	60,794	16
all adults 55 vs adults w high blood pressure	360,966 (1997)	638,384	2
Screening once for HIV			
HIV prevalence = 1.0%	30,800 (2004)	41,209	24
HIV prevalence = 0.1%	60,700 (2004)	81,214	12
Diet/exercise to prevent diabetes			
adults at high risk of diabetes	143,000 (2000)	227,495	4
adults with diabetes	35,400 (2000)	56,317	18
Smoking cessation			
15 programs weighted by % enrolled	3,294 (1995)	6,198	161

Sources: Tetanus boosters, Balestra and Littenberg, 1993[4] and Table 2; TKA, Losina, Walensky, Kessler, et al, 2009[7]; aspirin, Pignone, Earnshaw, Tice, et al, 2006[5]; mammography, Mandelblatt, Saha, Teutsch, et al, 2003;[8] diabetes screening, Hoerger, Harris, Hicks, et al, 2004;[9] HIV screening, Paltiel, Walensky, Schackman, et al, 2006;[10] diet/exercise, Eddy, Schlessinger, Kahn, 2005;[11], and smoking cessation, Cromwell J, Bartosch WJ, Fiore MC, et al, 1997.[6]

whatever health resources are available. By itself CEA will not reduce U.S. medical spending, nor even reduce its growth rate. It can help the United States allocate medical spending more effectively to improve health and move closer to the life expectancies achieved by other high-income countries, but it will not control costs.

The Scientific Evolution of CEA

The purpose of CEA is comparison – comparison of alternative health interventions in order to identify those that make the most productive use of resources. To serve that purpose, CEAs must be comparable. And, as Table 3 suggests, they need to be comparable across a wide range of very different interventions aimed at very different health conditions. When they are not comparable, they can lead to poorer choices, not better ones. Much of the methodological development over the 40-50 years since CEA was first applied in health has gone to making CEAs comparable. I will briefly describe two central developments: the Quality-Adjusted Life-Year (QALY) and the Reference Case.

The QALY was developed because a comprehensive measure of health is needed that allows comparisons of interventions that deal with different diseases. Cases of disease, used in some early CEAs, could not serve the purpose; the implications for health of a case of diabetes are very different from those of a case of influenza. Lives saved was an improvement but did not take into account that some measures save many more years of life than others. Years of life, used in the tetanus booster study, improved on lives saved by giving more credit to interventions that saved many years but did not give credit for improvements in the quality of life – less pain, better function – that are important contributions of some interventions.

And so the QALY was created.[12] QALYs are years of life with each year weighted by the state of health, or quality, of that year. Quality is measured on a 0-1 scale – 0 for death, 1 for perfect health. Weights are elicited from a representative sample of people using various methods: standard gamble, time tradeoff, paired comparisons, or rating scales. Early QALY systems were the Quality of Well-Being (QWB) scale, the Health Utilities Index, and the EuroQoL (now the EQ-5D and EQ-5D-5L).[a] Health utility measurement, as it is often called, continues to be an active field of research.

QALYs address a crucial issue, but CEAs need to be comparable in other ways as well: health states included in QALYs, resources counted and their valuation, discount rate, methods for exploring uncertainty, how results are reported, and much more. I was fortunate enough to co-chair, with Milton Weinstein, one of the early efforts to address the issue of the broader comparability of CEAs. The U.S. Public Health Service convened the first Panel on Cost-Effectiveness in Health and Medicine in 1993 and asked the panel to assess the state of the field and provide recommendations for the conduct of CEAs "in order to improve their quality and encourage their comparability."[12] Michael McGinnis, Assistant Surgeon General (retired), explained: "The overarching goal for this work has been to move the field forward so that over the next decade, State and Federal decisionmakers will have access to robust information with respect to the true cost per health effect gained for the continuum of health-related interventions – be they preventive, palliative, curative, or rehabilitative."[12] Thus, at the outset, the panel's charge directed it to take a broad view and consider how CEA could inform decisions across the entire range of health interventions.

[a] See the appendix to Gold, Siegel, Russell, et al, chapter 4, for a description of these systems.[12]

To meet that charge the panel's recommendations defined a "reference case" analysis.[12-15] In the words of the panel "To promote comparability of CEAs …, the panel proposes that studies include … a reference case. The reference case is defined by a standard set of methods and assumptions. It includes a set of standard results: the reference case results. While an investigator might also present results based on different methods and assumptions to serve the other purposes of the analysis, the reference case serves as a point of comparison across studies. It should be included whenever the CEA is intended to contribute to decisions about the broad allocation of health care resources."[13] The Second Panel on Cost-Effectiveness in Health and Medicine[b] was convened in 2012 to review and update the reference case; that panel's report is due in 2016.

In the last 20-25 years, CEA has become an integral part of decisionmaking in many health agencies around the world, and the reference case has been a powerful organizing principle and a popular name for the resulting standards. Those agencies, like Australia's Pharmaceutical Benefits Advisory Committee (PBAC), the Canadian Agency for Drugs and Technologies in Health (CADTH), and the United Kingdom's National Institute for Health and Care Excellence (NICE), each have their own guidelines for economic evaluations, which are designed to promote comparability across the studies that inform their recommendations.[16-18] In addition, the World Health Organization (WHO) has guidelines for CEA,[19] and the Gates Foundation has developed its own reference case.[c]

Since the first panel published its recommendations, many strands of research have contributed to the improvement of CEA. They include better methods for systematic review and meta-analysis of the health and medical literature, which have contributed to better estimates of the parameters needed for CEAs. Contributions from PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses),[d] the Cochrane Collaboration,[e] and the Agency for Healthcare Research and Quality (AHRQ)[f] have been particularly important, as well as reports like the Institute of Medicine's Finding What Works in Health Care: Standards for Systematic Reviews.[g] Advances in simulation modeling, another methodology fundamental to CEA, and standards for modeling, have also been important; I would mention NIH's CISNET modeling network[h] and the recommendations of the 2012 ISPOR-SMDM (International Society for Pharmacoeconomics-Society for Medical Decision Making) Modeling Good Research Practices Task Force.[20]

How CEA Results Align with National Choices

I will now turn to examples that illustrate the differences in medical choices between the United States and two other English-speaking countries, Canada and the United Kingdom. Both Canada

[b] See http://2ndcep.hsrc.ucsd.edu/ for more information.

[c] Bill and Melinda Gates Foundation Methods for Economic Evaluation Project (MEEP), prepared by NICE International; final report submitted to the Foundation in January 2014.

[d] For more information on PRISMA, go to http://www.prisma-statement.org/.

[e] For more information on the Cochrane Collaboration, go to http://www.cochrane.org/Cochrane-reviews.

[f] See http://www.effectivehealthcare.ahrq.gov/ to access information about products produced by AHRQ's Effective Health Care Program.

[g] Available at http://www.iom.edu/Reports/2011/Finding-What-Works-in-Health-Care-Standards-for-systematic-Reviews.aspx.

[h] See http://cisnet.cancer.gov/about/.

and the United Kingdom spend much less of their national income on medical care, yet have longer life expectancies at birth and at age 65 than the United States (Table 1). Both countries consider cost-effectiveness in setting medical policies, at the provincial level in Canada and the national level in the United Kingdom, and have increasingly formalized the use of CEA in recent years. The United Kingdom created NICE in 1999 to recommend services to be provided by the National Health Service (NHS). NICE includes an economic evaluation – a CEA – in every appraisal and has become a world leader in cost-effectiveness methods.[i] CEA also plays a role in decisions made by Canada's provincial health plans; more detail will be provided in the section on process, which includes the process in Ontario, home to almost 40 percent of Canadians.

Tetanus Boosters

Consider the recommendations for tetanus boosters in the United States, Canada, and the United Kingdom. The CEA presented earlier showed that a single booster at age 65 is good value, but that providing routine boosters every 10 years is an expensive way to produce better health.

All three countries recommend a series of doses of tetanus toxoid before the age of 6 – five doses in the United Kingdom and six in the United States and Canada – followed by a booster just before or during adolescence.[j] For adults, both the United States and Canada recommend boosters every 10 years. The United Kingdom does not recommend routine adult boosters. The NHS site says: "After the full course of five injections, you should have life-long immunity against tetanus. However, if you or your child has a deep wound, it's best to get medical advice."[21] The NHS also recommends that anyone traveling to a country with "limited medical facilities" get a booster first if their last one was more than 10 years ago.

The national recommendations are in line with national spending. The United Kingdom, which spends only 9.3 percent of its national income on medical care, recommends the most limited use of tetanus boosters, not even including one at age 65. Canada, at 11.2 percent, and the United States, at 17.9 percent, recommend routine adult boosters every 10 years.

The tetanus vaccine has been available for 90 years, and these recommendations were probably not influenced by an actual CEA in any country. The CEA mindset, however – how does this intervention contribute to better health – would be easy to apply informally. A few back-of-the-envelope estimates of the number of cases prevented and the cost of repeated boosters would point to the same conclusion. A formal analysis is not always necessary. Long before NICE, the United Kingdom led in applying the CEA mindset to decisions about the National Health Service.[22]

Cervical Cancer Screening

In 1981, Eddy published a CEA that evaluated the cost-effectiveness of screening for cervical cancer at different frequencies.[23] In my 1994 book,[24] I devote an entire chapter to cervical cancer

[i] In support of my earlier point that CEA is not a cost-control mechanism, NICE was created at the same time that the Blair administration, as a matter of policy, increased spending on the NHS to bring health spending's share of national income up to the average for Europe.

[j] Tetanus recommendations, see: United Kingdom: NHS, http://www.nhs.uk/conditions/vaccinations/pages/vaccination-schedule-age-checklist.aspx; Canada: http://www.phac-aspc.gc.ca/im/ptimprog-progimpt/table-1-eng.php and http://www.phac-aspc.gc.ca/im/vpd-mev/tetanus-tetanos/prevention-eng.php; United States: http://www.cdc.gov/vaccines/vpd-vac/tetanus/fs-parents.html.

screening and discuss the American Cancer Society's first attempt, in 1980, to reduce the frequency of cervical cancer screening based on Eddy's results. Similar results from a 1990 article are shown in Table 4, updated to 2012 dollars.[25]

Table 4. Opportunity costs: Cervical cancer screening by screening interval

Screening intervals compared	Cost per year of life gained (2012 $)	Life-years per $1 million
Every 3 years vs. no screening	$48,618	20.57
Every 2 years vs. every 3	$1,534,582	0.65
Annually vs. every 2 years	$3,890,556	0.26

Source: Adapted from Eddy DM, 1990.[25]

No one questions that screening for cervical cancer is effective, but the CEA showed that its cost-effectiveness depends strikingly on screening frequency. In most cases cervical cancer is a slow-moving disease; few new cases are discovered with annual screening, and the cost in health of not catching them immediately is small, while the cost in resources of screening so often is large. As screening grows more frequent, the cost of an additional year of life rises sharply. As Table 4 shows, screening every year instead of every 2 years costs, in 2012 dollars, almost $4 million for each year of life gained.

In response to Eddy's 1981 study,[23] which was available even earlier, the American Cancer Society (ACS) changed its recommendation for annual screening and recommended screening every 3 years. The reaction, subsequently repeated in response to other recommendations to limit screening, persuaded ACS to go back to its original recommendation despite the CEA results. In 2002, ACS tried again, recommending at least 3 years between screenings for women with negative tests.[k] Yet U.S. physicians continue to conduct more frequent screening.[26,27]

The United Kingdom and Canada, along with other high-income countries that achieve longer life expectancies while spending less than the United States on medical care, have long recommended cervical cancer screening at 3-5 year intervals. The NHS introduced screening in the 1980s: "Women aged between 25 and 49 are invited for testing every 3 years, and women aged between 50 and 64 are invited every 5 years."[28] The Canadian recommendations are to begin routine screening at age 25, screen women 25-69 every 3 years, and stop screening women 70 and older who have had three negative screens in the preceding 10 years.[29] These recommendations are based as much on the benefit/risk considerations as on cost-effectiveness. For example, the Canadian Preventive Services Task Force notes: "…the recommended 3-year interval balances the small incremental potential for benefit from shorter intervals against the greater potential for harm from increased testing and procedures with more frequent screening. Most countries outside North America use 3- or 5-year intervals." A 2004 Commonwealth Fund survey of five countries—which shows that the United

[k] See http://www.cancersocietyofamerica.org/cervical_cancer.html for current American Cancer Society guidance on cervical cancer screening.

States screens more frequently than Canada, the United Kingdom, Australia, or New Zealand for both cervical and breast cancer—underscores the point.[30] All those countries spend less than the United States to achieve longer life expectancies.

Cost-Benefit Analysis: Looking for Health Beyond Medical Care

For choices in the medical sector, a comprehensive measure of health like the QALY is sufficient to measure everything, or almost everything, of value that is gained from an alternative. Choices that fall outside the medical sector may involve substantial benefits of other kinds, in addition to better health. Good decisions require that everything of value be considered in making a choice. But how to aggregate the different benefits? Some form of aggregation, of summarizing benefits (and costs), makes it easier to compare alternatives and to give appropriate weight to each benefit. When benefits are not summarized decisionmakers, who are overwhelmed with information, may focus on one or two and lose sight of the rest.

CBA allows the aggregation of health and non-health benefits because it values benefits, as well as costs, in money terms. In the United States, the Office of Management and Budget requires that a CBA be performed for all 'economically significant' regulations proposed by Federal agencies.[31]

Under this requirement, many of the U.S. Environmental Protection Agency's (EPA) regulations, which involve substantial health benefits, are subject to CBA.[32] In 2011, for example, EPA published a CBA of the 1990 Clean Air Act amendments, which compared their costs and benefits to those of the legislation in force when they were enacted, the 1970 Clean Air Act and 1977 amendments.[33] The analysis valued a variety of benefits: reductions in mortality and illness; improvements in visibility at recreational sites and in residential areas; benefits to commercial timber, agricultural crops, and recreational fishing; and reduced materials damage. The report noted that other potential, and even known, benefits had been identified but left unvalued "… many beneficial outcomes involving human health or environmental improvement could not be expressed in terms of economic values because the scientific and economic studies to support such valuations remain inadequate or unavailable."[33]

Costs (public and private) and benefits were projected to rise over the three decades covered by the analysis, 1992-2020, with benefits consistently much greater than costs: by 2020, benefits were projected to reach almost $2 trillion, compared with $65 billion in costs.[33] Mortality benefits made up 90 percent of the monetized benefits. For those uncomfortable with valuing lives and health in money terms, the report included a rough cost-effectiveness ratio: the projected cost per life saved in 2020 was $280,000. If each premature death avoided means 5 more years of life, on average, that works out to $56,000 per year of life, or 18 years of life for $1 million. The calculation ignores, of course, non-health benefits.

Table 5 details the projected health benefits, in natural units, for the two pollutants for which epidemiological evidence is strong and supports projections, fine particles and ozone. For other pollutants, such as carbon monoxide, estimates of the health benefits were not possible, e.g., because their effects could not be distinguished from those of co-occurring pollutants.

Table 5. Health effects of the 1990 Clean Air Act Amendments, numbers of cases, 2010 and 2020

Health Effect Reductions (PM2.5 & Ozone Only)	Pollutant(s)	Year 2010	Year 2020
PM2.5 Adult Mortality	PM	160,000	230,000
PM2.5 Infant Mortality	PM	230	280
Ozone Mortality	Ozone	4,300	7,100
Chronic Bronchitis	PM	54,000	75,000
Acute Bronchitis	PM	130,000	180,000
Acute Myocardial Infarction	PM	130,000	200,000
Asthma Exacerbation	PM	1,700,000	2,400,000
Hospital Admissions	PM, Ozone	86,000	135,000
Emergency Room Visits	PM, Ozone	86,000	120,000
Restricted Activity Days	PM, Ozone	84,000,000	110,000,000
School Loss Days	Ozone	3,200,000	5,400,000
Lost Work Days	PM	13,000,000	17,000,000

Source: U.S. Environmental Protection Agency 2011.[33] Exhibit 8. Differences in key health effects outcomes associated with fine particles (PM2.5) and ozone between the *With-CAAA* and *Without-CAAA* scenarios for the 2010 and 2020 study target years (in number of cases avoided, rounded to 2 significant digits).

Note: The table shows the reductions in risk of various air pollution-related health effects achieved by the 1990 Clean Air Act Amendment programs, with each risk change expressed as the equivalent number of incidences avoided across the exposed population.

This CBA confirmed that a choice that had already been made was, and continued to be, beneficial: the benefits exceed the costs by a wide margin. It is more difficult to find evidence to show how CBA contributes earlier in the process, when choices are being considered but have not yet been made, because not all of the process is open to public view. As proponents of CEA and CBA often point out, the framework itself, and the experience of working with that framework, can shape the alternatives considered and the final choice in significant ways, leading participants to uncover new technological options, evidence for new health and non-health effects, better ways to value those effects, and new ideas for institutional arrangements to address a problem, in addition to testing their views of the costs and benefits of each option against careful estimates.[34,35]

Nichols' history[36] of an early CBA of regulations to reduce lead in gasoline shows how CBA can work prospectively to help identify promising alternatives and make decisions. Lead levels in gasoline had been substantially reduced by the early 1980s, in part because other environmental regulations required cars to be outfitted with catalytic converters, which fared better with unleaded gasoline. EPA was under no outside pressure to do more. But evidence suggested substantial benefit from reducing lead in gasoline still further. When a preliminary analysis supported the idea, the

agency proceeded with a full CBA and a regulatory proposal. The new regulation, to reduce lead in gasoline from 1.1 grams per leaded gallon to 0.1 grams, went into effect in 1985.[37]

In Nichols' words the CBA "played a key role in shaping the rule."[36] The CBA showed, for example, that reducing lead to 0.1 grams per leaded gallon would yield substantial benefits at reasonable cost but that a complete ban might cause problems for heavy agricultural machinery; so EPA considered, but did not propose, a ban. (Congress included a ban in the 1990 Clean Air Act Amendments).[38] Although sensitivity analyses indicated that only under very unlikely circumstances might refineries fail to meet the schedule for reducing lead, analysts designed a lead banking program to make the transition still easier: refineries that reduced lead ahead of schedule could save the difference to use later or to sell to other refineries. As a third example, after the analysis began, EPA staff learned of new evidence that linked lead in gasoline to hypertension in adults.

The final CBA considered four very different kinds of benefit: (1) health benefits to children because less lead in gasoline would reduce the number who suffered physical and cognitive problems from elevated blood lead levels; (2) less damage to catalytic converters and thus better control of conventional pollutants; (3) lower expenditures on car maintenance due to less mis-fueling (use of leaded gas in cars designed for unleaded); and (4) fewer deaths and less morbidity related to hypertension.

Valuing benefits in money terms allowed these different benefits to be aggregated and showed their importance relative to each other. The first three categories of benefit – children's health, conventional pollutants, and car maintenance – were projected to be $1,657 million in 1988, well above the projected additional refining costs of $532 million (both 1983 dollars). The benefits from reducing hypertension were far larger: an additional $5.4 billion, bringing total benefits to about $7 billion.[36] The agency's legal counsel advised, however, that the hypertension benefits could not be used to support the legal case for regulation because the scientific evidence for the lead-hypertension link had not yet been peer reviewed. So the results were presented two ways – with and without the benefits from reducing hypertension. The hypertension benefits served to make the conclusion more robust: benefits exceeded costs.

It may be worth spending a little space to discuss the science behind the valuation of benefits in CBA. The social sciences, as I mentioned earlier, study how people make decisions and the circumstances under which they make the best decisions, those that bring them what they want. Different social sciences have chosen different natural laboratories for this study. Economists' natural laboratory is the market. The theory with which they approach real-world markets is the theory of the competitive market. In competitive markets, well-informed buyers and sellers, none of them large enough to manipulate the market, exchange money for goods; the buyers receive all the benefit of the good, the sellers incur the full costs of producing the good.

When these conditions hold, and only when they hold, market prices represent, at one and the same time, how much people are willing to pay for one more unit of the good and how much it costs to produce one more unit of the good. Buyers' willingness to pay reflects a tradeoff – to obtain this good the buyer had to forego other goods. The buyer had to value it more highly than other goods. It is the tradeoff between this good and other goods that is the fundamental measure of value, with

money as the common medium for measuring tradeoffs; the money price of a good represents the amount of other goods given up. The willingness to give them up signals that the buyer values this good more than the other goods. On the supply side, the cost of producing the good reflects the value of the goods that could not be produced, the opportunity cost of producing this good instead of others. In short, the competitive market system is a system for working through values and allocating resources to their mostly highly valued uses.

In the real world, of course, many markets fall well short of the requirements for competitive markets – certainly most of the markets that make up the medical sector do. So they don't produce the kinds of outcomes that accurately reflect people's valuations of goods and services. The purpose of CBA – the reason it was developed – is to evaluate the pros and cons of choices that arise in poorly functioning markets, or that occur outside of markets altogether, and in tallying those pros and cons, to approximate the results that would come naturally from competitive markets: measures of benefits for each alternative that correctly represent the willingness to pay of those affected, their true valuation; and measures of cost for each alternative that correctly represent the costs of production, their true opportunity costs.

Fifty years ago, economists valued lives and health for purposes of CBA at the present value of future earnings.[39] Several economists, including Drèze, Schelling, and Mishan argued that this was wrong. They reasoned that life-saving should be valued at what the people affected would be willing to pay, the same principle that operates in the competitive market, and not by their discounted future earnings. By the 1980s, the concept "that social decisions should, so far as possible, reflect the interests, preferences and attitudes to risk of those who are likely to be affected by the decisions" was accepted as a better way to think about valuing life-saving.[40]

Since then, the basic principle in CBA has been that benefits should be valued as the people involved value them. Dollar valuations, like QALY weights, come from the people who will experience the benefits or from people like them. The methods used in the two CBAs I have described follow that principle: benefits are valued based on evidence from markets or, if the good is not sold in markets, from related markets, evidence that reveals people's tradeoffs, their valuations. Lives saved, for example, are valued based on the tradeoffs revealed in labor markets, between wages and risk of death or injury on the job.[32] The analysis of lead in gasoline was an early case of the use of this "value of a statistical life."

The willingness-to-pay principle in CBA has been criticized for its link to people's incomes – people with more money can be "willing" to pay more than people with less money. In actual CBAs, this problem is at least somewhat attenuated by using average willingness to pay. CBA has also been criticized for ignoring the distributional consequences of decisions based on CBA, just as QALYs have been criticized for treating all QALYs as equal. Such criticisms are reasons why the process to which CEA and CBA contribute is important.

The Process that Includes CEA/CBA

Everyone agrees that CEA and CBA are aids to decisionmaking, not machines that produce decisions all by themselves without human intervention. A CEA or CBA is one element in a larger

decisionmaking process. That process is often spelled out in legislation, agency guidelines, or the opening decisions of committees, and it involves many of the elements we associate with good science and representative democracy.

The process itself is as important as the CEA or CBA. This section sketches some features of the process, or processes, to which CEA and CBA contribute. This is not my area of expertise, so it is not a comprehensive summary; a few examples are provided that point to the importance of process and to some common elements that, in many cases, we simply take for granted will be present.

Nichols' account of the lead in gasoline regulations contains numerous references to the process: the posting of a preliminary regulatory proposal and CBA, with a public comment period; the changes in the analysis and in the regulatory proposal as those involved learned more, sometimes from the public comments; the advice from legal counsel.[36] Similarly, EPA's 2011 summary report on the 1990 Clean Air Act amendments makes frequent references to process.[33] An example from the acknowledgements: "the full integrated report and this summary report were reviewed by the EPA Science Advisory Board's Advisory Council on Clean Air Compliance Analysis … and its three technical subcommittees." The three technical subcommittees specialized in air quality modeling, health effects, and ecological effects.

In Canada, the Ontario Health Technology Advisory Committee,[1] created in 2003 to advise the Ontario Ministry of Health and Long Term Care about nondrug technologies, has developed a multi-criteria decision framework to guide its deliberations. An article describing the framework states: "Designing a process that weighs scientific evidence appropriately, takes cost and values into consideration, and makes decisions fairly is not straightforward."[41] Only technologies approved by Health Canada, whose duties include safety and efficacy regulation of drugs and health products, are considered. Once the technology comes to Health Quality Ontario, the committee considers four categories of criteria in making its recommendation: (1) overall clinical benefit, (2) consistency with societal and ethical values, (3) value for money, and (4) feasibility of adoption into the health system. CEA, item number 3 in the list, enters through only one of the four categories.

In the United Kingdom, NICE is known for giving CEA a central role in making recommendations to the National Health Service. Yet the overview of the decisionmaking process, the 93-page Guide to the Methods of Technology Appraisal 2013,[18] spends at least as much time on the process of which CEA is a part as on guidelines for the CEA itself. Section 4 (seven pages), "Involvement and Participation," deals with the processes for ensuring that all points of view are represented, including patients, clinicians and manufacturers. Section 6 (10 pages), "The Appraisal of the Evidence and Structured Decision-making," discusses how an appraisal committee should be composed, what evidence it should consider, how it should seek and incorporate public comment, the social and ethical values it is bound to take into account, the need for transparency in its deliberations, for clarity in explaining the reasons for its recommendations, and other conditions. NICE explicitly states that "life-extending treatment at the end of life" may be evaluated differently from the general guidelines for cost-effectiveness, using a weighting of QALYs different from that for other technologies, so the process spells out the general approach and exceptions to that approach.

[1] Now Health Quality Ontario. Go to http://www.hqontario.ca/ for more information.

Any scientific, empirically based method for evaluating choices requires projections of what will happen as a result of the choice. This is the contribution of CEAs and CBAs. Part of a good process is setting standards for those CEAs/CBAs—i.e., they are used to compare, so they must be comparable. I have mentioned some of the standards in the course of this chapter. Here, I want to point out that standards have been developed by every group that uses these decisionmaking tools. EPA, for example, explains that its Guidelines for Preparing Economic Analyses[42] provide "(3) … an overarching framework for economic analyses throughout the Agency and across EPA Program Offices; and (4) ensure that important subjects such as uncertainty, timing, and valuation of costs and benefits, are treated consistently in all economic analyses at EPA." Standards for the analyses that inform decisions are part of defining the decisionmaking process.

The processes aided by CEA/CBA involve what Ostrom calls "social dilemmas."[3] These are situations in which people need to act collectively to solve problems. Ostrom points out that behavioral experiments show that people act cooperatively much more often than would be predicted by theories based on pursuing individual self-interest. Further, "field research also shows that individuals systematically engage in collective action to provide local public goods or manage common-pool resources without an external authority to offer inducements or impose sanctions." Whether we are hardwired to cooperate, or have learned to do so, there are many situations in which cooperation leads to better outcomes for everybody, and human beings have the good sense to try to cooperate in those situations. Many of those decisions affect our health.

Ostrom[3] reports two experimental findings that begin to explain how people arrive at cooperative solutions. First, talking face-to-face is important: "simple, cheap talk allows individuals an opportunity to make conditional promises to one another and potentially to build trust that others will reciprocate." Notice, in this regard, the omnipresence of committees – committees that meet in person – in the decisionmaking processes I have described. Second, people set up structures and rules that encourage cooperation and discourage refusal to cooperate, or, as Ostrom puts it, they "solve second-order social dilemmas that change the structure of the first-order dilemma." … "Most robust and long-lasting common-pool regimes involve clear mechanisms for monitoring rule conformance and graduated sanctions for enforcing compliance." In this regard, notice the effort that goes into developing a process – who is represented on a committee, what evidence the committee is to consider, how the committee's work is to be reviewed – as well as structures and regulations designed to enforce a decision once it is adopted, which I have not covered in this chapter.

This is a story of people developing new techniques and new institutions to guide choices in social dilemma situations in medical care and public health. Social scientists who work with CEA and CBA are contributing to that process. They might also fruitfully devote some of their scientific attention to the rest of the process. They could ask, "What is this process that is aided by CEA/CBA? How well does it work? How could it be structured to work better?" Some social scientists have been approaching the issue from the other side – the process rather than the CEA/CBA – and conversations among those working on the two sides might be useful. Ostrom[3] has said: "…individuals temporarily caught in a social-dilemma structure are likely to invest resources to innovate and change the structure itself in order to improve joint outcomes." CEA and CBA are part of changing the structure to improve the outcomes of decisions about health. What more can we do, as social scientists, to help?

Implications for Practice

Although the United States spends far more of its national income on medical care than any other high-income country, U.S. citizens live shorter lives. If this situation is to change, U.S. decisionmakers in health and medicine need to make more use of the decision practices and supports used by other high-income countries, particularly CEA and CBA, to help direct resources to more productive ways to improve health.

Implications for Research

To support better health in the United States, more research needs to focus on providing decisionmakers with the information that will help them identify the uses of medical spending that contribute the most to better health. Government agencies and medical professional societies should expand their efforts to consult with decisionmakers on a research agenda, conduct the necessary studies, and publicize the results.

Acknowledgments

The opinions presented herein are those of the author and do not necessarily reflect the position of the Agency for Healthcare Research and Quality, the National Institutes of Health, or the U.S. Department of Health and Human Services.

Author's Affiliation

Institute for Health, Health Care Policy, and Aging Research and Department of Economics, Rutgers University, New Brunswick NJ.

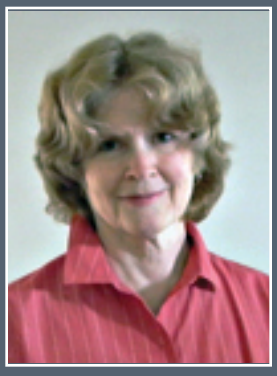
Address correspondence to: Louise B. Russell, PhD, Institute for Health, 112 Paterson Street, New Brunswick, NJ; email: lrussell@ifh.rutgers.edu

References

1. Institute of Medicine. U.S. health in international perspective: shorter lives, poorer health. Washington, DC: National Academies Press; 2013.
2. OECD. Health at a glance 2011: OECD indicators. Paris: OECD Publishing; 2011. Available at http://dx.doi.org/10.1787/health_glance-2011-en. Accessed February 24, 2015.
3. Ostrom E. A behavioral approach to the rational choice theory of collective action: presidential address. Am Polit Sci Rev 1998;92(1):1-22.
4. Balestra DJ, Littenberg B. Should adult tetanus immunization be given as a single vaccination at age 65? A cost-effectiveness analysis. J Gen Intern Med 1993;8:405-12.
5. Pignone M, Earnshaw S, Tice JA, et al. Aspirin, statins, or both drugs for the primary prevention of coronary heart disease events in men: a cost-utility analysis. Ann Intern Med 2006;144:326-36.
6. Cromwell J, Bartosch WJ, Fiore MC, et al. Cost-effectiveness of the clinical practice recommendations in the AHCPR guideline for smoking cessation. JAMA 1997;278:1759-66.
7. Losina E, Walensky RP, Kessler CL, et al. Cost-effectiveness of total knee arthroplasty in the United States: patient risk and hospital volume. Arch Intern Med 2009;169:1113-21.
8. Mandelblatt J, Saha S, Teutsch S, et al. The cost-effectiveness of screening mammography beyond age 65 years: a systematic review for the U.S. Preventive Services Task Force. Ann Intern Med 2003;139:835-42.
9. Hoerger TJ, Harris R, Hicks KA, et al. Screening for type 2 diabetes mellitus: a cost-effectiveness analysis. Ann Intern Med 2004;140:689-99.
10. Paltiel AD, Walensky RP, Schackman BR, et al. Expanded HIV screening in the United States: effect on clinical outcomes, HIV transmission, and costs. Ann Intern Med 2006;145(11):797-806.
11. Eddy DM, Schlessinger L, Kahn R. Clinical outcomes and cost-effectiveness of strategies for managing people at high risk for diabetes. Ann Intern Med 2005;143:251-64.
12. Gold MR, Siegel JE, Russell LB, et al (eds). Cost-effectiveness in health and medicine. New York: Oxford University Press; 1996.
13. Russell LB, Gold MR, Siegel JE, et al. The role of cost-effectiveness analysis in health and medicine. JAMA 1996;276:1172-7.
14. Weinstein MC, Siegel JS, Gold MR, et al. Recommendations of the Panel on Cost-Effectiveness in Health and Medicine. JAMA 1996;276(15):1253-8.
15. Siegel JS, Weinstein MC, Russell LB, et al. Recommendations for reporting cost-effectiveness analyses. JAMA 1996;276(16):1339-41.
16. Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (Version 4.3). Canberra: Australian Government, Department of Health; December 2008. Available at http://www.pbac.pbs.gov.au/. Accessed February 23, 2015.
17. Guidelines for the economic evaluation of health technologies: Canada (3rd ed). Ottawa: Canadian Agency for Drugs and Technologies in Health; 2006.
18. Guide to the methods of technology appraisal 2013**.** London: National Institute for Health and Care Excellence; 2013. Available at http://publications nice.org.uk/pmg9. Accessed February 3, 2015.
19. WHO guide to cost-effective analysis. Geneva: World Health Organization; 2003. Available at http://www.who.int/choice/publications/p_2003_generalised_cea.pdf. Accessed February 23, 2015.
20. Caro JJ, Briggs AH, Siebert U, et al. Modeling good research practices−−overview: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force−1. Med Decis Making 2012;32:667-77.
21. NHS Choices: Your Health, Your Choices. Website. London: National Health Service. Available at http://www.nhs.uk/conditions/tetanus/pages/introduction.aspx. Accessed February 25, 2015.
22. Russell LB. Technology in hospitals: medical advances and their diffusion (chapter 6). Washington, DC: The Brookings Institution; 1979.
23. Eddy DM. Appropriateness of cervical cancer screening. Gynecol Oncol 1981;12(2 Pt 2):S168-87.
24. Russell LB. Educated guesses: making policy about medical screening tests. Berkeley, CA: University of California Press; 1994.
25. Eddy DM. Screening for cervical cancer. Ann Intern Med 1990;113:214-26.
26. Berkowitz Z, Saraiya M, Sawaya GF. Cervical cancer screening intervals, 2006 to 2009: moving beyond annual testing. JAMA Intern Med 2013;173:922-4.
27. LeFevre ML. Swimming upstream: doing less in health care is hard comment on "No Papanicolaou tests in women younger than 21 years or after hysterectomy for benign disease" and "Cervical cancer screening intervals, 2006 to 2009." JAMA Intern Med 2013;173(10):856-8.
28. NHS Choices: Your Health, Your Choices. Website. London: National Health Service. Available at http://www.nhs.uk/conditions/cervical-screening-test/pages/introduction.aspx. Accessed February 26, 2015.

29. Canadian Preventive Services Task Force. Recommendations on screening for cervical cancer. Available at http://www.cmaj.ca/content/185/1/35 full. Accessed February 26, 2015.
30. Schoen C, Osborn R, Huynh PT, et al. Primary care and health system performance: adults' experiences in five countries. Health Aff 2004;Web exclusive:W4-487-W4-503.
31. Office of Information and Regulatory Affairs, Office of Management and Budget. https://www.whitehouse.gov/omb/oira (Web site). Available at https://www.whitehouse.gov/sites/default/files/omb/assets/a11_current_year/s15.pdf. Accessed February 26, 2015.
32. Robinson LA. How U.S. government agencies value mortality risk reductions. Rev Environ Econ Policy 2007;1(2):283–99.
33. The benefits and costs of the Clean Air Act from 1990 to 2020: summary report. Washington, DC: Environmental Protection Agency, Office of Air and Radiation; March 2011.
34. Robinson LA, Hammitt JK. Behavioral economics and the conduct of benefit-cost analysis: towards principles and standards. J Benefit Cost Anal 2011;2:5.
35. Robinson LA. Personal communication; January 2014.
36. Nichols AL. Lead in gasoline. In Morgenstern RD (ed), Economic analyses at EPA: assessing regulatory impact. Washington, DC: Resources for the Future; 1997.
37. Costs and benefits of reducing lead in gasoline: final regulatory impact analysis. Washington, DC: Environmental Protection Agency; 1986. Available at http://yosemite.epa.gov/ee/epa/eerm.nsf/vwGA/4242532644520FAB8525660400177383. Accessed February 26, 2015.
38. U.S. Congress. Clean Air Act Amendments of 1990. Available at http://www.epw.senate.gov/envlaws/cleanair.pdf. Accessed February 26, 2015.
39. Russell LB. Do we really value identified lives more highly than statistical lives? Med Decis Making 2014;34(5):556-9.
40. Jones-Lee MW, Hammerton M, Philips PR. The value of safety: results of a national sample survey. Econ J 1985;95(March):49-72.
41. Johnson A1, Sikich NJ, Evans G, et al. Health technology assessment: a comprehensive framework for evidence-based recommendations in Ontario. Int J Technol Assess Health Care 2008;25:2:141–50.
42. Guidelines for preparing economic analyses. Washington, DC: U.S. Environmental Protection Agency; 2010. Available at http://yosemite.epa.gov/ee/epa/eerm nsf/vwAN/EE-0568-50.pdf/$file/EE-0568-50.pdf. Accessed February 26, 2015.

Louise B. Russell, PhD, is Distinguished Professor at the Institute for Health, Health Care Policy and Aging Research, and in the Department of Economics, Rutgers University, where her research is focused on the methods and application of cost-effectiveness analysis. Previously, she was a Senior Fellow at the Brookings Institution. Dr. Russell co-chaired the first U.S. Public Health Service Panel on Cost-Effectiveness in Health and Medicine and currently is working with four other leaders in the field who are organizing and facilitating a followup second panel. Dr. Russell was elected to membership in the Institute of Medicine in 1983.



The Contribution of Behavior Change and Public Health to Improved U.S. Population Health

Susan T. Stewart and David M. Cutler

Abstract

Adverse behavioral risk factors contribute to a large share of deaths. We examine the effects on life expectancy (LE) and quality-adjusted life expectancy (QALE) of changes in six major behavioral risk factors over the 1960-2010 period: smoking, obesity, heavy alcohol use, and unsafe use of motor vehicles, firearms, and poisonous substances. These risk factors have moved in opposite directions. Reduced smoking, safer driving and cars, and reduced heavy alcohol use have led to health improvements, which we estimate at 1.82 years of quality-adjusted life. However, these gains were roughly offset by increased obesity, increased firearm deaths, and increased deaths from poisonous substances, which together reduced QALE by 1.77 years. We model the hypothetical effects of a 50 percent decline in morbid obesity and in poisoning deaths and a 10 percent decline in firearm fatalities, roughly matching favorable trends in smoking and increased seat belt use. These changes would lead to a 0.92 year improvement in LE and a 1.09 year improvement in QALE. Thus, substantial improvements in health by way of behavioral improvements and public health are possible.

Introduction

While health is often thought of in terms of diagnosed medical conditions, it is modifiable behavioral risk factors such as obesity and smoking that account for the largest portion of deaths each year.[1,2] For example, Mokdad et al.[1] report that in 2000, 43 percent of total deaths were accounted for by six behavioral risk factors: smoking (18 percent), obesity (17 percent), alcohol consumption (4 percent), motor vehicle accidents (2 percent), firearms (1 percent), and illicit drug use (1 percent).

In addition to affecting mortality, these modifiable behavioral factors have significant effects on health-related quality of life (HRQOL).[3-6] Smoking, heavy alcohol use, and obesity have been causally linked to a myriad of diseases and symptoms,[6-8] and injuries from motor vehicle accidents and firearms can be severe. When feasible, considering this nonfatal impact can provide a more comprehensive picture of the health outcomes attributable to these factors.

Behavioral and public health interventions have been very successful in reducing the harms of some of these factors.[1,9] Deaths from smoking and (to a lesser extent) motor vehicle accidents have declined markedly over time.[9] Other factors, such as obesity[1] and drug overdose,[10] have worsened.

On net, there is little understanding of how these risk factor changes as a whole have contributed to U.S. health trends.

In this chapter, we will examine the effects of changes in six major behavioral factors on HRQOL and mortality from 1960 to 2010: smoking, obesity, heavy alcohol use, and unsafe use of motor vehicles, firearms, and poisonous substances. For the factors that have improved, we evaluate how much improvement there has been in length and quality of life. For those that have deteriorated, we simulate the hypothetical gain achievable through effective public health interventions and behavioral change targeting these factors.

In each case, we try to differentiate medical from non-medical changes. In the case of smoking, for example, cigarette consumption has declined by over half, and medical care has extended life for people with cardiovascular disease.[11] Our primary focus is on the non-medical changes. We separate the medical from the non-medical changes as we are able to do so. Unfortunately, we generally do not have enough evidence to differentiate amongst sources of non-medical improvement or decrement: behavioral health, public health measures, or other causes.

Methods

For each of the six behavioral risk factors that we considered—smoking, obesity, heavy alcohol use, and unsafe use of motor vehicles, firearms, and poisonous substances—our goal was to measure changes in HRQOL, including both mortality and morbidity, over the past several decades.

The data sources and methods for each of the risk factors that we considered are shown in Table 1. For smoking, obesity, and alcohol use, we had consistent measures of prevalence over time and could examine the effects of both fatal and nonfatal exposures. For these conditions, there were multiple steps to the analysis: measuring the historical trend in the prevalence of the risk factor and evaluating how it has affected both length and quality of life. We describe each of these further below.

For motor vehicle accidents, firearms, and accidental poisoning, there are multiple underlying behaviors and factors that can affect a person's risk. For example, in the case of motor vehicles, these can include reckless, distracted, or impaired driving; failure to use a seat belt; and the safety features of roads and vehicles.[12] For firearms, they include factors such as secure storage, locking devices, safety training, community-based prevention, identification of individuals at risk of misuse, and laws regarding firearm acquisition and possession.[13,14] These factors can be difficult or impossible to track reliably over time, and they can affect both the risk of an incident and the extent of resulting injury. In addition, for nonfatal motor vehicle accidents, firearm injury, and drug misuse, HRQOL is difficult to account for due to a lack of consistent data on non-fatal outcomes and to changes over time in the number and severity of injuries. Thus, as a measure of the effect of these risk factors over time, we used mortality data from U.S. vital statistics.[15] Motor vehicle deaths include those resulting from collisions of all types of road vehicles, including collision of these vehicles with pedestrians and bicycles. For firearm deaths, we considered homicides and suicides. In the accidental poisoning category, the vast majority of deaths since the 1980s were due

to unintentional overdose of prescription or illicit drugs; the remainder resulted from exposures to other substances such as alcohol, noxious fumes, and ingested chemicals. Since we used data on mortality only, we omitted quality of life estimates for these factors.

Table 1. Data sources and methods for risk factors considered

	Mortality and HRQOL			Mortality		
	Obesity	Smoking	Alcohol	Motor Vehicles	Firearms*	Accidental Poisoning**
Data for prevalence	NHANES, NHES	NHIS	NHANES			
Data for impact on mortality	NHANES mortality followup					
Data for actual deaths	n/a			Vital statistics		
Data for impact on HRQOL	MEPS	MEPS	NHIS	n/a		
Methods	Public health affects prevalence. Effects of risk factor on Mortality and HRQOL measured in national data and held constant over time.			Public health affects both prevalence and the effects of risk factor on mortality over time; actual mortality used. HRQOL difficult to account for due to changing pool of nonfatal exposures.		

*Suicide and homicide only. **Includes all poisonous substances, though drugs accounted for most deaths.

Note: HRQOL = health related quality of life; NHANES = National Health and Nutrition Examination Survey; NHES = National Health Examination Survey; NHIS = National Health Interview Survey; MEPS = Medical Expenditure Panel Survey

Motor Vehicle Accidents, Firearms, and Poisoning—Trends and Assessment

We began by measuring trends in mortality from motor vehicle accidents, firearms, and accidental poisoning for all years from 1960 through 2010. To adjust for changes in the definition of motor vehicle deaths in the different versions of the International Classification of Diseases used over this period (ICD-7, 8, 9, and 10), we used comparability ratios provided by the National Center for Health Statistics (NCHS).[16-18] Comparability ratios were not provided for the subcategories of firearm suicide and homicide, however these deaths are more straightforward to measure over time, and comparability ratios for the broader categories of suicide and homicide were close to 1. For poisoning, comparability ratios were not provided by NCHS due to insufficient sample sizes for stable estimates. However, specific categories of poisoning match well over time. Thus, we followed the method of a prior study examining child poisoning,[19] using accidental poisoning by solid and liquid substances in all years and excluding deaths from poisonous vapors and gases, which includes carbon dioxide poisoning.

Modeling the Portion of Change Due to Medical Care

An important consideration when using mortality as a proxy for exposure to these risk factors is that there are medical contributions to survival. Advances in trauma care have reduced the probability of death from motor vehicle collisions, drug overdoses, and shootings. The proportion of population health improvement attributable to medical care is unknown, with varying estimates derived or used in the literature, including 10, 20, and 50 percent.[20-22] The proportion also depends on whether improved access to medical care is counted as part of medical improvement or as public health improvement. For example, improvements in emergency response such as "enhanced 911" (which provides the precise location of callers) have sped up response times and improved patient survival.[23,24]

To address this uncertainty, we used the approximate median of the existing estimates and assumed that 25 percent of the mortality improvements for motor vehicle accidents and poisonings have resulted from medical care. (We did not assume a medical contribution to reduced firearm deaths, since gun suicides are nearly always fatal and account for more than half of firearm deaths. The analysis did not allow estimation of a contribution of medical care to only the homicide portion of deaths in the firearm category.) We performed sensitivity analyses using values of 10 percent and 50 percent for the contributions of medical care to mortality change, with any remaining improvement attributable to behavioral and public health factors. In the case of motor vehicle accidents, these include impaired driving prevention and surveillance, seat belt and child safety seat use, graduated driver licensing, vehicle and road safety advances, and many more. For prevention of drug overdose, interventions are being designed at the Federal, State, local, and insurance company levels to address and prevent dangerous prescription drug addictions.[10,25-27]

Prevalence and Effects of Smoking, Obesity, and Heavy Alcohol Use

To measure the historical distribution of body mass index (BMI), we used physical measures of height and weight from the National Health and Nutrition Examination Survey (NHANES),[28] a comprehensive national health survey that combines interviews with physical examinations. We also used its precursor, the National Health Examination Survey (NHES). Respondents were classified using World Health Organization (WHO) criteria[29] as underweight (BMI < 18.5 kg/m^2), normal weight (18.5 to 24.9 kg/m^2), overweight (25.0 to 29.9 kg/m^2), obese (30.0 to 34.9 kg/m^2, WHO obesity class I), or morbidly obese (≥ 35.0 kg/m^2, WHO obesity classes II and III).

Sample size considerations did not permit exact estimation of BMI by single year of age, which is needed for the mortality estimates. To form estimates by single year of age, we used regression analysis to assess the likelihood of being in each BMI category based on age, gender, race, and their interactions, and predicted the distribution of categories by year of age for each time period covered by the data: 1959-1962 (NHES), 1971-75 (NHANES I) 1976-1980 (NHANES II), 1988-1994 (NHANES III), and the continuous NHANES (2000-2010, in 2-year cycles). In essence, this just smooths the observed data across nearby ages.

To measure the historical distribution of current and former smoking, we used data from the National Health Interview Survey (NHIS), an ongoing health survey of the U.S. civilian non-institutionalized population.[30] Current smokers and former smokers were defined as those who had

ever smoked at least 100 cigarettes and who still smoked, or had quit, respectively. Smoking rates were predicted by year of age in 1965 (the first year in which smoking questions were asked in the NHIS) and 2010. To estimate smoking rates by single year of age, we used the same smoothing process as described above for BMI.

We measured the prevalence of alcohol use in NHANES I (1971-75), NHANES III (1988-1994) and in the continuous NHANES (2000-2010, in 2-year cycles). (Alcohol was not asked about in the 1959-62 NHES, and we did not use NHANES II due to a difference in alcohol measurement in that year.) We defined heavy alcohol use as 15 or more drinks per week for men and 8 or more drinks per week for women, consistent with the Centers for Disease Control and Prevention (CDC) definition.[31] As with BMI and smoking, smoothed heavy alcohol use rates were predicted by year of age from a regression of heavy drinking on age, gender, race, and their interactions.

To measure how each of these risk factors affects length of life, we estimated Cox proportional hazard models relating each risk factor to subsequent all-cause mortality, controlling for age, race, gender, and their interactions. For this we used data from the combined NHANES I, II, and III surveys, matched to subsequent death records (through 1992 for NHANES I and II and 2006 for NHANES III).

To estimate the impact of BMI and smoking on HRQOL, we used data from the 2002 Medical Expenditure Panel Survey (MEPS),[32] which measures health status, medical expenditures, and socioeconomic characteristics in a sample nationally representative of the U.S. community population. Following a method that we previously developed,[33] we related each risk factor separately to a 100-point rating of overall health (the 'visual analog scale' from the Euroqol EQ-5D[34] health measurement instrument), controlling for age, gender, and race. From this we predicted HRQOL weights for each risk factor by 10-year age group.

To estimate the effects of heavy alcohol use on HRQOL, we used data from the 2010 NHIS. The NHIS does not contain a 100-point self-rating of health, but it does include a large number of impairments and symptoms that enabled the estimation of HRQOL scores using techniques that we developed in previous work.[33,35] First, we identified the full set of impairments and symptoms asked in both NHIS and MEPS: depressive symptoms, anxious symptoms, and limitations in primary activity (e.g. work), self-care (activities of daily living), routine needs (instrumental activities of daily living), walking, bending, standing, dexterity, vision, and hearing. We then regressed the 100-point rating of overall health in MEPS on this set of impairments and symptoms. Using these regression coefficients, we calculated predicted HRQOL scores using this same set of impairments and symptoms in the NHIS data. Previous analyses[33] supported holding constant the impact of each impairment and symptom over time, since the effect of a particular impairment on HRQOL remains relatively stable over time, whereas its prevalence can change more rapidly.

Life Expectancy and QALE Calculation

To calculate life expectancy at age 18, we used U.S. life tables for the base year, 1960 (1973 for alcohol).[36,37] To measure change in life expectancy through 2010 due to each risk factor, we used the 2010 mortality rate for that risk factor, otherwise holding life expectancy constant at the 1960 level (1973 for alcohol). This provided an estimate of the life expectancy change that would have occurred

if only this risk factor had changed. For obesity, smoking, and alcohol, the 2010 mortality rate was the product of the relative risks of mortality associated with each risk factor and the predicted share of the population with the risk factor at each year of age in that year.

To calculate quality-adjusted life expectancy (QALE), mean HRQOL for each 10-year age group was used to adjust remaining life expectancy at each age. For motor vehicle accidents, firearms, and poisoning, where we were able to measure mortality but not quality of life, we estimated a QALE value using the rough assumption that HRQOL remained constant over time. Since not all years are in perfect health, this reduces the quantity of quality-adjusted life years relative to life years. To place a dollar value on QALE, we valued each year of life at $100,000 and used a 3 percent discount rate for future years lived.

Finally, for those factors found to have worsened rather than improved over time, we performed simulations of the potential benefit of successfully curbing these problems. Based on what was achieved for factors that improved, we modeled the effects of similar improvements for the factors that did not improve, simulating the hypothetical effect of this progress.

Results

Trends in Behavioral Risk Factors

Of the six factors that we examined, three improved: smoking, motor vehicle deaths, and alcohol use. The others – obesity, firearms, and poisoning – worsened over time. Table 2 shows the percentage change in prevalence or the deaths from each risk factor over time, age-adjusted to the 2000 population. Smoking prevalence and motor vehicle deaths showed the largest declines, while obesity and poisoning showed the greatest increases.

Table 2: Percent change in each risk factor

Percent Change in Prevalence, 1960-2010*	
Smoking	-54
Heavy Alcohol Use	-22
Obesity (BMI ≥ 30)	161
Percent Change in Deaths, 1960-2010**	
Motor Vehicles	-46*
Firearm Suicide Firearm Homicide	4 25
Accidental Poisoning (primarily drug overdose)	1005

Notes: *Heavy alcohol use change is from 1973-2010. **Accounting for the increase in miles driven over this time period, motor vehicle deaths would have increased by 234 percent if the death rate per mile had not changed. Percent change in each factor using rates that are adjusted to the age distribution of the population in the year 2000, to account for changes in the age distribution of the population over time.

Figure 1 (a-f) shows the trends in each factor, age-adjusted to the year 2000 population. Smoking is the leading preventable cause of death; a regular smoker has a life expectancy that is about 7 years below that of a never smoker.[7] Figure 1a shows the smoking rate among U.S. adults over time. The rate declined by more than half, from 42 percent in 1965 to 19 percent in 2010.[38]

Motor vehicle accident deaths have also declined over time, as shown in Figure 1b. Unadjusted, the decline was 46 percent. However, more people had cars over time, and vehicle miles driven increased markedly. Thus, Figure 1b also shows a counterfactual mortality rate assuming no change in deaths per mile during this time of increased driving. That mortality rate is forecast to have increased by 234 percent. Thus, the net reduction in motor vehicle mortality adjusted for miles driven was 280 percent. This is believed to result from many successful public health interventions, including safer automobiles and roads and enforcement of seat belt laws, motorcycle helmet laws, graduated driver licensing, and impaired driving laws and penalties.[12,39]

Heavy use of alcohol has declined over time (Figure 1c). The change from 1971-75 to 2009-10 was -22 percent in relative terms, or 3 percentage points. Given the relatively low prevalence of heavy drinking and small change in its prevalence, this did not contribute a great amount to population health.

Obesity has increased over time (Figure 1d); among behavioral risks, this is the major contributor to worse health. The share of the population that is obese or morbidly obese increased from 14 percent in 1959-62 to 36 percent in 2009-10. The increase in obesity is attributed to many causes, including reductions in the time cost of food preparation.[40]

Firearm suicide and homicide (Figure 1e) increased into the 1970s and declined in the 1990s. Firearm homicides also declined in the mid-1970s and 1980s but rose again in the late 1980s and early 1990s. This rise is largely attributed to the crack cocaine drug epidemic at that time.[41] Firearm homicides have declined by nearly 50 percent since that peak in the 1990s.[42] Firearm homicides changed little through the 2000s and declined in the most recent years. Firearm suicides declined slightly in the mid-2000s, with an uptick in the most recent years. The net change over the past half-century was a slight increase 0.7 percentage points for gun homicide and 0.2 percentage points for gun suicide.

Finally, deaths due to poisoning (Figure 1f) increased dramatically over the past 20 years. The vast bulk of this is overdose of drugs, particularly prescription opioid medications, which has more than quadrupled since 1999.[10,43] This increase overwhelmed a decline in infant and child poisoning since the 1960s (findings by age not shown). Poisoning deaths do not include intentional suicides or deaths due to accidents of undetermined intent, which are measured separately by the CDC. While the dramatic rate of increase in drug deaths slowed slightly from 2006 to 2010, during this time the rate of alcohol poisoning suddenly increased among all age groups. Explanations for this include increased frequency of binge drinking during the Great Recession of late 2007 to mid-2009[44] and increased use of caffeinated energy drinks containing alcohol; these energy drinks increase consumption of alcohol during binge drinking by masking the depressive effect of alcohol.[45]

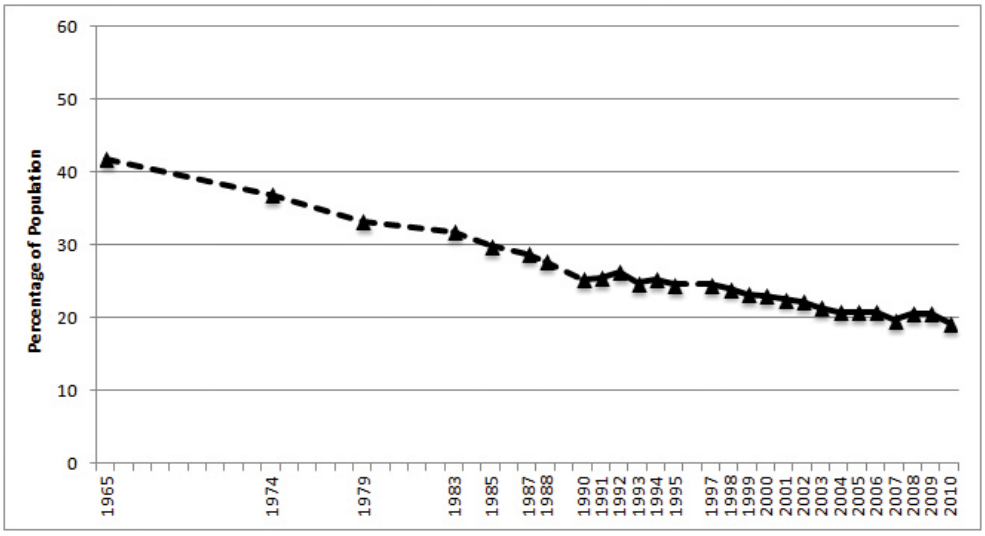


Figure 1a: Trend in current smoking from NHIS data, age 18+, 1965-2010

Note: Current smokers were those who had ever smoked at least 100 cigarettes and still smoked, as reported in the National Health Interview Survey, National Center for Health Statistics. Data directly from National Center for Health Statistics. Health, United States, 2011: With Special Feature on Socioeconomic Status and Health. Hyattsville, MD. 2012. Rates are age-adjusted to the 2000 population.

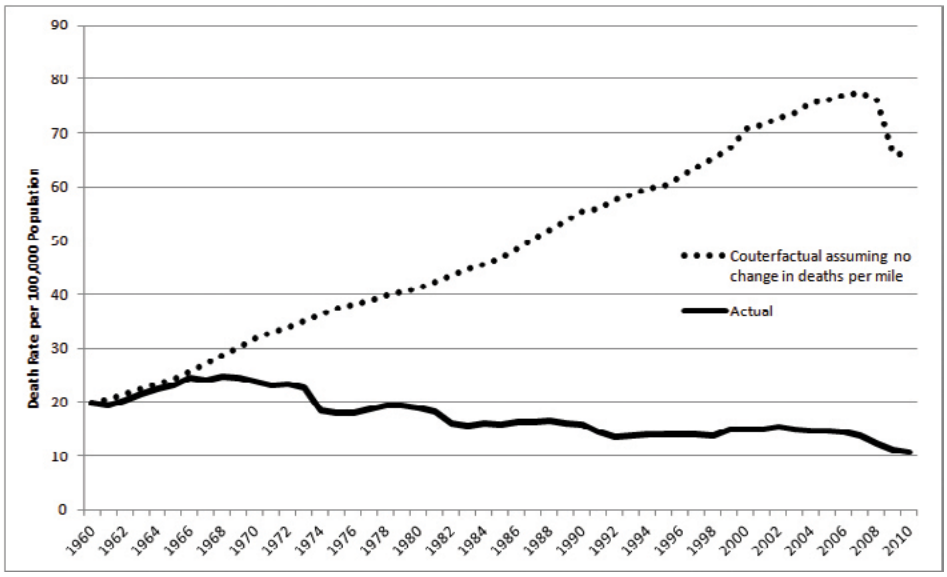


Figure 1b: Trend in mortality rate from motor vehicle accidents, 1960-2010

Note: Mortality rates from U.S. vital statistics (National Center for Health Statistics[15]) are adjusted for changes in the age distribution of the population over time and for changes to the definition of motor vehicle accident (MVA) mortality across ICD (International Classification of Diseases) versions 7-10. Includes deaths resulting from collisions of all types of road vehicles, including collision of these vehicles with pedestrians and bicycles. The counterfactual trend reflects the increase in deaths that would have occurred if the death rate per mile had remained at the 1960 rate while driving rates (number of miles driven) increased over time. The drop in this trend in 2007-2009 reflects a decline in miles driven, coinciding with the Great Recession.

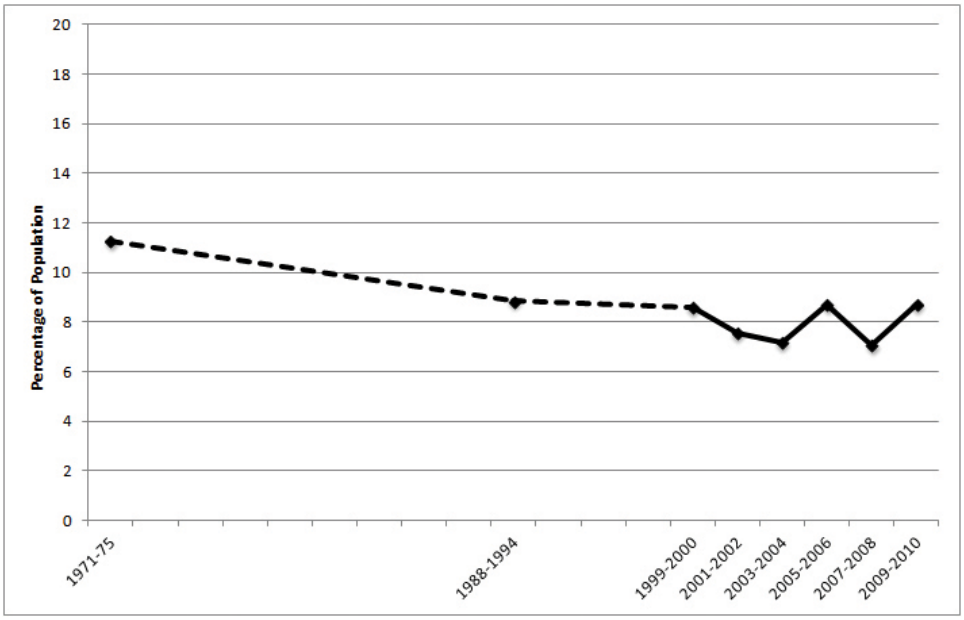


Figure 1c: Trend in heavy alcohol use, age 25+, NHANES 1973-2010

Note: Heavy alcohol use is defined as 15+ drinks/week for men and 8+ drinks/week for women, as reported in the National Health and Nutrition Examination Survey, National Center for Health Statistics. Rates in each year are age-adjusted to the 2000 population. Note difference in scale compared to Figures 1a and 1d.

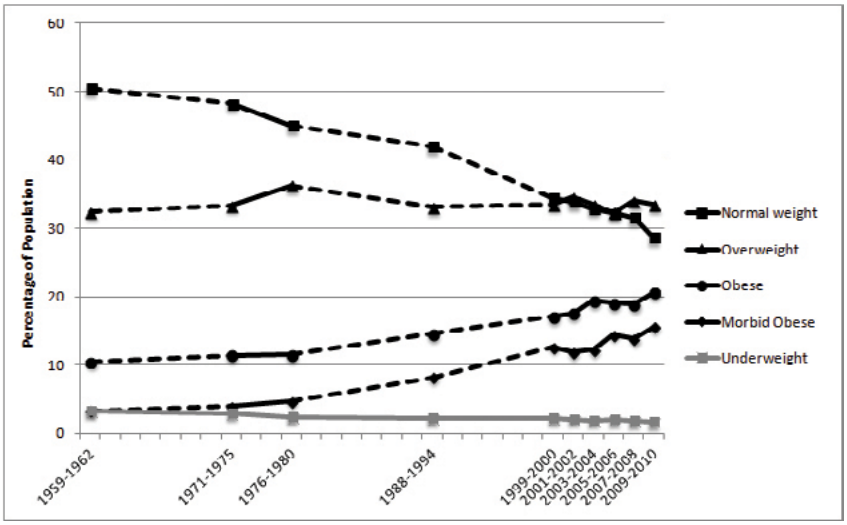


Figure 1d: Trends in body-mass index distribution, NHES/NHANES, 1960-2010

Note: Categories of body-mass index (BMI, the weight in kilograms divided by the square of the height in meters) were defined according to the World Health Organization criteria: underweight (BMI <18.5), normal weight (BMI 18.5 to 24.9), overweight (BMI 25.0 to 29.9), obese (BMI 30.0 to 34.9; obesity class I), and morbidly obese (BMI ≥35.0; obesity classes II and III). Underweight is shown in figure but not used in life expectancy (LE) and quality-adjusted life expectancy (QALE) calculations because low BMI can be indicative of pre-existing illness.[46] Rates in each year are age-adjusted to the 2000 population. From the National Health Examination Survey and the National Health and Nutrition Examination Survey, National Center for Health Statistics.

Figure 1e: Trend in mortality rate from firearm suicide and homicide, 1960-2010

Note: Includes intentional self-harm and assault using all types of firearms. Rates in each year are age-adjusted to the 2000 population. Note difference in scale compared to Figure 1b. Mortality rates from U.S. vital statistics, National Center for Health Statistics.[15]

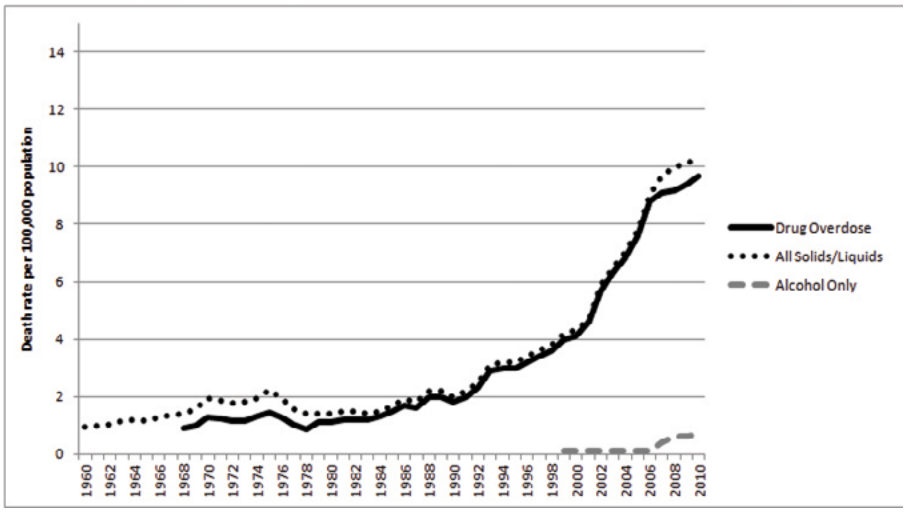


Figure 1f: Trend in mortality rate from accidental poisoning, 1960-2010

Note: Drug overdose includes accidental death from prescription, over-the-counter, and illicit "drugs, medicaments and biological substances."[47] The broader category of deaths due to accidental "poisoning by and exposure to all solids and liquids" includes these drug substances as well as alcohol, pesticides, "organic solvents and halogenated hydrocarbons and their vapours," and "other and unspecified chemicals and noxious substances," but excludes deaths due to "other gases and vapors"[48] Rates in each year are age-adjusted to the 2000 population. Note difference in scale compared to Figure 1b. Mortality rates from U.S. vital statistics, National Center for Health Statistics.[15]

Changes in QALE

Table 3 shows relative risks of all-cause mortality by smoking status, BMI status, and alcohol consumption calculated using separate Cox proportional hazards models for each risk factor. Current smokers had the greatest relative risks of death, followed by those who were morbidly obese and heavy drinkers. The effects on HRQOL of smoking, BMI category, and heavy alcohol use are shown in Table 4. As expected, mean quality of life was worse for current smokers than never smokers. Those in the morbid obese category had HRQOL similar to smokers. The HRQOL of those in the obese category was higher than for smokers but not as high as that of those in the normal weight or overweight categories. Heavy alcohol users had HRQOL scores essentially the same as non-heavy users.

Table 3: Hazard ratios: Relative risks of death from all causes among respondents aged 25-70 in NHANES I, II, and III linked mortality data

		Attained age under 65	Attained age 65+
Body Mass Index	Normal weight	1.00	1.00
	Overweight	0.76	0.87
N= 35,542 (50% age 65+)	Obese	1.08	1.05
Deaths: 1,340/ 8,609	Morbid obese	1.77	1.50
Smoking			
	Never smoked	1.00	1.00
N = 29,408 (51% age 65+)	Former smokers	1.08	1.26
Deaths: 1,097 / 6,666	Current smokers	2.51	2.52
Alcohol			
	Non-Heavy Drinkers	1.00	1.00
N = 22,985 (44% age 65+)	Heavy Drinkers	1.74	1.42
Deaths: 990 / 5,393			

Notes: body mass index (BMI) was based on measured weight and height and categorized according to the World Health Organization criteria, as described in note for Figure 1d. Current smokers were those who had smoked at least 100 cigarettes in their lifetime and smoked now. Heavy alcohol use was defined as 15+ drinks/week for men and 8+ drinks/week for women. Separate regression models were fit for deaths that occurred before and at/after age 65. Deaths among those under age 65 were left censored in the first model, and deaths among age 65+ were right censored in the second. (The SAS entry= option was used with proc phreg) Fifty percent of the sample in our analyses attained age 65 or greater. Those lost to followup were assigned a survival time of ½ the possible survival interval before the time when they were lost but were censored on mortality. Models also included race: white (including Hispanic), black, and other, and age in 5-year age groups (coefficients not shown). These analyses omit those under age 25 and those who died within the first 4 years of followup. BMI analyses excluded those who were underweight at baseline because this can be indicative of pre-existing illness.[46] Analyses used the National Health and Nutrition Examination Surveys I, II, and III, linked to the National Death Index (NDI).

Table 4: Effects on health-related quality of life of smoking, body mass index, and heavy alcohol use

	Mean predicted HRQOL Score
Normal weight	0.85
Overweight	0.82
Obese	0.79
Morbid obese	0.74
Nonsmokers	0.85
Current smokers	0.77
Non-Heavy Drinkers	0.828
Heavy Drinkers	0.825

Notes: HRQOL = health-related quality of life. The effects of BMI and smoking were assessed in the 2002 Medical Expenditure Panel Survey (MEPS), and the effect of heavy drinking was assessed in the National Health Interview Survey (NHIS) 2010. BMI was based on self-reported weight and height and categorized according to the World Health Organization criteria, as described in note for Figure 1d. Smokers were those reporting that they currently smoked. Heavy alcohol use was defined as 15+ drinks/week for men and 8+ drinks/week for women.

The changes in life expectancy (LE) and QALE for each risk factor that we estimate to be attributable to public health interventions and behavioral change are shown in Table 5, which separates factors into those that improved and those that worsened and ranks them by the magnitude of these effects. Reduced smoking has had the largest effect, improving life expectancy by 1.26 years and QALE by 1.42 years. Reductions in motor vehicle fatalities affected fewer individuals but also had an important effect of improving population life expectancy, by 0.43 years. To put the magnitude of these improvements in context, the total life expectancy improvement at age 18 over the past half-century was about 7 years; we do not have a comparable QALE estimate because quality of life data are not available for 1960. In terms of life expectancy, therefore, the factors that improved LE accounted for 25 percent of the total. Valuing the total effect of the factors that improved in dollars, the QALE gain is over $65,000 per person over the course of his or her adult life.

Of the factors that worsened, obesity had the largest effect, reducing LE by 1 year and QALE by over a year and a half. Increases in accidental poisoning deaths held back LE improvement by over a quarter of a year. Accounting for the contribution of medical advances to improved survival of poisoning victims over time means that poisoning events occurred even more frequently than the mortality data indicate. In this case because the effect of behavioral change for poisoning was negative, we accounted for the improvement in medical care by increasing the behavioral effect to 125 percent. Firearm suicide and homicides, which worsened slightly overall, held back LE by a small amount, 0.03 years. The total dollar value associated with QALE not gained due to the three worsening factors combined was about $64,000; only slightly less than the $65,000 gained from factors that improved.

Table 5: Change in life expectancy and quality adjusted life expectancy at age 25 attributable to behavioral risk factors between 1960 and 2010

Risk Factor	Change (years) LE	QALE	Value of Improvement	
Overall	6.9			
Areas of Improving Health:				
Smoking	1.26	1.42	$32,495	$52,011
Motor Vehicle Accidents*	0.43	0.34†	$15,241	$12,450
Alcohol‡	0.06	0.05	$1,065	$967
Total	**1.75**	**1.82**	**$48,801**	**$65,428**
Areas of Declining Health:				
Obesity	-1.00	-1.53	-$23,685	-$57,950
Poisoning*	-0.26	-0.21†	-$8,811	-$7,171
Firearms	-0.03	-0.03†	-$1,677	-$1,401
Total	**-1.30**	**-1.77**	**-$34,173**	**-$66,522**

*Excluding 25% of improvement assumed to be attributable to improvements in medical care.
†Assuming no change in HRQOL; QALE change reflects only the LE change.
‡Alcohol measured from 1973-2010.
Note: LE = life expectancy; QALE = quality-adjusted life expectancy; HRQOL = health-related quality of life.

The results of our sensitivity analyses on the proportion of mortality improvement attributable to medical care are shown in Table 6. Whether medical care accounts for as little as 10 percent or as much as 50 percent of the improvement, the remaining improvement in motor vehicle mortality attributable to public health and behavioral change remains large, accounting for a gain in life expectancy between one-half and one-third of a year from 1960 to 2010. For poisoning deaths, the extent of the public health problem varied little with the magnitude of the assumed medical benefit; poisoning held back LE improvement by between 0.23 and 0.31 year, depending on the extent to which medical care may have prevented additional overdose deaths. The total net effect of public health and behavioral change on life expectancy for all six factors was estimated to vary between 0.57 and 0.25 year, depending on the assumed effect of medical care in preventing motor vehicle and poisoning deaths.

Table 6: Sensitivity analyses on the proportion of change in life expectancy at age 25 due to medical care improvement between 1960 and 2010

	Life Expectancy Improvement		
Risk Factor	Point Estimate	Lower Bound	Upper Bound
	25%	10%	50%
Motor Vehicle Accidents	0.43	0.52	0.29
Poisoning	-0.26	-0.23	-0.31
Total net effect of all six factors	0.45	0.57	0.25

Note: For motor vehicle deaths, we model results assuming that a portion of the observed mortality improvement (25%, 10%, and 50%) was due to medical care. For poisoning deaths, we assume that mortality would have worsened even more than observed if not for the effect of medical care; thus we model an increase in the effect of behavioral change and public health interventions (125%, 110%, and 150%). Last row shows the net effect on life expectancy change of all six risk factors we considered.

Impact of Reversing Trends

Table 7 shows the results of simulated improvements for obesity, poisoning, and firearm suicide/homicide. These were hypothetical scenarios in which we modeled the effect of progress that might be possible if public health and behavioral changes could be achieved in these areas at levels similar to those seen with smoking or motor vehicle accidents. We did not assume that we could completely reverse the negative factors, returning the population to 1960 rates, since this would be unrealistic; looking back at the dramatic reduction in smoking, it was far from completely eliminated. Rather, it was reduced by about half. Thus, we chose to model a 50 percent decline in morbid obesity (and a corresponding increase in normal weight), since morbid obesity was the major driver of ill effects of high BMI. Similarly, we modeled a 50 percent reduction in accidental poisoning (primarily drug overdose) deaths.

In estimating what would be a comparable rate of progress for firearm fatalities, we considered firearm safety as approximately equivalent to convincing people to wear a seat belt. The literature on motor vehicle interventions estimates that primary seat belt laws (allowing ticketing for not wearing a seat belt with no other traffic offense) resulted in about a 10 percent reduction in motor vehicle deaths.[39] We thus modeled the effect of a hypothetical firearm intervention with a similar impact. The positive effect on overall life expectancy was fairly small, primarily because firearm morality had already declined since its peak in the 1980s and 1990s.

As Table 7 shows, these simulated improvements yielded 0.92 year of improved life expectancy together and 1.09 years of increased quality-adjusted life expectancy. These are substantial. For example, the gain from reducing obesity would be over half of the gain from reduced smoking since 1960. The gain in reduced poisoning deaths would be about one-third of the gain from reduced motor vehicle deaths since 1960.

Table 7: Effects of hypothetical reductions in 2010 mortality rates from obesity, poisoning, and firearms

Risk Factor	Improvement in 2010 vs actual		Scenario	
	LE	QALE		
Obesity	0.76	0.96	50% reduction in morbid obesity	(comparable to
Poisoning	0.14	0.11*	50% reduction in deaths	smoking decline)
Firearms	0.03	0.02*	10% reduction in deaths	(comparable to seat-belt use)
Total	**0.92**	**1.09**		

* Assuming no change in HRQOL; QALE change reflects only the LE change.

Note: LE = life expectancy; QALE = quality-adjusted life expectancy; HRQOL = health-related quality of life.

Discussion

Declines in just two factors—smoking and motor vehicle accident deaths—account for a substantial portion (one-quarter) of the improvement in population life expectancy over the past half century, demonstrating the important impact of behavior change and public health interventions. Though the mortality decline from reduced smoking is well known, we also quantify the effect of this decline on nonfatal health.

Of the other factors that we consider, slight reductions in heavy alcohol use also had a small positive effect, while slight increases in firearm suicide and homicide had a negative effect on population health. Substantially holding back population QALE improvement over the past 50 years were increases in obesity and drug overdose deaths. Our simulations indicate the extent of improvement that could be achieved with effective public health interventions to address the worrisome trends in these harmful factors, if we can repeat in these areas the success we've had in others. Our simulation suggests that progress on obesity and accidental poisoning comparable to that made on smoking in the past half century would have an effect on health almost three-fourths the size of the large (1.26 year) effect of the reduction in smoking on life expectancy.

Implications for Practice

Following the examples of the various effective interventions in the areas of tobacco,[7] motor vehicle accidents,[39] and child poisoning,[19] Hemenway and colleagues[13,14] have proposed many analogous ways in which similar progress could be made in reducing firearm deaths, including regulation, taxation, safety training, and identification and counseling of at-risk individuals. In the case of accidental poisoning, rapidly growing abuse of prescription medications, which is also known to serve as a gateway to the use of illegal drugs, is currently gaining widespread attention, garnering efforts to better understand and address the problem, in part by detecting and reducing over-prescription of the most problematic medications.[10,25,49] The recent increases in alcohol poisoning deaths underline the importance of increased efforts to reduce binge drinking and discourage the mixing of alcohol with energy drinks.[45]

Prevention of obesity, particularly among children, has been a key focus in recent years, but it has been difficult to combat the roots of obesity, including sedentary lifestyles, insufficient physical activity, widespread availability of high-calorie food in large portions, and reduced time available for food preparation in the home[50-54] Those with low incomes face unique challenges in adopting healthful behaviors, including reduced or inconsistent access to affordable healthy food of good quality, greater availability of fast food restaurants, reduced physical activity resources, reduced access to quality health care, and greater exposure to advertising of obesity-promoting products.[55] Continued efforts to address these problems will be essential in order to reduce death and disability resulting from high BMI.

Among those already obese, the clinical efficacy of a number of weight loss interventions has been demonstrated,[56] and there is encouraging evidence that even modest weight loss can have important impacts on public health.[57] The challenge is to increase the use of these interventions, as well as ensure ongoing support to help individuals maintain weight loss in the long-term. Complementary interventions include those affecting food consumption (e.g. nutrition labeling, taxation, and subsidy

of specific foods[57] or nutrients[58]) and increasing access to healthy foods and exercise opportunities within communities.[59]

Alcohol use can also be reduced through taxation and restrictions on alcohol availability.[60] However, in order to further combat the effects of heavy and binge drinking, these strategies must be combined with others that are known to be effective, including information dissemination, early intervention by primary physicians, and behavioral and pharmacological interventions to treat alcohol dependence.[60]

Finally, continued progress is crucial in the areas where we have seen success; increasing the use of proven methods to prevent youth smoking and achieve smoking cessation,[7] and broadening the adoption of laws and practices that improve the safety of vehicles, roads, drivers, passengers, pedestrians, and those riding bicycles.[12] Measures to address the safety of older drivers and pedestrians will be increasingly important as the population ages.[12] Promising new technological interventions to save lives in motor vehicle accidents include features to help first responders more quickly access vital information, such as Google Glass,[61] vehicles with the ability to automatically contact emergency workers in the case of a collision, and quick response (QR) codes that provide internal vehicle diagrams to emergency rescue workers.[62]

Implications for Research

Our findings underline the importance of research to understand the roots of behavioral health problems and develop successful interventions and implementation strategies. The dramatic lifesaving effect of prior research is demonstrated most poignantly for smoking and motor vehicle fatalities.[7,39] The skyrocketing rates of poisoning deaths, driven in particular by prescription drug overdose, point to an urgent need for research on the prevention and treatment of prescription drug addiction. Additional research on the impact of firearm interventions, which has lagged behind research on other behavioral causes of death, will also be critical to addressing this social problem. Finally, continued research is required to understand specific pathways to the development of obesity, heavy alcohol use, and binge drinking habits and to test the most effective ways to broaden the use of strategies already known to be effective in reducing the harms of all of these behavioral risk factors.

Limitations

Our study has some limitations. While it quantifies the effects of some major causes of mortality and morbidity, it does not account for the many other factors that determine life expectancy and HRQOL, such as genetics, environmental toxins, and stress. Also, the impact of different BMI levels on mortality is a matter of ongoing debate.[63] However, we have calculated our relative risks of mortality directly in nationally representative data and have found our past results to be robust to sensitivity analyses using other published relative risks.[3,46,64]

We measured HRQOL associated with obesity, smoking, and heavy alcohol use at a single point in time, since quality-of-life data were not available over our extended time period of analysis. While we would not expect the ill-effects of these behaviors to change nearly as rapidly as their prevalence, the morbidity associated with them may have been worsened or improved by various

factors, including intensity of smoking and alcohol use, diet quality and activity level, and improved medical treatment of diseases associated with behavioral risk factors.[4,6,65,66]

Finally, our study uses the period life table approach, which assumes that as cohorts age, the distribution of their risk factors will change to reflect the distribution among people currently living at these older ages. However in the case of BMI, though obesity prevalence may be stabilizing,[50,67] population BMI levels will rise with the aging of current younger cohorts, among whom obesity rates are (or were recently) at a historical high.[67] The effects of obesity on health occur primarily via disease, and the prevalence of diabetes in particular is at an all-time high and continues to increase. Our prior work forecasting alternate future scenarios concluded that the obesity increases would outweigh the benefits of smoking declines.[3] This is where medical care can have an important effect; an encouraging finding is the improved control of cardiovascular risk factors such as high cholesterol and hypertension, particularly among those who are overweight and obese.[66] More recently, a sharp decrease in complications has been found among those with diabetes.[67] Improved treatment of obesity-related diseases is thus vital to improving the Nation's health.

Conclusion

The benefits of behavioral and public health advances, where they have occurred, have been large; nearly 2 additional years of life, primarily from improvements in just two main factors—smoking and motor vehicle accidents. In addition to reduced mortality, this accounts for quality of life improvements due to reduced smoking. However, these improvements are counterbalanced by declines in health due to other factors, primarily obesity and accidental drug overdose. These are fairly untapped areas in which progress could provide significant benefits, but they have proven more difficult to address. Our study demonstrates the enormous benefit of public health and behavioral change in improving population health, underscoring the importance of continued advances in these areas of research and practice.

Acknowledgments

Kaushik Ghosh and Jean Roth provided some advice and assistance with data analysis. The opinions presented herein are those of the authors and do not necessarily represent the position of the Agency for Healthcare Research and Quality, the National Institutes of Health, or the U.S. Department of Health and Human Services.

Authors' Affiliations

Susan T. Stewart, PhD, is with the National Bureau of Economic Research (NBER) in Cambridge, MA. David M. Cutler, PhD, is with the Department of Economics, Harvard University, and the NBER.

Address correspondence to: Susan T. Stewart, PhD, National Bureau of Economic Research, 1050 Massachusetts Avenue, Cambridge, MA 02138; email sstewart@nber.org.

References

1. Mokdad AH, Marks JS, Stroup DF, et al. Actual causes of death in the United States, 2000. JAMA 2004;291:1238-45. Erratum in: JAMA 2005;293:293-4, 298.
2. McGinnis JM, Foege WH. Actual causes of death in the United States. *JAMA* 1993;270(18):2207–2212.
3. Stewart ST, Rosen AB, Cutler DM. Forecasting the effects of obesity and smoking on U.S. life expectancy. New Engl J Med 2009;361(23):2252-60.
4. Coste J, Quinquis L, D'Almeida S, et al. Smoking and health-related quality of life in the general population. Independent relationships and large differences according to patterns and quantity of smoking and to gender. PLoS One 2014 Mar 17;9(3):e91562.
5. Jia H, Lubetkin EI. The impact of obesity on health-related quality-of-life in the general adult U.S. population. J Public Health (Oxf) 2005;27(2):156-64.
6. Rehm, J, Room R, Graham K, et al. The relationship of average volume of alcohol consumption and patterns of drinking to burden of disease: An overview. Addiction 2003;98:1209-28.
7. The health consequences of smoking—50 years of progress: a report of the Surgeon General. Atlanta, GA: Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; 2014.
8. The health effects of overweight and obesity. Atlanta, GA: Centers for Disease Control and Prevention. Division of Nutrition, Physical Activity, and Obesity, National Center for Chronic Disease Prevention and Health Promotion; 2013. Available at http://www.cdc.gov/healthyweight/effects/. Accessed October 14, 2014.
9. Centers for Disease Control and Prevention. Ten great public health achievements — United States, 1900-1999. MMWR 1999 Apr 2;48(12):241-3.
10. Addressing prescription drug abuse in the United States: current activities and future opportunities. Atlanta, GA: Centers for Disease Control and Prevention; 2013. Available at http://www.cdc.gov/HomeandRecreationalSafety/pdf/HHS_Prescription_Drug_Abuse_Report_09.2013.pdf. Accessed October 14, 2014.
11. Cutler DM. Your money or your life: strong medicine for America's health care system. Oxford: Oxford University Press; 2004.
12. Sleet DA. Achievements in public health, 1900-1999 motor-vehicle safety: a 20th *century public health achievement.* MMWR 1999 May 14;48(18):369-74.
13. Hemenway D. The public health approach to motor vehicles, tobacco, and alcohol, with applications to firearms policy. J Public Health Policy 2001;22(4):381-402.
14. Mozaffarian D, Hemenway D, Ludwig, DS. Curbing gun violence: lessons from public health successes. JAMA 2013;309(6):551-2.
15. Vital statistics of the United States. Volume II, mortality, part A. Website, accessed July 19, 2014 for all years. Available at http://www.cdc.gov/nchs/products/vsus.htm. Accessed October 14, 2014.
16. Klebba AJ, Dolman AB. Comparability of mortality statistics for the seventh and eighth revisions of the International Classification of Diseases. Vital Health Stat 2 1975;(66):1-93. Available at http://www.cdc.gov/nchs/data/series/sr_02/sr02_066.pdf. Accessed October 14, 2014.
17. Estimates of selected comparability ratios based on dual coding of 1976 death certificates by the eighth and ninth revision of the International Classification of Diseases, Monthly Vital Statistics Report (Supplement). Washington, DC: National Center for Health Statistics; 1980 Feb 29;28:1. Available at http://198.246.124.29/nchs/data/mvsr/supp/mv28_11s.pdf. Accessed October 14, 2014.
18. Anderson RN, Minino AM, Hoyert DL, et al. Comparability of cause of death between ICD-9 and ICD-10: preliminary estimates. Natl Vital Stat Rep 2001;49(2):1–32. Available at http://www.cdc.gov/nchs/data/nvsr/nvsr49/nvsr49_02.pdf. Accessed October 14, 2014.
19. Pediatric poisoning fatalities from 1972 through 2010. Bethesda, MD; U.S. Consumer Product Safety Commission; 2013.
20. Schroeder SA. We can do better — improving the health of the American people. New Engl J Med 2007;357:1221-8.
21. Catlin B, Jovaag A, Remington P, et al. County health rankings 2014: Wisconsin. Madison, WI: University of Wisconsin Population Health Institute; 2014.
22. Cutler DM, Rosen AB, Vijan S. The value of medical spending in the United States, 1960–2000. N Engl J Med 2006;355:920–7.
23. Athey S, Stern S. The impact of information technology on emergency health care outcomes. Rand J Econ 2002;33(3):399-432.
24. Wilde ET. Do emergency medical system response times matter for health outcomes? Health Econ 22(7):790–806.
25. Radnofsky L, Walker J. DEA restricts narcotic pain drug prescriptions: Government seeks to curb abuse of hydrocodone combination pills. Wall Street Journal, 2014 Aug 22. Available at http://online.wsj.com/articles/dea-restricts-narcotic-pain-drug-prescriptions-1408647617. Accessed October 14, 2014.

26. Mass. health insurer cuts narcotic prescriptions to combat abuse. Insurance Journal 2014 April 10. Available at http://www.insurancejournal.com/news/east/2014/04/10/325950 htm. Accessed October 14, 2014.
27. Barrett D. U.S. to allow pharmacies to take back unused prescription drugs. Wall Street Journal 2014 Sept 8. Available at http://online.wsj.com/articles/u-s-to-allow-pharmacies-to-take-back-unused-prescription-drugs-1410186602. Accessed October 14, 2014.
28. National Health and Nutrition Examination Survey (NHANES). Centers for Disease Control and Prevention, National Center for Health Statistics. Available at http://www.cdc.gov/nchs/nhanes htm. Accessed July 7, 2014.
29. Obesity: preventing and managing the global epidemic. Report of a WHO consultation on obesity. WHO technical report series: 894; 2000. World Health Organization. Available at http://whqlibdoc.who.int/trs/WHO_TRS_894.pdf?ua=1. Accessed October 15, 2014.
30. National Health Interview Survey (NHIS). Centers for Disease Control and Prevention, National Center for Health Statistics. Available at http://www.cdc.gov/nchs/nhis htm. Accessed October 15, 2014.
31. Preventing excessive alcohol use. Fact sheet; March 14, 2014. Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion. Available at http://www.cdc.gov/alcohol/fact-sheets/prevention.htm. Accessed October 15, 2014/
32. Medical Expenditure Panel Survey (MEPS). Rockville, MD: Agency for Healthcare Research and Quality; 2012. Available at http://www.ahrq.gov/research/data/meps/index html. Accessed October 15, 2014.
33. Stewart ST, Woodward RM, Rosen AB, et al. The impact of symptoms and impairments on overall health in U.S. national health data. Med Care 2008;46(9):954-62.
34. Brooks R, Rabin RE, de Charro F (Eds). The measurement and valuation of health status using EQ-5D: a European perspective. Dordrecht: Kluwer Academic Publishers; 2003
35. Stewart ST, Cutler DM, Rosen AB. U.S. trends in quality-adjusted life expectancy from 1987 to 2008: combining national surveys to more broadly track the health of the nation. Am J Public Health 2013;103(11):e78–e87.
36. United States life tables: 1959-61. Public Health Service Pub. No. 1252. Washington, DC: U.S. Department of Health, Education, and Welfare; 1964. Available at http://www.cdc.gov/nchs/data/lifetables/life59-61_1_1acc.pdf. Accessed October 15, 2014.
37. Vital statistics of the United States, 1973 life tables. Vol. II, section 5. Rockville, MD: U.S. Department of Health, Education, and Welfare; 1975. Available at http://www.cdc.gov/nchs/data/lifetables/life73_2acc.pdf. Accessed October 15, 2014.
38. Health, United States, 2011: with special feature on socioeconomic status and health. HHS Pub. No. 2012-1232. Hyattsville, MD: National Center for Health Statistics; 2012.
39. Miller TR, Bhattacharya S, Zaloshnja E. Fruits of 20 years of highway safety legislative advocacy in the United States. Ann Adv Automot Med 2011;55:357–63.
40. Cutler DM, Glaeser EL, Shapiro JM. Why have Americans become more obese? J Econ Perspect 2003;17(3):93-118.
41. Fryer R, Heaton P, Levitt S, et al. Measuring crack cocaine and its impact. Econ Inquiry 2103;51(3):1651-81.
42. Cohn D, Taylor P, Lopez MH, et al. Gun homicide rate down 49% since 1993 peak; public unaware: pace of decline slows in past decade. Pew Research Center, May 7, 2013.
43. Jones CM, Mack KA, Paulozzi LJ. Pharmaceutical overdose deaths, United States, 2010. JAMA 2013;309:657-9.
44. Bor J, Basu S, Coutts A, et al. Alcohol use during the great recession of 2008-2009. Alcohol Alcohol 2013 May-Jun;48(3):343-8.
45. Caffeine and alcohol. Fact sheet. National Center for Chronic Disease Prevention and Health Promotion; March 14, 2014. Available at http://www.cdc.gov/alcohol/fact-sheets/cab.htm. Accessed October 15, 2014.
46. Freedman DM, Sigurdson AJ, Rajaraman P, et al. The mortality risk of smoking and obesity combined. Am J Prev Med 2006;31(5):355-62.
47. Age-adjusted death rates for 282 selected causes by race and sex using year 2000 standard population: United States, 1979-98. Hyattsville, MD: National Center for Health Statistics; 2000. Available at http://www.cdc.gov/nchs/data/mortab/aadr7998i.pdf. Accessed October 15, 2014.
48. Underlying cause of death 1999-2011 on CDC WONDER online database, released 2014. Data taken from the Multiple Cause of Death Files, 1999-2011, as compiled from data provided by the 57 vital statistics jurisdictions through the Vital Statistics Cooperative Program. Hyattsville, MD: National Center for Health Statistics; 2014. Available at http://wonder.cdc.gov/ucd-icd10 html . Accessed October 15, 2014.
49. Saint Louis C. Surge in narcotic prescriptions for pregnant women. The NY Times: Science, April 13, 2104.
50. Komlos J, Breitfelder A, Sunder M. The transition to post-industrial BMI values among U.S. children. Am J Hum Biol 2009;21(2):151-60.
51. Cutler DM, Glaeser EL, Shapiro JM. Why have Americans become more obese? J Econ Perspect 2003;17(3):93-118.
52. Bleich S, Cutler D, Murray C, et al. Why is the developed world obese? Annu Rev Public Health 2008;29:273-95.

53. Swinburn BA. Obesity prevention: the role of policies, laws and regulations. Aust New Zealand Health Policy 2008;5:12.
54. James WP. The epidemiology of obesity: the size of the problem. J Intern Med 2008;263(4):336-52.
55. Why low-income and food insecure people are vulnerable to overweight and obesity. Washington, DC: Food Research and Action Center; 2010. Available at http://frac.org/initiatives/hunger-and-obesity/. Accessed October 15, 2014.
56. Orzano AJ, Scott JG. Diagnosis and treatment of obesity in adults: an applied evidence-based review. J Am Board Fam Pract 2004;17(5):359-69.
57. Powell LM, Chriqui JF, Khan T, et al. Assessing the potential effectiveness of food and beverage taxes and subsidies for improving public health: a systematic review of prices, demand and body weight outcomes. Obes Rev 2013 Feb;14(2):110-28.
58. Bishai D. Generalized nutrient taxes can increase consumer welfare. Health Econ 2014 Sep 19; epub ahead of print.
59. Kettel Khan L, Sobush K, Keener D, et al. Recommended community strategies and measurements to prevent obesity in the United States. MMWR 2009;58(RR07):1–26.
60. Room R, Babor T, Rehm J. Alcohol and public health. Lancet. 2005;365(9458):519-30.
61. Kelly H. Fighting fires with the help of Google Glass. CNN Tech (Web site); 2014 Feb 23. Available at http://www.cnn.com/2014/01/21/tech/innovation/google-glass-firefighter/. Accessed October 15, 2014.
62. Mercedes-Benz adds QR codes to save car crash victims. BBC News, Technology; 2013 May 27. Available at http://www.bbc.com/news/technology-22682186. Accessed October 15, 2014.
63. Manson JE, Bassuk SS, Hu FB, et al. Estimating the number of deaths due to obesity: can the divergent findings be reconciled? J Women Health 2007;16(2):168-76.
64. van Baal PH, Hoogenveen RT, de Wit GA, et al. Estimating health-adjusted life expectancy conditional on risk factors: results for smoking and obesity. Popul Health Metr 2006;4:14.
65. Saydah S, Bullard KM, Cheng Y, et al. Trends in cardiovascular disease risk factors by obesity level in adults in the United States, NHANES 1999-2010. Obesity 2014;22(8):1888-95.
66. Gregg EW, Li Y, Wang J, et al. Changes in diabetes-related complications in the United States, 1990–2010. N Engl J Med 2014; 370:1514-23.
67. Ogden CL, Carroll MD, Kit BK, et al. Prevalence of childhood and adult obesity in the United States, 2011-2012. JAMA 2014 Feb 26;311(8):806-14.

Susan Stewart, PhD, is a research specialist at the National Bureau of Economic Research, where since 2003, she has been working on health measurement for a Satellite National Health Account. Formerly, she was a postdoctoral research fellow at the University of California San Diego's Health Outcomes Assessment Program. She served as a consultant to the Institute of Medicine's 2005 Committee to Evaluate Measures of Health Benefits for Environmental, Health, and Safety Regulation.

David Cutler, PhD, is the Otto Eckstein Professor of Applied Economics in the Department of Economics at Harvard University. While at Harvard, he has served as an Assistant Professor of Economics and was named John L. Loeb Associate Professor of Social Sciences in 1995. He holds secondary appointments at Harvard's John F. Kennedy School of Government and T. H. Chan School of Public Health. Dr. Cutler has authored numerous publications focused on health care and other public policy topics. Currently, he is a member of the Massachusetts Health Policy Commission and was a key advisor in the formulation of the recent cost control legislation in Massachusetts.

Health Economics and Improvements in Behavioral Health

Richard Frank and Sherry Glied

Abstract

Economic research on behavioral health policy has informed the design of public policies that directly prevent the development of health burdens; contributed to improvements in the efficiency and effectiveness of care delivery; and led to improved outcomes for people with behavioral health problems through an understanding of the relationships between behavioral health and other services. In this chapter, we review the history of economic research on behavioral health and particularly its relationship to Arrow's analysis of market failures in the health care system and to Becker and Grossman's models of human and health capital. The theory of health capital formed the basis of a strand of economic research focused on unhealthy behavior, including an influential set of analyses showing that taxes and regulations on alcohol could reduce crime, accidents, and child abuse. Research building from Arrow's analysis of market failures has contributed to significant improvements in insurance design and provider payment and in enabling a shift in the locus of care from institutions to the community. Finally, the health production model provided a theoretical basis for research analyzing programs that focus on the interaction between mental health and other services, including research on supported housing and supported employment and research on spillovers from behavioral health to general health. The significant impact of economic research on policy stems from the discipline's emphasis on providing empirical estimates of the likely impacts of policy on both people and budgets.

Introduction

Economic research on behavioral health policy has contributed to well-being in three broad ways. First, this research has informed the design of behavioral health-oriented public policies that directly prevent the development of behavioral and physical health burdens. Second, research has informed the design of policies that affect the efficiency and effectiveness of care delivery and thus help more people gain the benefits of effective treatments. Third, economic research has improved outcomes by explicitly recognizing the interactions between behavioral health and other services and institutions that contribute to well-being.

Research has documented how social programs improve behavioral health and augment mental health care. It has also illustrated how behavioral health interventions improve social outcomes, leading to improvements in well-being that depend on access to both social programs and health interventions. In this chapter we review how the field of health economics has contributed to the development of behavioral health policy and how those policies have made a difference in the outcomes, efficiency, and fairness of behavioral health care in the United States. The chapter is

organized into five sections. We begin with a brief overview of the intellectual traditions in health economics that serve as a foundation for applications to the behavioral health area. The subsequent three sections address the connection of research to policymaking in the areas of prevention of behavioral health problems, efficiency and effectiveness of care delivery, and interactions between behavioral health care and social policy. In the final section, we offer some concluding observations, discuss implementation issues, and consider implications for future research.

How Health Economics Has Approached Behavioral Health Issues

Economists have long been interested in problems of health and health policy. Economic research on behavioral health policy is a subset of this more general work. The earliest work in applying economics to health care took one of two approaches. The first, spearheaded by scholars such as Reuben Kessel[1] and Milton Friedman,[2] examined how the institutions of the health sector compared to those that would be expected in a perfectly competitive market for health services. These analyses often pointed to guilds and excessive government involvement as departures from this theoretical competitive market and from (ostensibly) efficient outcomes. A second line of analysis consisted of careful studies of how resources were used in the health sector.[3] This line of research focused on providing rich, statistically-informed descriptions of the demand and supply of health care services and the financing of those services. Several researchers in this tradition examined issues of behavioral health policy. Rashi Fein provided the first systematic assessment of how money was spent on mental health care and a description of how mental health services were financed in the 1950s. These analyses were used to establish a foundation for efforts to shift behavioral health funding from institutional to community care.[4]

Beginning in the 1960s, two new lines of analysis were pioneered in economics that had profound effects for expanding the understanding of the health sector. The first flowed from an analysis of health care markets by Kenneth Arrow.[5] That work, an application of the new economics of information, led to a philosophical shift away from a normative approach that viewed departures from a perfectly competitive market as inefficient policy failures. Instead, Arrow tried to understand the basis for institutions in the health care system that differed from those seen in perfectly competitive markets and to consider whether these impeded or promoted social well-being.

The second line of work stemmed from the research of Gary Becker on the concept of human capital. That area of research sought to understand how "human capital" – the productivity attributes of workers (education, health, experience) – contributed to the overall growth in economic productivity and national output.[6] Human capital research was extended to include the analysis of how households produced and accumulated human capital, including health.

In the 1970s, researchers building on the framework set out by Arrow, began a systematic program of research into the health sector, including behavioral health. They developed theoretical models describing – among many other issues – how insurance affects the use of services (moral hazard); how private information can affect the functioning of insurance markets (adverse selection); how payment incentives affect physician and hospital behavior; how regulation of markets can improve efficiency when information on quality of care is difficult and costly to obtain; and the interplay among financing mechanisms. More recent theoretical research has incorporated new concepts,

such as behavioral economics and economic sociology to the tool set of health economists. The insights of these various theoretical models suggested new directions for policy development.

The research flowing from the human capital tradition considered how individuals make trade-offs among leisure, consumption, and health in allocating their time and resources, a process labeled health production. The health production model was initially developed by Michael Grossman in 1970. His ideas led to a sea change in thinking about risky and unhealthy behaviors.[7] Grossman argued that people made trade-offs among products and services that gave them enjoyment, subject to constraints that arose from their income, time, and knowledge. Good health could be understood as one among these enjoyment-producing products and services – but, products and services that were bad for health, such as tobacco, alcohol, and laziness, might also be among these enjoyment-producers. In this rational utility maximizing framework, raising the price of such "bads" could induce individuals to change the tradeoffs they made, without requiring that they alter their underlying preferences about what did or did not give them enjoyment. Recently, this approach has been extended to take account of addictive behavior, consumer short-sightedness, and other forms of irrational behavior. The applications have focused largely on cigarettes and addictive drugs.[8,9] This line of work is becoming increasingly influential with respect to policy addressing addictive behaviors.

In health economics, the application of theoretical models was almost immediately tied to empirical analyses that could directly inform public policy through estimation of the magnitude of theoretically-posited effects. The best known example of these efforts was the 1979 Health Insurance Experiment, which randomly assigned households to different levels of insurance coverage and measured spending and health outcomes.[10]

Improvements in data collection and in computing capacity further enabled this type of research. As we discuss in detail below, economists, using a variety of experimental and quasi-experimental methods, have assessed the magnitude of many of these effects, and these findings have better equipped private and public policymakers to design institutions and health policies. In some cases, empirical analyses of existing behaviors have been used to forecast the effects of future policy changes. In other cases, evaluations of policy changes that had been implemented have led to improvements and modifications. Empirical estimates have also been incorporated into cost-benefit and cost-effectiveness analyses that directly assess the net impact of a new technology or a policy change.

Health economics research has not produced new pills or procedures. But in similarly tangible ways, the theoretical and empirical findings of health economics research in behavioral health have given consumers, providers, insurers, and local, State, and Federal policymakers insights and a medicine chest of tools and techniques to improve behavioral health outcomes.

Preventing the Development of Behavioral Health Problems

Grossman's theory of health production[7] is among the most influential conceptual approaches to prevention and early intervention for behavioral health problems. It led to a rich body of research on how changes in the price of "bads," induced by changing taxes or penalties, might affect individual

behavior. In the behavioral health arena, much of this research focused on alcohol use. Grossman's model implied that, whatever the psychological and physiological antecedents of excessive alcohol use might be, changes in the price of alcohol would change patterns of utilization. Changes in alcohol prices could be induced by increases in alcohol taxes. Changes in use could be generated through legal restrictions that raised the cost of obtaining alcohol for at-risk populations (teenagers), and they could also be produced through increases in penalties associated with excessive alcohol use (sanctions on driving, criminal penalties).

One strand of this research took advantage of existing variation in the level of alcohol and beer taxes across States and localities within the United States. Quasi-experimental analyses exploiting this variation found that raising alcohol taxes led to reductions in alcohol consumption, as might be expected. More unexpected – but consistent with Grossman's model[7] – were a series of findings indicating that significant reductions in alcohol use had occurred among heavy drinkers, including those at greatest risk of experiencing (or inflicting) substantial harm through alcohol use. For example, zero tolerance laws, adopted by a number of States, reduced heavy episodic drinking by under-age males by 13 percent.[11]

Studies of the impact of raising alcohol taxes consistently show that tax increases contribute to meaningful reductions in motor vehicle fatalities.[12] For example, the consensus estimates suggest that a roughly 6 percent increase in the beer tax reduces highway fatalities by about 2 percent.[13] Raising alcohol taxes also reduces other pathological behaviors associated with heavy drinking. States that raise alcohol taxes see reductions in teenage pregnancies and abortions.[14] Higher alcohol taxes are also associated with reduced arrests for child abuse.[15] A 10 percent increase in alcohol taxes is estimated to reduce violence towards children by women by about 2.1 percent. Higher alcohol taxes even reduce the rates of child homicide.[16] A State with 10 percent higher alcohol taxes can expect 1.9 percent fewer deaths of children by homicide than one with lower taxes.

Similar results were found from studies of other interventions that raise the effective price of alcohol to particular groups. Raising the minimum drinking age, which increases the price at which adolescents and young adults can gain access to alcohol, is associated with reductions in traffic fatalities and in crime.[17,18] In States with higher minimum drinking ages, there were significant reductions in non-violent crimes like vandalism. The implementation of zero-tolerance laws reduced property crime arrests by 3.4 percent and saved tens of billions of dollars over time.[18]

Rational actor models, such as those described above, generate the implication that policies designed to curb the use of "bads" reduce consumer well-being and, by extension, overall economic welfare, unless the policies have benefits for non-consumers (as is the case when alcohol taxes reduce drunk driving, for example). More recent research building on this framework has relaxed the assumption that actors are always fully rational.[20] In this newer literature, policies that lead to reductions in the consumption of "bads" also improve the well-being of the consumers themselves (and economic welfare). In these analyses, policies enable myopic and otherwise rationally constrained consumers to improve their own self-interested behavior.[9] Empirical analyses of the health effects of policies to reduce "bads" support policy aimed at raising or equalizing alcohol taxes.

Improving the Efficiency and Effectiveness of Care Delivery

The analysis of markets for health insurance and health care services derives from the work of Arrow.[5] From an economic standpoint, health care is plagued by problems of uncertain and asymmetric information. People know more about their own health than their insurers or providers do, and providers know more about the care they are providing than consumers do. The problems that Arrow identified in the general health sector, however, are compounded in the arena of behavioral health. Because the extent of behavioral health problems can rarely be measured with a physical diagnostic test and because behavioral health treatments often cannot be readily quantified either, information asymmetries are generally more complex and serious in behavioral health than in general health.

The particular challenges of behavioral health make it more difficult to design optimal financing and delivery incentives in this arena. As a consequence, considerable economic research has focused on addressing problems of moral hazard in the design of behavioral health insurance benefits; adverse selection in competitive behavioral health insurance markets; and agency in the payment of providers and programs that deliver behavioral health services.

Insurance coverage for behavioral health problems has long been far less complete than the coverage for other types of health care.[21] The principal economic forces behind this result were two: moral hazard and adverse selection. Moral hazard is the term economists use for the increased utilization that occurs in the presence of insurance because insured consumers obtain care at a discount and do not face the full cost of services. The result of the insurance discount is a tendency to use "too much" of those services. The amount of the "overuse" depends on how responsive consumer demand for services is to their price – the elasticity of demand. An intensive program of research on the demand for mental health services, including both observational and experimental studies, found that under fee-for-service indemnity insurance arrangements, the elasticity of demand for ambulatory mental health services was about twice that for general medical care.[22,23] This implies that the cost of increasing the generosity of insurance coverage was considerably greater for mental health than for general medical care. This argument provided an efficiency rationale for the unequal mental health coverage that prevailed in the 1970s and 1980s.

Adverse selection also had significant effects on the form of coverage of mental health services in private insurance. Adverse selection occurs when consumers who expect that they are more likely than average to use a particular type of service buy the most comprehensive insurance coverage for these services.[24] This means that insurers with the most generous coverage will attract the least healthy segments of a population. This problem is particularly prevalent in behavioral health because there is strong predictability in mental health use.[25] The presence of adverse selection creates an incentive for insurers to adopt measures that will discourage unhealthy (and costly) enrollees from choosing their plans. They compete to avoid the highest risk clients. In behavioral health this pattern is consistent with the health insurance adage that the last thing one wants to be is the best mental health plan in a market. This competition leads to reduced coverage of behavioral health services, which provides an economic rationale for mandating plans to include behavioral health benefits and for parity statutes that require plans to cover behavioral health benefits at

parity with general health benefits. Thus, health economics research in the 1980s contributed to understanding the efficiency rationales for (a) differential cost sharing for ambulatory mental health care under fee-for-service arrangements and (b) mandated behavioral health benefit statutes.

The fundamental economics of insurance coverage for behavioral health services were dramatically altered by the advent of managed behavioral health care in the 1990s. Managed behavioral health arrangements offered insurers and employers a method of controlling utilization and costs without resorting to reduced coverage. Managed care introduced a set of administrative and technological tools that permitted health plans to control utilization, negotiate lower prices, and apply clinical algorithms to the delivery of behavioral health care. These processes reduced moral hazard in the presence of more generous insurance coverage. Many managed behavioral health care contracts were structured so that a single managed behavioral health care vendor was selected to manage the care for all behavioral health services regardless of which general health plan among several was chosen by an enrollee, which eliminated competition to avoid high-cost cases. This addressed the problem of adverse selection.[22] The results were impressive.[26] Spending growth in behavioral health was sharply reduced, and for some services (inpatient care), absolute spending was also reduced. The ability to control service use without changing cost sharing arrangements under insurance meant that the demand response to coverage was altered, changing the efficiency rationale for differential coverage for behavioral health. This was confirmed by empirical studies of demand.[27,28] Quasi-experimental studies of parity statutes further buttressed this logic, showing that coverage for behavioral health services could be expanded without increasing the level of total spending on behavioral health care over what it otherwise would have been.[29,30] These studies proved to be highly influential in Congressional consideration of legislation that expanded and mandated behavioral health coverage (Mental Health Parity and Addictions Equity Act[a]; mental health provisions of the Affordable Care Act[b]).

The policy impact of this line of research was that over a 10-year period, the Congressional Budget Office reduced its estimates of the expected costs of parity legislation by an order of magnitude, and the Congress became receptive enough to the parity idea that it enacted the Mental Health Parity and Addictions Equity Act in 2008.[c] The Act did not constitute a mandate for coverage, but it did require that coverage for mental health and substance use disorder benefits, if offered, be equal to that for general benefits. The Affordable Care Act and its Essential Health Benefits regulations mandated mental health and substance use disorder coverage as an essential benefit. As the economics literature suggested would happen, the managed behavioral health care carve-out, a separate insurance sub-contract for mental health services, became a dominant approach to organizing insurance for behavioral health services.

Economic research also played a central role in the evolution of payment arrangements for insurers and providers. Economic analyses suggest that the incentives in fee-for service and other cost-based

[a] See Employee Benefits Security Administration, Department of Labor, at http://www.dol.gov/ebsa/mentalhealthparity/.

[b] See U.S. Department of Health and Human Services, Affordable Care Act, About the Law, at http://www.hhs.gov/healthcare/rights/.

[c] See the Mental Health Parity and Addiction Equity Act of 2008 at https://www.federalregister.gov/articles/2013/11/13/2013-27086/final-rules-under-the-paul-wellstone-and-pete-domenici-mental-health-parity-and-addiction-equity-act.

reimbursement arrangements can contribute to rapid cost increases and "overuse' of hospital care. Public policy began to depart from these reimbursement arrangements, shifting towards toward greater use of prospective payment for paying health care providers, including hospital prospective payments (the diagnostic related group, or DRG, system) in the 1980s and physician prospective payment (e.g., capitation) plans in the 1990s. Behavioral health services were largely insulated from this shift toward prospective payment because of several concerns.[31] First, it is more challenging to develop relatively homogeneous diagnostic categories for behavioral health services than it is for other types of health care services. This makes it more difficult to adjust prospective payments for the riskiness of patients within the category, leaving more incentives for providers to avoid people with mental and addictive disorders.[32] Second, empirical evidence based on provider behavior and the behavior of capitated health plans suggested that behavioral health care providers are more responsive to these new incentives than are general medical care providers, so that reductions in service use are much greater, and possibly excessive.[25] This may be due to the greater variability in clinical practice, to less developed quality assurance and measurement systems, or to weaker consumer knowledge. These findings created concern among policymakers that greater caution would be needed in applying prospective payment systems to behavioral health care. The Federal Government exempted specialty hospital and general hospital psychiatric units from Medicare's DRG case-based prospective payment system. Most managed behavioral health care carve-out contracts used a less intensive form of prospective payment – risk corridors – rather than purely capitated payment arrangements in their initial years. Improvements in case mix adjustment methodologies and monitoring practices have led to a greater use of prospective payment mechanisms over time. Capitation is used more frequently today to pay for managed behavioral health care, and Medicare now uses a modified prospective per diem payment system for inpatient psychiatric care.

More recently, these same economic considerations have prompted the use of pay-for-performance systems in general medical care. Pay-for-performance programs reward providers for meeting specific quality targets. As was the case for prospective payment, however, pay-for-performance has been introduced more slowly into the behavioral health arena.[33] The challenges of case-mix adjustments and of measuring relevant outcomes are more difficult in behavioral health than in general health care, and this area requires further research and policy development.

Another area where economic research on provider behavior has been important is in the balance between inpatient and community programs. Policymakers have many reasons to disfavor inpatient hospitalization, including issues of individual rights, concerns over coercive treatment, and the desire to integrate people with severe mental illnesses into their home communities. Publicly owned and operated inpatient psychiatric hospitals and specialty units constitute a significant source of inpatient service supply. These publicly-owned providers have historically been funded directly from State budgets and were frequently decoupled from community programs that were supposed to be providing community-based treatment for people with severe mental disorders.[21] This created perverse incentives: since the community programs could rely on the hospital – an off-budget resource – to care for the sickest patients, community programs were often oriented to serving less severely ill segments of the population. Economic research on this pattern led some States to begin experimenting with new ways of transferring funds to local treatment programs so that there were stronger incentives to treat even the sickest people in community settings. Building on these

insights, policy designers began to write contracts in which they charged community facilities for the costs they imposed on publicly-funded services.[34] For example, in Ohio, the State transferred funds to local programs based on the total expected cost of serving the population and then charged those programs the average per diem cost of publicly provided inpatient care for each day of care used. In addition, bonus payments were awarded to programs based on the number of clients with a severe and persistent mental illness they served. The result was a clear decline in inpatient care use attributable to the policy change. Other States used related approaches to restructure the intergovernmental transfer of mental health funds.[35]

Together, shifting payment policy and the use of managed behavioral health care in insurance-like financing systems, alongside changes in intergovernmental transfer methods, were important contributors to the shift of treatment resources from institutional to community-based behavioral health care that mostly was realized in the period 1985 to 2000. Of course, economic policy was not the only factor leading to these changes. Other important factors included judicial rulings on the rights of patients, new treatments, and the expansion of disability programs.

As noted earlier, new financing approaches coupled with new approaches to treatment of the most severely mentally ill people in community programs allowed for major changes in mental health care delivery. Social sciences and economic research have been influential in understanding the value of the new approaches. Cost-benefit and cost-effectiveness studies of innovative treatment approaches provided evidence about the efficiency of community treatments.[36] That research documented which population segments were most efficient to target with the new programs.[37] It also showed some important limitations to the application of high intensity community-based treatment programs.[38] This research has contributed to widespread adoption of Assertive Community Treatment programs and efforts aimed at targeting these treatments to the populations that generate the greatest cost-effectiveness.

The net result of the confluence of these forces and the policy responses to the expanded understanding has been a mental health system that offers individuals with mental disorders greater autonomy in pursuing treatment, much more complete insurance coverage than at any time in U.S. history, a set of payment arrangements that promote community-based treatment of the most severely ill people, and a set of clinical and financing arrangements that offer more rewards for efficient delivery of care. It is striking to note that there has been a dramatic increase in the share of people with a mental disorder treated since the 1980s (over a 65 percent increase) while the share of GDP devoted to mental health care has hardly changed.[39]

Interactions Between Behavioral Health and Social Policy

Several economic theories suggest that there are connections between behavioral health policy and social policy. As in the case of inpatient facilities and community treatment centers described above, the existence of several programs with different funding streams serving the same populations provides incentives to shift costs. Several analyses have examined the cost-shifting impact of cuts to one type of program on the delivery of services in others. Grossman's health production model,[7] which provided the theoretical basis for empirical examinations of policies to reduce "bads,"

also offers two additional components of a framework for understanding the interactions between behavioral health and social welfare policy. First, the corollary of the effects of "bads" on health is that many goods, including but not limited to medical care, are inputs into the production of health. In general health, the model captures the idea that diet and exercise can substitute for medical care in the production of health. In behavioral health, the range of non-medical factors that can contribute to health is broader still, encompassing housing, employment, and security and stress or insecurity. Picking up on these ideas, health economists have examined the health and well-being benefits, as well as spillover effects, of several aspects of non-clinical care.

Inpatient facilities historically provided some people with serious and persistent mental illness with some treatment. They also provided this population with housing. In the era of deinstitutionalization, it quickly became evident that identifying and subsidizing housing options for people with serious mental illness would be critical to keeping them in the community. Housing was a key input into well-being; the cost of housing might offset the cost of keeping a patient in a psychiatric institution. Following this insight, a range of studies have documented that housing, especially supported housing, both enables people to live in the community and reduces hospitalization and other social costs.[40,41] At least seven controlled studies of permanent supported housing have been conducted, encompassing different populations and a range of settings across the United States, including residents of urban and rural locations, veterans, people with severe mental disorders, severe alcohol abusers, and those with a range of disabilities.[41-49] Only one out of the eight studies showed no significant savings. The other seven estimated savings that ranged from $1,514 per person to as much as $32,302 (converted to 2015 dollars).[46,49] Overall, these studies suggest that clinical and other service savings offset 10 percent to 100 percent of the costs of permanent supportive housing. Savings on the cost of homelessness shelters account for between 18 percent and 45 percent (local funding); Medicaid health services account for between 20 percent and 34 percent of savings; and mental health services (mostly inpatient) account for 20 percent to 40 percent of savings (mix of Federal and State funds), depending on the study. Only one or two studies have measured savings from detoxification (mostly State and local funds) and nursing home use (mix of Federal and State funds).[44,47] However, when they have been measured, these savings tend to be large.

In a similar vein, mental health policy advocates have argued that the experience of work is a vital component to well-being. Paid work may also have spillover effects, defraying the cost of maintaining a person in a hospital or in the community. Analyses of supported employment programs suggest that these programs can generate modest improvements in work.[50,51] In general, however, these increases in paid employment are not sufficient to reduce reliance on other income supports.

The availability of publicly-funded mental health services provides people with implicit insurance coverage for mental health needs. Expansions of insurance coverage can offer an alternative source of financing for this care. Recent economic research on an expansion of Medicaid coverage in Oregon suggests that providing people with symptoms of depression with health insurance has substantial effects on their well-being. Those who gained coverage through the expansion of Medicaid had a much higher rate of being diagnosed with depression (an increase from 4.8 percent to 8.6 percent, $p<0.04$) and were more likely to be receiving antidepressants (16.8 percent to 22.3 percent, $p<0.07$). In addition, the increased security that people experienced by having insurance coverage also appears to have had a direct effect on depression and mental health well-being.[52]

The natural complement of spillover effects from social programs to mental health programs is that mental health may have spillover benefits on other dimensions of well-being. Likewise, a second aspect of Grossman's model[7] is that it formalizes the idea that health is both a good that is desirable in itself (a consumption good) and also one that acts as an investment, enabling people to produce other valuable goods. The perspective of mental health as an investment good lies behind economic analyses of the role of better behavioral health in producing healthier children, more education, and better employment outcomes.

A growing body of evidence suggests that parental depression can harm young children. Postpartum depression has been related to poor parenting, childhood developmental delays, weaker cognitive skill development, attention disorders, and a greatly elevated rate of behavioral problems in children.[53] Research by economists has bolstered psychological studies showing these effects.[54] Economic analyses have strengthened the case that interventions that improve the circumstances of young children can be a very cost-effective strategy for improving adult outcomes.

Another dimension of spillovers from behavioral health to other outcomes relates to the labor market. Several studies have shown that improving depression care can reduce absenteeism and improve the productivity of employees at work. These improvements in productivity are positive spillovers from mental health treatment, and they offset a portion of the increased cost of that treatment.[55,56]

Economic analyses of spillovers from mental health to other aspects of child and adult well-being have contributed to interest in the development and dissemination of evidence-based programs that address behavioral health in these contexts. Recent expansion of Home Visiting,[d] and changes in Head Start standards, note the importance of parental mental well-being on child health.[e]

Economic analyses of the spillover benefits of behavioral health intervention programs have been most explicitly incorporated in decisionmaking in the State of Washington, where the University of Washington Evidence-Based Practice Institute and the Washington State Institute for Public Policy consider the broad fiscal impacts of diverse behavioral health programs and make recommendations to the State legislature based on these analyses.[57] In that context, research documenting the benefits to the criminal justice system of programs to reduce conduct disorder in adolescents has led the State to fund such programs.

Conclusions

Economic analysis has contributed to behavioral health policy in several broad ways. Economists have provided a framework that has broadened and enhanced our thinking about policies to reduce dangerous behaviors, improve the design of coverage and provider payment, and link social programs and behavioral health treatments. Economic theory has pushed empirical research in the direction of studying causal mechanisms. The resulting empirical work has been useful

[d] See Maternal, Infant, and Early Childhood Home Visiting at http://mchb.hrsa.gov/programs/homevisiting/.

[e] See https://www.childwelfare.gov/topics/preventing/programs/homevisit/maternal-mental-health/ and http://eclkc.ohs.acf.hhs.gov/hslc/tta-system/family/docs/depression-pfce-rtp.pdf.

and influential in informing public policy design. Economic evaluations of programs have led to improvements in policy design and implementation. Finally, an economic analysis of the magnitude of potential policy and budgetary effects has often been a key ingredient in the design and enactment of policies related to behavioral health. These policies, in turn, have had striking, large, and tangible effects on the well-being of people with mental illness.

As the discussion above suggests, many of the insights of economic research on behavioral health have been disseminated broadly and incorporated into policy and practice. Two features of economic research have contributed to its policy success. First, the discipline's emphasis on providing empirical estimates of the likely impacts of policy has made economic research valuable to assessments of specific potential policy interventions. Second, economic research often describes outcomes in budgetary and monetary terms (or can be easily adapted to do so), making the research readily comprehensible to policymakers.

Nonetheless, the path from economic research to policy has not always been smooth. Economic research results often challenge large and established interests. For example, while economic analysis has provided a very strong argument for raising alcohol taxes, these taxes remain well below the levels that research suggests would be optimal. Second, economic analysis often challenges conventional wisdom. It took some 15 years until the idea that parity coverage could be accomplished at little to no additional cost was accepted by policymakers. Third, research showing that one or another action improves the efficiency of the provider or insurance markets does not always trump distributional considerations. Changing incentives and funding streams generates both winners and losers, and this can stall even efficient policy changes.

Despite the significant contributions of economic research to behavioral health policy to date, many gaps remain in our understanding of how to design incentives and institutions to best serve people with behavioral health problems. The relatively new field of behavioral economics offers considerable potential for analyses that will improve the design of policies for preventing the development of mental health and substance use problems. The expansion of choice in health insurance markets through the development of the Affordable Care Act Marketplaces and private insurer counterparts, and new forms of prospective payment such as accountable care organizations, make further improvements in case mix and risk adjustment design even more vital. More research on assessing and paying for outcomes in behavioral health ought to accompany the increased attention to the use of incentives to generate improvements in quality through pay-for-performance throughout health care. Finally, much more research is needed to understand how best to address the incentives generated through the overlaps in services provided in the social service and behavioral health systems, including overlaps between behavioral health and disability income, housing, employment, and education.

Acknowledgments

We thank Stephanie Ma for excellent research assistance. The opinions presented herein are those of the authors and may not necessarily represent the position of the Agency for Healthcare Research and Quality, the National Institutes of Health, or the U.S. Department of Health and Human Services.

Authors' Affiliations

Richard Frank, PhD, is Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services. Sherry Glied is Dean and Professor of Public Service, Robert F. Wagner Graduate School of Public Service, New York University.

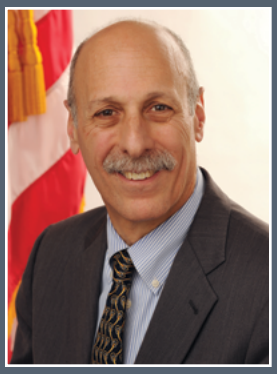
Address correspondence to: Sherry Glied, PhD, 295 Lafayette Street, 2nd Floor, New York, NY 10012; email Sherry.Glied@nyu.edu

References

1. Kessel RA. Price discrimination in medicine. J Law Econ 1958;20-53.
2. Friedman M. Capitalism and freedom. Chicago, IL: University of Chicago Press; 2009.
3. Committee on the Cost of Medical Care. Medical care for the American people: The final report of the Committee on the Costs of Medical Care, Adopted October 31, 1932. Washington, DC: U.S. Government Printing Office; 1932.
4. Fein R. Economics of mental illness. New York, NY: Basic Books; 1958.
5. Arrow KJ. Uncertainty and the welfare economics of medical care. Am Econ Rev 1963;941-73.
6. Becker GS. Investment in human capital: a theoretical analysis. J Polit Econ 1962;9-49.
7. Grossman M. The demand for health: a theoretical and empirical investigation. Cambridge, MA: National Bureau of Economic Research; 1972.
8. Bernheim BD, Rangel A. Addiction and cue-triggered decision processes. Am Econ Rev 2004;94(5):1558-90.
9. Gruber J, Köszegi B. Is addiction "rational"? Theory and evidence. Q J Econ 2001;116(4):1261-1303.
10. Newhouse JP, Rand Corporation Insurance Experiment Group. Free for all?: lessons from the Rand health insurance experiment. Cambridge, MA: Harvard University Press; 1993.
11. Carpenter C. How do zero tolerance drunk driving laws work? J Health Econ 2004;23(1):61-83.
12. Cawley J, Ruhm C. The economics of risky health behaviors. In Barros PP, McGuire T, Pauly M (Eds), Handbook of health economics, Vol. 2. Philadelphia, PA: Elsevier; 2011.
13. Young DJ, Likens TW. Alcohol regulation and auto fatalities. Int Rev Law Econ 2000;20(1):107-26.
14. Sen B. Can beer taxes affect teen pregnancy? evidence based on teen abortion rates and birth rates. Southern Econ J 2003;328-43.
15. Markowitz S, Grossman M. The effects of beer taxes on physical child abuse. J Health Econ 2000;19(2):271-82.
16. Sen B. The relationship between beer taxes, other alcohol policies, and child homicide deaths. Topics Econ Analysis Policy 2006;6(1).
17. Wagenaar AC, Toomey TL. Effects of minimum drinking age laws: review and analyses of the literature from 1960 to 2000." J Stud Alcohol Suppl 2002;14:206-25.
18. Carpenter C. Heavy alcohol use and crime: evidence from underage drunk-driving laws. J Law Econ 2007;50(3):539-57.
19. Kenkel DS, Sindelar JL. Economics of health behaviors and addictions: contemporary issues and policy implications. In Glied S, Smith PC (Eds), Oxford handbook of health economics (pp. 206-31). Oxford, UK: Oxford University Press; 2011.
20. Frank RG, Glied SA. Better but not well: mental health policy in the United States since 1950. Baltimore, MD: Johns Hopkins University Press; 2006.
21. Frank RG, McGuire TG, Newhouse JP. Risk contracts in managed mental health care. Health Aff 1995;14(3):50-64.
22. Frank RG, McGuire TG. A review of studies of the impact of insurance on the demand and utilization of specialty mental health services. Health Serv Res 1986;21(2, Pt 2):241-65.
23. Rothschild M, Stiglitz J. Equilibrium in competitive insurance markets: an essay on the economics of imperfect information. In Dionne G, Harrington SE (Eds), Foundations of insurance economics. Hoboken, NJ: John Wiley & Sons; 1976.
24. Frank RG, McGuire TG. Economics and mental health. In Culyer AJ, Newhouse JP (Eds), Handbook of health economics. Philadelphia, PA: Elsevier; 2000.
25. Sturm R. Tracking changes in behavioral health services: how have carve-outs changed care? J Behav Health Serv Res 1999;26(4):360-71.
26. Meyerhoefer CD, Zuvekas SH. New estimates of the demand for physical and mental health treatment. Health Econ 2010;19(3);297-315.

27. Lu C, Frank RG, McGuire TG. Demand response of mental health services to cost sharing under managed care. J Mental Health Policy Econ 2008;11(3):113-25.
28. Goldman HH, Frank RG, Burnam MA, et al. Behavioral health insurance parity for federal employees. New Engl J Med 2006;354(13):1378-86.
29. McConnell KJ, Samuel HN, Gast M, et al. Behavioral health insurance parity: does Oregon's experience presage the national experience with the Mental Health Parity and Addiction Equity Act?" Am J Psychiatry 2012;169(1):31-8.
30. Ellis RP, McGuire T. Provider behavior under prospective reimbursement: cost sharing and supply. J Health Econ 1986;5(2):129-51.
31. Freiman MP, Ellis RP, McGuire TG. Provider response to Medicare's PPS: reductions in length of stay for psychiatric patients treated in scatter beds. Inquiry 1988;26(2):192-201.
32. Ettner SL, Schoenbaum M, Williams JA. The role of economic incentives in improving the quality of mental health care. In Jones AM (Ed), The Elgar companion to health economics. Northampton, MA: Edward Elgar Publishing; 2012.
33. Frank RG, Gaynor M. incentives, optimality, and publicly provided goods: the case of mental health services. Public Finance Rev 1995;23(2):167-92.
34. Ganju V, Bouchard C. Funding incentives for community based programs: the impact on local services of Texas' $35.50 program. Austin, TX: Texas Department of Mental Health and Mental Retardation; 1990.
35. Weisbrod BA. A guide to benefit-cost analysis, as seen through a controlled experiment in treating the mentally ill. J Health Politics Policy Law 7 1983;4:808-45.
36. Burns BJ, Santos AB. Assertive community treatment: an update of randomized trials. Psychiatr Serv 1995;46(7):669-75.
37. Salkever D, Domino ME, Burns BJ, et al. Assertive community treatment for people with severe mental illness: the effect on hospital use and costs. Health Serv Res 1999;34(2):577-601.
38. Frank RG, Glied SA. Better but not well: mental health policy in the United States since 1950. Baltimore, MD: John Hopkins University Press; 2006.
39. Newman SJ, Reschovsky JD, Kaneda K, et al. The effects of independent living on persons with chronic mental illness: an assessment of the Section 8 certificate program. Milbank Q 1994;72(1): 171-98.
40. The "value added" of linking publicly assisted housing for low-income older adults with enhanced services: a literature syntheses and environmental scan. Washington, DC: U.S. Department of Health and Human Services, Office of The Assistant Secretary for Planning and Evaluation; 2012.
41. Cheng AL, Lin H, Kasprow W, et al. Impact of supported housing on clinical outcomes analysis of a randomized trial using multiple imputation technique." J Nerv Ment Dis 2007;195(1):83-8.
42. Culhane DP, Metraus S, Hadley T. Public service reductions associated with placement of homeless persons with severe mental illness in supportive housing. Hous Policy Debate 2002;13(1):107-63.
43. Kessell ER, Bhatia R, Bamberger JD, et al. Public health care utilization in a cohort of homeless adult applicants to a supportive housing program. J Urban Health 2006;83(5):860-73.
44. Mares AS, Greenberg GA, Rosenheck RA. HUD/HHS/VA collaborative initiative to help end chronic homelessness, national performance outcomes assessment: is system integration associated with client outcomes. West Haven, CT: Northeast Program Evaluation Center; 2007.
45. Martinez T, Burt M. Impact of permanent supportive housing on the use of acute care health services by homeless adults. Psychiatr Serv 2006;57(7):992-9.
46. Mondello M, Bradley J, Chalmers McLaughlin T, et al. Cost of rural homelessness: rural permanent supportive housing cost analysis, State of Maine. Augusta, ME: Maine Department of Health and Human Services; 2009.
47. Sadowski LS, Kee RA, VanderWeele TJ, et al. Effect of a housing and case management program on emergency department visits and hospitalizations among chronically ill homeless adults: a randomized trial. JAMA 2009;301(17):1771-8.
48. Larimer ME, Malone DK, Garner MD, et al. Health care and public service use and costs before and after provision of housing for chronically homeless persons with severe alcohol problems. JAMA 2009; 301(13):1349-57.
49. Bond GR, Drake RE, Becker DR. An update on randomized controlled trials of evidence-based supported employment. Psychiatr Rehabil J 2008;31(4):280-90.
50. Drake RE, Frey W, Bond GR, et al. Assisting Social Security Disability Insurance beneficiaries with schizophrenia, bipolar disorder, or major depression in returning to work. Am J Psychiatry 2013;170(12):1433-41.
51. Baicker K, Taubman SL, Allen HL, et al. The Oregon Experiment—effects of Medicaid on clinical outcomes. N Engl J Med 2013; 368(18):1713-22.
52. Dennis CL, Hodnett E. Psychosocial and psychological interventions for treating postpartum depression. Cochrane Database Syst Rev 2007;4:CD006116.
53. Case A, Fertig A, Paxson C. The lasting impact of childhood health and circumstance. J Health Econ 2005;24(2):365-89.

54. Timbie JW, Horvitz-Lennon M, Frank RG, et al. A meta-analysis of labor supply effects of interventions for major depressive disorder. Psychiatr Serv 2006;57(2):212-8.
55. Wang PS, Simon GE, Avorn J, et al. Telephone screening, outreach, and care management for depressed workers and impact on clinical and work productivity outcomes: a randomized controlled trial. JAMA 2007;298(12):1401-11.
56. Aos S, Mayfield J, Miller, M, et al. Evidence-based treatment of alcohol, drug, and mental health disorders: potential benefits, costs, and fiscal impacts for Washington State. Olympia, WA: Washington State Institute for Public Policy; 2006.



Richard G. Frank, PhD, is the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services, where he advises the Secretary on development of health, disability, human services, data, and science policy and provides advice and analysis on economic policy. He is on leave from his position as the Margaret T. Morris Professor of Health Economics in the Department of Health Care Policy at Harvard Medical School. Previously, Dr. Frank has served as the Deputy Assistant Secretary for Planning and Evaluation, directing the Office on Disability, Aging, and Long-Term Care Policy; research associate with the National Bureau of Economic Research; editor for the Journal of Health Economics; and Professor of Health Economics in the Department of Health Policy, Harvard Medical School. He was elected to the Institute of Medicine in 1997.

Sherry Glied, PhD, is the Dean of New York University's Robert F. Wagner Graduate School of Public Service. Her previous positions include Professor of Health Policy and Management at Columbia University's Mailman School of Public Health; Chair of the Department of Health Policy and Management at Mailman School of Public Health; Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services; and Senior Economist for Health Care and Labor Market Policy on the President's Council of Economic Advisors. Dr. Glied has been elected to the Institute of Medicine and the National Academy of Sciences.

Evidence-Based Psychotherapies: Novel Models of Delivering Treatment

Alan E. Kazdin

Abstract

Remarkable progress has been made in developing psychosocial interventions for a broad range of psychiatric disorders and domains of social, emotional, and behavioral impairment for children, adolescents, and adults. At this time, over 340 interventions have been identified as evidence based in light of rigorously controlled studies, replication of treatment effects, and other criteria. The vast majority of these interventions are delivered in one-to-one, in-person sessions provided by a mental health professional. This model of delivery cannot begin to reach the large numbers of individuals in need of services. Multiple models and novel models of delivery are needed to provide interventions on a large scale in order to reduce the burdens of mental illness. Novel models of treatment delivery are presented that have emerged from global health care, business, economics, and the media and are well outside of mainstream mental health professions. Two models (task-shifting and best-buy interventions) illustrate how different ways of delivering treatment are essential to reach large and diverse groups of unserved individuals. Translational research is presented as a broad framework to convey the scope of the research agenda from basic treatment research to large-scale community applications.

Introduction

The development of psychosocial interventions for the treatment of psychiatric dysfunction and related domains of impairment has made remarkable gains. Currently, there are many psychotherapies or psychosocial treatments that have strong research support and are designated as evidence based. In this chapter, the term psychosocial interventions refers to procedures based on psychological methods that alter functioning by focusing on affect, cognition, and behavior. These are distinguished here from more biologically based interventions (e.g., diet, medication, brain stimulation). The interventions focus on the full range of psychiatric disorders but also on multiple domains that may not focus on a diagnosis (e.g., stress, bereavement) where there is impairment in daily functioning. Psychosocial interventions that are supported by evidence have been designated with many different terms, including empirically supported treatments, empirically validated treatments, evidence-based treatments, evidence-based psychotherapies, and others.[1,2] For this chapter, I will use the term evidence-based psychotherapies to cover the full range of interventions that draw on psychological methods to effect change.

The criteria to define treatment as evidence based vary among multiple disciplines (e.g., psychology, psychiatry, social work), professional organizations within and among countries (e.g., in the

Americas, European Union), and private and public agencies within a given country. Although there is no single consensus definition of evidence-based psychotherapy (EBP), most of these criteria are commonly invoked:

1. Careful specification of the patient population.
2. Random assignment of participants to intervention and comparison or control conditions.
3. Use of treatment manuals that document the procedures.
4. Statistically significant differences at the end of the intervention period between treatment and control or comparison conditions.
5. Replication of outcome effects, ideally by an independent investigator or research team.

There is no single count of the accumulated treatments given the varied definitions. Yet, one U.S. government agency has identified over 340 such interventions for mental health and substance abuse disorders.[3] EBPs encompass interventions for children, adolescents, and adults, although some populations (e.g., ethnic minorities) and developmental stages (e.g., adolescents, elderly) tend to be less well studied. Among the critical issues are extending EBPs to patient care and to do so on a scale sufficient to have impact on the personal and social burdens of mental illness. There are multiple challenges, but they begin with recognition that in the United States, and also worldwide, most people in need of mental health services do not receive them.

The purpose of this chapter is to convey next steps in moving from models of treatment (what techniques are used to alter clinical dysfunction) to models of delivery (how that treatment is provided so that it reaches people in need and on a large scale). The chapter begins by highlighting the dominant model of applying EBPs and how that model is evolving to expand service delivery. The main focus of the chapter is on what is needed to reach large numbers of individuals and novel models of treatment delivery that are available from multiple disciplines. The chapter ends with a discussion of translational research as a framework to help conceptualize the challenges and the next steps in research.

Background: Dominant Model and New Variations of Administering Treatment

EBPs encompass many different therapy techniques and procedures (e.g., mindfulness, cognitive therapy for depression, graduated exposure for anxiety), and these have been developed and evaluated for a period now extending decades. A pervasive consistency in developing EBPs has been the model through which the techniques are delivered. The dominant model has been administration of treatment by a highly trained (e.g., doctoral, Master's level) mental health professional in one-to-one, in-person sessions with a client.[a] The model applies to well-developed EBPs as well as the much larger number of interventions yet to be evaluated empirically. The

[a] Some therapies focus on groups (e.g., 8-10 individuals, families, or couples) and involve more than one individual. Typically, these units are seen individually, in person by a mental health professional. The scope and use of these treatments are not such that they alter the central thesis of the chapter or materially alter the impact of the dominant model of delivering treatment.

one-on-one, in-person model has been enduring, is in demand, and can deliver several EBPs. Indeed, the now vast research on EBPs not only supports scores of specific interventions, but by implication is the dominant model through which they are delivered.

There are several lines of work that have altered facets of the dominant model. First, within the past decade efforts have expanded to take advantage of technology and online versions of treatment that draw on the Internet and other media, including video, texting, application software ("apps"), and feedback of various types via smartphone, smart watches, tablets, and related devices. These variations can extend one-to-one treatment to places (e.g., rural areas, across country borders) where services might not otherwise be available,[4] and many of these address a range of significant clinical problems. For example, "apps" are available for treating anxiety, preventing suicide, and helping individuals recovering from alcohol abuse.[5]

Second and related, other extensions of the model of delivery include self-help interventions and a vast array of techniques (e.g., online interventions, expressive writing techniques) that are available 24/7 and require little or no assistance from a trained professional.[6] Several Internet, computer-based, and self-help psychological interventions are evidence based, achieve outcome effects on par with treatment administered in person by a trained mental health professional, and are high in client adherence and satisfaction.[7-9] Use of technology and self-help interventions are important developments and part of the solution of reaching the community that is otherwise unserved.

Third, recent changes in how EBPs are conceived have implications for delivery of treatment. The development of EBPs has focused on matching treatments to specific psychiatric dysfunctions, such as anxiety disorders, major depression, bipolar disorder, and a few hundred more.[10,11] An alternate view to disorder-specific treatments has emerged and involves two interrelated areas: transdiagnosis and transtreatment, and it serves as a model for understanding clinical dysfunction, underlying processes, and intervention. Transdiagnosis refers to the study of processes that span many different manifestations of clinical dysfunction.[12,13] Several factors have served as the impetus for a transdiagnostic conceptualization of psychopathology, including findings that:

- There are high rates of comorbidity so that individuals (children or adults) who meet criteria for one disorder are likely to meet criteria for at least one other disorder as well.[14,15]
- Underlying processes that maintain "different disorders" often are quite similar.[16,17]
- Several disorders share common biological underpinnings as reflected in brain structures, neurotransmitters, and genes.[18,19]
- A number of EBPs (e.g., cognitive behavior therapy, mindfulness) are effective across a range of disorders, suggesting some common mechanisms or core processes (e.g., working memory, emotion regulation).[20]
- Broad characteristics such as a general psychopathology factor (a "p factor"), neuroticism, perfectionism, and tolerance of uncertainty might serve as underlying or mediating characteristics of many different disorders.[21–24]

Transdiagnostic conceptualizations have altered treatment and treatment research. Rather than adding to the burgeoning list of EBPs for specific disorders, attention is now being directed to the

search for transdiagnostic treatments. Terms such as transdiagnostic treatment, unified treatment protocol, and unified cognitive behavior therapy are increasingly evident in the treatment literature. It is likely that research will continue to explore treatments that are broad in applicability and study basic biological and psychological processes on which these treatments depend.[25–27]

Finally, another effort is designed to facilitate training of clinicians and delivery of EBPs. With hundreds of EBPs now recognized, which ones do we train? In mental health graduate and professional training programs (e.g., psychology, psychiatry, social work), it is rare to provide training in even one or two EBPs. Possible solutions have been advanced and implemented.[28–30] These include identifying a few components of treatments or modules that can be more readily and fluidly used as clinical cases require. A few modules might be widely applicable to the range of disorders and comorbidities that are seen in clinical practice.[31] Also, with a limited number of modules, rather than a long list of EBPs, training clinicians in practice is much more feasible. A smaller set of treatments or modules that can be flexibly deployed raises the prospect of more readily placing EBPs in the hands of clinicians.

These variations of treatment highlighted here, including the use of technology, self-help, transdiagnostic treatments, and modular-type treatments, are critical topics in their own right and can only be mentioned in passing. However, many of them retain core features of the dominant model, i.e., one-to-one treatment by a therapist. Overall, it is still the case that most psychotherapy services, evidence based or not, transdiagnostic or not, and modular or not are administered in the one-to-one model of treatment delivery.

Novel and Expanded Models of Delivery Are Needed

Although the dominant model can deliver EBPs, several findings have emerged to make it increasingly obvious that the model cannot begin to address the challenge of reducing the burdens of mental illness.[32] Among the factors are:

- High rates of psychiatric disorders in community samples, as evident from major epidemiological studies.
- The absence of mental health services for most individuals in need of treatment.
- Paucity of services for populations with some of the greatest needs (e.g., children and elderly individuals, individuals from ethnic minority groups, and those who live in rural areas).
- A limited number of mental health professionals both in developing and developed countries who are available to deliver services.

To make these points more concrete, approximately 25 percent of children, adolescents, and adults living in the community in the United States meet criteria for at least one psychiatric disorder.[33,34]

Currently, that would amount to approximately 80 million people. Apart from prevalence, the personal and social burdens of mental disorders are astounding. Mental disorders are more impairing than common chronic medical disorders.[35] For example, in 2004, the burden of depressive disorders (e.g., years of good health lost because of disability) was ranked third on the list of mental and physical diseases worldwide.[36] By 2030, depression is projected to be the number one cause of disability, ahead of cardiovascular disease, traffic accidents, chronic pulmonary disease, and HIV/AIDS.[37]

Most people in need of services do not receive any treatment or care. In the United States, approximately 70 percent of individuals in need of treatment do not receive services.[38] The significance of this is difficult to overstate. It means that if all clinicians began using EBPs tomorrow, it still would not be pertinent to the majority of individuals who do not receive treatment. As I have noted elsewhere, "treatment as usual" in the United States but also worldwide has another name, namely, "no treatment."[39]

Perhaps an increase in the workforce of trained mental health professionals would address the problem of providing care with the dominant model. Yet the geographical distribution, interests, and composition of the highly trained professional workforce convey why sheer numbers alone will not mitigate the problem of reaching many unserved individuals in need.[40] In the United States, mental health professionals are concentrated in highly populated, affluent urban areas, which limits the ability to reach large numbers of people (e.g., those in rural areas, small towns). In addition, most mental health professionals do not provide care to populations for clinical problems where there is a great need (children, adolescents) and an increasing need (the elderly) need.[41,42] Too few mental health professionals are trained to provide services to these groups. Finally, disproportionately few mental health professionals reflect the cultural and ethnic characteristics of those in need of care. Merely expanding the workforce with additional mental health professionals would not bring treatment to the community on a sufficient scale to have a major impact on the burdens of mental illness. One-to-one, in-person treatment, while useful as a model of delivery, is not very helpful as the dominant or primary model if there is going to be any palpable reduction in the burdens of mental illness. New ways of delivering services are sorely needed.

Requirements to Reach People in Need of Services

We already have proven we can develop interventions, but we have not established that we can extend them very widely in a given country or throughout the world. As a point of departure, it would be valuable to begin by considering what would be the key characteristics or requirements of a model of delivery that could in fact reach people. Table 1 lists several characteristics that, if met, could reach and have impact on large numbers of people.

Table 1. Key characteristics of models of treatment delivery to reach people in need of services

Characteristic	Defined
Reach	Capacity to reach individuals not usually served or well served by the traditional dominant service delivery model
Scalability	Capacity to be applied on a large scale or larger scale than traditional service delivery
Affordability	Relatively low cost compared to the usual model that relies on individual treatment by highly trained (Master's, doctoral degree) professionals
Expansion of the nonprofessional workforce	Increase the number of providers who can deliver interventions
Expansion of settings where interventions are provided	Bring interventions to locales and everyday settings where people in need are likely to participate or already attend
Feasibility and flexibility of intervention delivery	Ensure the interventions can be implemented and adapted to varied local conditions to reach diverse groups in need
Flexibility and choice of alternatives for clients within a particular type or class of effective interventions	Allow choice or alternative ways to meet the criteria for what would be an effective intervention. Exercise and meditation, for example, two very broad classes of intervention that affect mental health and clinical dysfunction. Yet, there are multiple options of precisely what is done to achieve similar outcomes

Multiple models have emerged from global health care, business, economics, and the media—all well outside of traditional psychological and psychiatric care.[32,43] Table 2 presents several models of delivery and their key characteristics. Some of the models make clear the distinction between treatment technique (the procedures or means of altering a clinical problem) from the model of delivery (how that technique is dispensed or provided). Task-shifting is one example where EBPs well established in the dominant model (individual, one-to-one treatment by a mental health professional) are delivered by nonprofessional individuals. This is a case of little or no change in the intervention (e.g., cognitive behavior therapy) but in how or in this case who provides the treatment. In other models, the distinction between models of technique and delivery is blurred because they are connected, and the technique is not at all like one of the traditional or evidence-based treatments. Best buy is an example where advertising and imposition of taxes on products are among viable interventions to address clinical problems, as highlighted below.

Table 2. Illustrations of novel models of delivering health services to expand the reach to the community level

Model	Key Characteristic	Examples	Sample References with Illustrations
Task-Shifting	Expanding the workforce by using lay individuals to administer interventions that otherwise might be delivered by health professionals.	Used worldwide for treatment and prevention of HIV/AIDS. Recently extended to mental health service delivery.	WHO[37] Patel, et al.[44]
Disruptive Innovations	A process in which services or products that are expensive, complicated, and difficult to deliver move in novel ways to alter these characteristics. In health care, services are brought to people more than bringing people to the services.	Delivery of health screening and treatment in shopping malls, drug stores, and grocery stores. Use of smartphones, apps, tablets to assess and deliver mental health interventions.	Christensen et al.[45] Rotheram-Borus, et al.[46]
Interventions in Everyday (unconventional settings)	Expansion of health care beyond clinics and traditional settings to places that people normally attend for other reasons. Overlaps with disruptive innovation but comes from a different tradition and draws on different settings (e.g., schools, workplace, churches, hair salons, barber shops).	Delivery of health screening and education messages in hair salons.	Linnan, et al.[47] Madigan, et al.[48]
Best-buy Interventions	Interventions selected based on their cost-effectiveness, affordability, feasibility for the setting (e.g., country, city), and other criteria. Conceived as an economic tool to help countries select among evidence-based strategies to have impact, where impact is quantified (estimated) for different strategies.	To reduce tobacco use: raising taxes, protecting people from cigarette smoke, warning about the dangers of smoking, and enforcing bans.	Chisholm et al.[49] WHO[50]
Lifestyle Changes	A range of behaviors individuals can engage in that are known to have an impact on physical and/or mental health, including diet control, exercise, meditation, and interaction with nature.	Exercise has broad impact on health and physical health.	Deslandes, et al.[51] Walsh[52]

Table 2. illustrations of novel models of delivering health services to expand the reach to the community level (continued)

Model	Key Characteristic	Examples	Sample References with Illustrations
Use of Social Media	Use of widely available material that includes social networking (e.g., Facebook, Twitter, texting, YouTube, Skype) and brings people together in novel ways to present information, obtain assessment, and provide feedback or delivery of interventions. Interventions can be brought to people wherever they are through these media connections.	Writing regularly as part of blogging to draw on many evidence-based expressive writing interventions; meeting with a therapist or support group in a virtual social world.	Baker & Moore[53] Gorini, et al.[54]
Entertainment Education	Use of television or radio to deliver health care messages and to model health-promoting behaviors. A culturally sensitive long-running series (e.g., TV series) in which different characters take on different roles, deal with the challenges related to the focus of the intervention, and model adaptive strategies.	Early application focused on reducing the birthrate and use of birth control in Mexico.	Singhal & Rogers[55] Singhal, et al.[56]
Use of Technologies	Use of Web-based interventions delivered remotely. Several self-help procedures rely on Web-based treatment. Overlaps with social media but has a separate literature.	Use of Internet-based treatment for cigarette smoking. Web-based self-help treatment for clinical depression.	Harwood and L'Abate[9] Muñoz[57]
Community Partnership Model	Developing partnerships between academics and community members for close collaboration on developing and then implementing action plans for providing community services. The model is a comprehensive process of planning through tracking and evaluating the services.	Development of an implementation strategy to engage agencies to provide services for individuals who are underserved.	Bluthenthal, et al.[58] Wells, et al.[59]

Note. These models occasionally have overlapping characteristics (e.g., bringing interventions to the people in need rather than asking individuals in need to come to special settings) but are worth distinguishing because they come from different traditions, disciplines, and collaborations. Each of the models in the table is elaborated and illustrated in greater depth elsewhere, beyond the specific references listed next to each model.[32,43]

Illustrations

Table 2 provides key characteristics of several delivery models, but in general the models are not widely familiar among mental health researchers and practitioners. I highlight two examples mentioned previously to illustrate their novelty and departure from the dominant model of delivery.

Task-Shifting

Task-shifting is a method to strengthen and expand the health care workforce by redistributing the tasks of delivering services to a broad range of individuals with less training and fewer qualifications than traditional health care workers (e.g., doctors, nurses).[37] This redistribution allows an increase in the total number of health workers (e.g., nonprofessionals, lay individuals) to scale up the scope of providing services. The concept and practice of task-shifting are not new and currently are in place in many developed countries (e.g., Australia, England, United States) where nurses, nurse assistants, and pharmacologists provide services once reserved for doctors. Also, community health workers, a term defined long before task-shifting was developed, have provided specific health services (e.g., birthing, neonatal care, immunization) in developing and developed countries and with demonstrated efficacy.[60,61]

Task-shifting emerged from global health initiatives, particularly in developing countries. These initiatives focused on treating and preventing infectious (e.g., malaria, HIV/AIDS, tuberculosis) and non-communicable disease (e.g., cardiovascular disease, diabetes, cancer, respiratory disease) and improving living conditions and education.[49,62–64] These initiatives provide an important context because they contended with key challenges of meeting health care needs in many cultures, under a variety of conditions (e.g., enormous resource constraints, geographical obstacles), and where people in need of services were not receiving them. Key strategies to address the problems included reorganizing and decentralizing health services to accommodate the limited traditional resources and infrastructure (e.g., medical personnel, hospitals). The majority of task-shifting applications have focused on physical health in developing countries (e.g., Ethiopia, Haiti, Malawi, and Namibia) where shortages of human resources and the burden of illness (e.g., HIV/AIDS) are acute. Empirical evaluations have shown task-shifting to rapidly increase access to services, reach large numbers of individuals in need, yield good health outcomes, and have high levels of patient and counselor satisfaction.[37]

Task-shifting was extended to mental health problems because of (1) its ability to be scaled up to provide services to individuals who otherwise did not have access to care and (2) its adaptability to diverse countries, cultures, and local conditions. An exemplary application of task-shifting in mental health was a randomized controlled trial of treatment of anxiety and depression in India.[43,65] More than 2,700 individuals with depression or anxiety (being served by 24 public and private facilities) received a stepped-care intervention beginning with psychoeducation and then interpersonal psychotherapy, as needed and as administered by lay counselors. The lay counselors had no health background and underwent a structured 2-month training course. Medication was available, as was specialist attention (health professional) for suicidal patients. At 6 and 12 months after treatment, the intervention group had higher rates of recovery than did a treatment-as-usual control group administered by a primary health care worker, as well as lower severity symptom scores, lower disability, fewer planned or attempted suicides, and fewer days of lost work. Overall, the study showed that lay counselors could be trained to administer interventions with fidelity,

and that their interventions reduced the rates of disorder in a large sample. This is an excellent example of extending EBPs developed in controlled research settings to community applications but with a change in the model of delivery of those treatments. An EBP, interpersonal psychotherapy, constituted one of the treatments, but the novelty was in the model of delivery that allowed the intervention to reach many more people than is typically the case in the dominant model where a mental health professional delivers the intervention.

Other studies have demonstrated the impact of task-shifting as a model of delivery for the treatment of depression and schizophrenia.[66,67] These demonstrations not only establish the clinical utility of task-shifting but add to the evidence that lay counselors can deliver effective treatment, and that outcome effects are not sacrificed in the process. Moreover, studies evaluated outcomes on a larger-than-usual scale for psychological intervention studies, evaluated and monitored treatment fidelity, and included followup, among other features.

Best-Buy Interventions

Economics of physical health care have added to the impetus to identify novel models of providing services, and these have been extended to mental health care. A survey of world business leaders by the World Economic Forum indicated that chronic disease (e.g., cardiovascular disease, cancer) is a major threat to economic growth globally.[68,69] Disability and mortality not only exert economic impact on individuals, families, and households, but also on industries and societies through consumption of health care services, loss of income, reduced productivity, and capital expenditures that could otherwise support public and private investment. Best-buy interventions have emerged from this context to designate interventions for physical illnesses, particularly the control of chronic diseases globally.[61]

Best buy refers to an intervention for which, "there is compelling evidence that it is not only highly cost effective, but also feasible, low cost (affordable), and appropriate to implement within the constraints of a local health system."[49] Best buy also considers features such as appropriateness for the setting (e.g., culture, resources), capacity of the health system to deliver a given intervention to the targeted population, technical complexity of the intervention (e.g., level of training that might be required), and acceptability based on cultural, religious, and social norms.

Identifying best buy interventions was conceived as an economic tool to help countries assess how to achieve a given amount of change, given the number of eligible individuals in need of the intervention, the potential savings of those changes, and the cost differences of alternative strategies, among other variables.[48,68,70] For example, in one analysis, four criteria (health impact, cost-effectiveness, cost of implementation, and feasibility of scaling up) were used to identify best-buy interventions that would have significant public health impact on noncommunicable diseases, including cardiovascular disease, cancers, diabetes, and chronic lung disease.[49] Best buys for cardiovascular disease and diabetes were counseling, multi-drug therapy, and aspirin. These were selected in light of the reduction of disease burden and very low cost. The example is not necessarily one that applies to all locales. The best-buy interventions can vary for a given disorder and country because of the cost of delivering a particular intervention in light of varied health care resources and infrastructure.

Best-buy interventions for physical diseases often focus on domains of functioning that overlap with and are part of behavior, lifestyle, and mental health as reflected in substance use and abuse (e.g., alcohol and tobacco). For example, for alcohol use, best-buy interventions include enhanced taxation of alcoholic beverages and comprehensive bans on advertising and marketing, based on their favorable cost-effectiveness, affordability overall, and feasibility. Excessive alcohol use was identified as a best buy for reducing the incidence of cardiovascular diseases and cancers, but it also extends to other burdensome conditions (e.g., cirrhosis of the liver, depression, traffic injuries and deaths).[49] More explicit designations of best buys have been identified for select mental disorders. For example, for clinical depression, generically produced antidepressant medication, brief psychotherapy, and treating depression in primary care qualified as best buys.[48] For psychoses, treating people with antipsychotic drugs and with psychosocial support are regarded as best buys.

Best-buy interventions are based on estimates of utilization and impact, relying on mathematical models.[48] Direct tests are critical to ensure that well-intended, feasible, and scalable interventions yield the intended outcomes and in fact are best buys. Also, as in any large-scale intervention, sustaining the integrity of the intervention can be a challenge. Yet, some best-buy interventions (e.g., selective taxes, bans on advertising to reduce substance use and abuse) differ from the usual psychological interventions and do not require compliance by clients or adherence to a specific treatment protocol by therapists. Of course, taxes and advertising have their own problems (e.g., black market sales that are not taxed, advertising not reaching the target population), but these different problems are precisely the reasons why we need multiple models of delivering interventions. No single model or small set of models is likely to reach the vast majority in need of services.

General Comments

My illustrations do not do full justice to the interventions and models of delivery. Yet the purpose in noting these was to convey that "novel models of delivery" is not an abstraction or a class with no members. Rather, there is already evidence for some of the models (see Table 2) that in fact they can be applied, achieve the desired outcomes, and be scaled to meet the enormous and still unmet need for services.

The importance of addressing mental illness has been well recognized in its own right. Yet, accelerated attention emerged from global health initiatives to treat physical disease (e.g., chronic, infectious).[61,68] Initiatives to provide physical health care services revealed gaps in mental health services. Moreover, it has become increasingly clear that mental and physical health are inextricably intertwined, with bidirectional, reciprocal, and comorbid relations. Reducing the burdens of physical health cannot neglect mental health, as reflected in the oft-cited statement there is "no health without mental health."[71,72]

Several models or strategies for delivering treatment and preventive interventions emerged to address physical diseases. Many barriers for delivering care for physical health care to large numbers of individuals in need, particularly in developing countries, were recognized to be similar to the barriers encountered in providing mental health care.[73,74] Consequently, models for delivering treatment proved to be applicable to both mental and physical health services. What remains to be accomplished is moving these models of delivery to mainstream applications of services.

Stepped Care: How to Deploy Treatments and Models

Determining how the different models and treatments within the range of available models will be deployed is related to expanding the models into mainstream care. That is, treatments and models of delivery vary along several dimensions, and some guidance would be needed for their application. Stepped care provides a useful point of departure for considering both treatments and delivery models.

Stepped care is the notion that we should begin with less intensive and less costly interventions that are more easily disseminated, and then move to more aggressive, costly, and specialized treatment as needed. The concept of stepped care has been around for some time and continues to be advocated as a model for providing treatment.[75–79] Also, earlier in this chapter I mentioned one example in which stepped care was used.[43,64]

Currently, cognitive therapy and interpersonal psychotherapy are two EBPs for individuals seeking treatment for depression. Typically, these are individually administered by a professional in a clinic setting. Yet, stepped care raises the question, "Are there opportunities to intervene to reduce and treat depression effectively without moving to these individually based therapies at least as an initial point of departure?" Stepped care would consider a range of interventions that could be applied as needed if a less costly intervention did not achieve change.

Where can we begin with low-cost interventions that might reach large numbers of individuals and be effective? Two come to mind, merely to make the point and again in the context of depression. First, a psychoeducational intervention might be systematically tried. This would consist of telling people about depression and what they can do about it on their own, including activities in which they could engage. We know already that information and education in relation to the treatment of disorders (or as behavior-change techniques more generally) typically are weak in the magnitude of change and number of individuals affected as interventions go. Yet, in a stepped-care model that is not necessarily problematic at all. It is likely that some number of individuals in need of treatment would be helped sufficiently with systematic efforts to provide psychoeducation. Psychoeducation can be delivered widely through many media (e.g., smartphone, TV, brochures, the Web), provided where people are in everyday life (e.g., grocery stores, shopping malls, doctors' offices), adapted culturally (e.g., language, customs, use of photos or illustrations of population-specific or –neutral individuals on the medium); made available 24/7, and if necessary, disseminated by lay individuals rather than professionals.

As another possibility, physical exercise might be promoted as a first or early line of attack to intervene for depression (and for other domains of mental disability as well). There is already evidence about its effectiveness clinically in nonhuman models to suggest the mechanisms through which it operates.[50,80,81] Also, exercise can take many different forms to suit individual tastes and preferences, and it can vary across the developmental spectrum (e.g., childhood through old age). Incentives to exercise could be integrated in schools, the workplace, or at public events (e.g., warm-up with the players before athletic events to receive a monetary rebate on the ticket) as part of combined preventive and treatment strategies. Again, exercise would not be expected to eliminate clinical depression for all or even most individuals. Yet the benefits would be expected to reduce the

incidence and prevalence of depression, leaving aside the enormous individual and societal benefits related to physical health. Will psychoeducation or exercise "cure" or "eliminate" depression? That is not the appropriate stepped-care question. Rather, can these interventions, and others like them (less costly, more easily disseminated) be deployed and be effective with a subgroup of individuals in need as an initial line of attack, and can they reduce the need for more intensive and costly interventions?

Stepped care develops as a sequence of interventions that vary in effort and cost. A given client might not necessarily go through the sequence in order. So, we do not take a depressed, suicidal patient and say, "Let us see what a couple of brochures could do for you." Moderator research might well be helpful in deciding where in a stepped-care sequence of treatment would be an optimal place to begin with this client. What is missing now is stepped-care thinking in routinely providing services and developing EBPs to serve that overall model. It is true currently that most people receive the least costly intervention, i.e., nothing. (Actually, with emergency room visits and physical health correlates and consequences of mental illness, "no treatment" is probably very expensive.) Much could be done to treat and reduce the burden of mental illness by moving to expanded, more easily disseminated, and stepped-care models of providing interventions.

Stepped care involves dual considerations of not only the treatment technique, but also the model of delivery. And, as we learned from best-buy interventions, treatment options might need to vary with local conditions (e.g., resources, health delivery infrastructure). Yet the initial goal is to move beyond the current dominant EBP treatment model. Considering options among the treatments and among the models of delivery and adapting these to local conditions would mean that adopting stepped care is not straightforward. Understandably, in deciding which interventions to apply, the best-buy model has relied on mathematical modeling to help with many of these complexities for making initial treatment recommendations.[48]

General Comments

Novel models, as illustrated by task-shifting and best-buy interventions, address many of the characteristics noted previously, such as reach, scalability, and cost. They add to the dominant model and increase the likelihood of reaching more people who are not being served but are in need of mental health care. Those who are unserved within a country or among different countries are heterogeneous in culture, ethnicity, geography, resources, infrastructure for providing and receiving care, and many other characteristics that can influence treatment delivery. Any one model will miss key segments of the population in need of services. But multiple models, particularly those that begin with the characteristics that are needed to provide treatment on a large scale, are likely to have the needed impact.

The cultural sensitivity issue warrants further comment. Many of the EBPs have been developed, evaluated, and implemented largely in Western cultures and could readily vary in applicability and effectiveness among diverse cultures. It is true that many EBPs do not vary in effectiveness across the few ethnic cultural groups (out of thousands internationally) to which they have been extended.[82] Add to that a small number of EBPs that began with the cultural and ethnic groups of interest as a basis for developing treatment.[83]

The novel models I have mentioned begin with a different and complementary point of departure for developing ethnically and culturally sensitive interventions. They begin with a global perspective and, as part of that, are designed to accommodate local conditions including what is feasible, not just economically, but what is acceptable to those who would be the recipients of the intervention. In task-shifting, for example, lay members of the communities in which treatment is provided are directly involved in delivery of the care. Thus, one is delivering and receiving interventions among one's peers of the same culture, ethnicity, and traditions. In best-buy interventions, precisely which interventions are likely to be appropriate is determined by local conditions and resources (e.g., government, political, likely impact) and, in that sense, are also compatible with the culture and society. A seemingly great best-buy intervention (e.g., taxes, advertising, medication) might not fit at all for a given country and culture, not just for feasibility or relevance but because these may not be an effective way to exert influence in that culture. The unique contribution of the multiple-models approach is that it begins with the goal, namely, reaching people in need to reduce the burdens of mental illness.

Translational Research: Conceptualizing the Research Terrain

An obstacle in moving to novel models of treatment delivery may be the absence of a broader framework that places current work into a larger context and also points to needed directions. Current research evaluates and develops EBPs in the model of delivery I mentioned and in well-controlled treatment trials. The needed work I have outlined is to develop models of delivery that can provide treatment on a larger scale. The now familiar concept of translational research includes concepts that provide a useful framework. Moreover, translational research spans areas of science and technology (e.g., medicine, agriculture, engineering) where there is interest in moving findings from the lab to application.[b]

Definitions of translational research vary in the kinds of research that qualify and their emphases.[84,85] Critical concepts from translational research that help cast the challenges for EBPs are bench, bedside, and community and the movement from one to the other. The phrase "bench to bedside" refers to the translation of research under well-controlled laboratory situations (where "bench" is equivalent to "laboratory" or "basic" research) to patient care (where "bedside" is direct application).

Although "bench to beside" is the key phrase that characterizes translational research, community conveys a broader thrust and a special challenge for EBPs. Translational research includes "bedside

[b] With translational research, a major concern is that findings from basic research take a long time, often decades, to move from the lab to helping people in everyday life. Translational research emerged in an effort to move findings from the lab to clinics more systematically and quickly. This is slightly different from the move to novel models of delivery, but key concepts from translational research are useful in casting the challenge for moving forward. Another area of work closely related to translational research is called Implementation Science (www.fic nih.gov/News/Events/implementation-science/Pages/faqs.aspx). Implementation science focuses on the movement from EBPs to application (e.g., how to do that, what implementation strategies are effective, how to adapt findings from controlled research to "real" world settings). Sometimes this is characterized as "research to programs" and "research to policy." Although the topic is beyond the goals of this chapter, the delineation of implementation science conveys the attention and concern in moving findings from research to application.

to community," which refers to bringing the findings and applications to others on a larger scale. This means taking bedside findings, i.e., research that can help individual patients or groups of patients in relatively small studies to the level of the community. Community here refers to interventions that can be scaled up perhaps at the level of public health. Vaccinations may be among the most familiar examples to convey the full range from bench and bedside to community in which very basic studies are done (e.g., nonhuman animal studies, evaluations of underlying processes), next these are moved to small scale or isolated applications to monitor their effects with individuals or small groups, and eventually they move to community-wide applications.

Bench to Bedside

EBP research by and large has been exemplary at the "bench" level, as it comprises well-controlled trials in laboratory-like rather than clinical practice settings. For some of the treatments, the bench part includes studies using animal models (e.g., to evaluate extinction of anxiety or reduction of depression),[79,86] and that qualifies as the usual meaning of "bench" in discussions of translational research. Yet, let us begin at the level of randomized controlled trials that have been the core focus of developing EBPs. What is accepted as routine and indeed exemplary treatment research involves careful screening of the sample using inclusion and exclusion criteria to recruit clients, development of manuals that specify the treatment, extensive training and supervision of therapists to administer treatment, and so on. Although this is not animal laboratory research, it has a "bench" feature because of the highly controlled, small-scale application. Patients can get better in such trials of course, so the research extends beyond a "proof of concept" demonstration. Yet, the high levels of experimental control when added to the dominant model make the treatment not very applicable beyond the confines of the study. This is the bench part.

Currently a major research priority is to extend treatment from the highly controlled conditions of the lab to "bedside" (patient care). The challenges of extending EBP findings from bench to clinical practice while retaining the effectiveness of treatment are enormous. Current efforts to disseminate EBPs is exactly that step, namely get the treatments in the hands of clinicians so that patient care benefits from the research.

Clinicians often integrate EBPs into their practice based on continuing education or workshop experiences. Typically, these training experiences are not up to the challenge of imparting the skill sets that many EBPs require. Understandably, without this training or supervision, one cannot expect EBPs to be optimally effective when clinicians try to adopt them. A now well-established finding is that when EBPs are used in clinical practice, their effectiveness drops sharply from what the results show in controlled treatment studies.[30,87] The standard explanation for why treatment effects are diminished has been that in clinical practice patients are more severely impaired and diverse than the "pure" cases seen in controlled trials. There are enough exceptions with direct applications of EBPs to patients with multiple comorbidities to challenge this interpretation.[88–90] Even so, it remains the case that many if not most studies with EBPs do not include clinically referred samples, or they are conducted in clinical practice settings. There may be many reasons why treatment outcome effects drop off when bench-to-bedside extensions are made. Among the likely candidates is the lack of training of the practitioners, dilution of the treatment (e.g., fewer sessions, combined with other treatments), and overall fidelity of the treatment in practice, compared to highly controlled research settings.

Bedside to Community

Bench-to-bedside remains important and is the incubator of interventions that may include principles and practices that serve as the bases for larger-scale interventions. Yet, current EBPs are not likely to have much impact as they are developed, studied, and disseminated at the present time. As I have mentioned, the problem with current EBPs is the dominant model of treatment delivery. Now it is critical to attend to the larger community of individuals in need of services and what we can do to deliver available treatments or draw on new ones.

The move from bedside to community does not merely require scaling up an intervention in the usual way. Scaling introduces special features that change the very nature of the task. In treatment, very well trained and supervised therapists are usually part of a clinical trial (bench), and training and supervision are two of the components that fall down in extensions to clinical practice (bedside). Scaling up now involves many more individuals administering treatment, under the most diverse circumstances. This is not a matter of doing more of the same but changing the model of delivery. Problems and challenges to administering treatment effectively are new, different, and formidable when providing an intervention on scale, even when the treatment is really well specified, clear, and not so difficult to administer (e.g., polio vaccinations). It was for this reason that my discussion began by considering the requirements of what is needed to provide a treatment that is to be administered on a large scale.

Translational research emphasizes bench, bedside, and community, but the progression need not be unidirectional and move from bench, to bedside, to community. There would be enormous value to beginning in the community with interventions that can be administered on scale and that seem to be working. These interventions can also be moved to the bench to evaluate critical features (e.g., mechanisms, mediators).

General Comments

Key concepts of bench, bedside, and community help convey the different levels of interest and our foci. Each type of work is ctitical to reduce the burden of mental illness. We want laboratory, experimental, and controlled studies (bench), and we want tests of how and whether a treatment is effective when extended to more routine practice settings (clinical work and patient care). These facets (bench, bedside) are being studied.

What is missing is more concern about extending treatment on scale. It is important to conceptualize critical goals of treatment with the community as an end point of our efforts. We develop treatments to have impact on the burden of mental illness. This goes beyond identifying EBPs alone but also ensuring that at some point, we are extending these in such a way that they will have impact on the scale as does the impact of mental illness.

It is useful in this conceptualization to begin with identifying the demands of models of delivery that can meet community needs. This does not begin at the bench level but rather looks at community needs, resources, and options. The ability of the model to scale up treatment, bring treatment to those in need, expand the workforce, and address other dimensions mentioned previously (e.g., scalability, affordability) are not currently considered in the context of bench

and bench-to-bedside research of psychosocial treatments. Not only do we need different models of delivery, we also need a different mindset in our research efforts. The current mindset is to demonstrate the effectiveness of a one-to-one delivered treatment and see if it can be generalized (extended to clinic settings). That part already has its own obstacles (e.g., training clinicians in large numbers, ensuring the integrity of treatment, preventing the decline in treatment effects) that remain to be resolved. The central reason to raise translational research has been to turn attention to the need to scale up interventions to the community. Even if bench to bedside were to be successful, there is little inherent in that process to help the community, i.e., the large and diverse population in need of care, in diverse contexts, settings, and cultures and for whom individual one-to-one, in person (or Web-based) treatment is not likely to be an option.

Conclusions

EBPs represent an enormous research advance. The comments in this chapter are not a challenge to that at all. Indeed, we are now at the first point in history where behavioral and social sciences have established a large set of treatments (a few hundred and growing) with rigorous scientific evidence supporting those treatments. This accomplishment has to be savored as an evolutionary leap that allows us to consider what is needed for the next breakthrough.

The vast majority of EBPs rely on a model of providing mental health services that is one-to-one, in-person treatment delivered by a mental health professional. This model has proven itself as a platform for effective treatments. What is clear now is that multiple models of treatment delivery are needed to ensure that the large and diverse numbers of individuals needing care can be reached.

In this chapter, I have discussed novel models of delivering treatment that are not merely "potential" options that could be used; in actuality, they are being used now in different contexts but not very often to deliver mental health services. The models draw on advances from multiple disciplines beginning with those in delivering physical health care but also drawing on health care, business, economics, and the media—all well outside of traditional psychological and psychiatric care. I listed several models and illustrated two (task-shifting, best buy), both of which have emerged in the context of providing physical health care but have also entered into mental health care.

The diverse models begin with the requirements of scaling up and sensitivity to local conditions (e.g., resources, geography, culture) that may influence care delivery. It is not just one model of delivery that is needed; indeed, elsewhere I have argued for a portfolio of delivery models to ensure coverage and multiple opportunities to reach those in need.[43] Thus, no single model among those I have noted (Table 2) is intended to be the new "dominant" model. Also, there is no need to replace or eliminate the dominant model of one-to-one, in-person therapy administered by a mental health professional. That model is quite fine, but it is limited, in that it reaches very few of those in need of services.

In addition to different models, I conveyed the importance of stepped care, i.e., providing both the intervention and model of delivery in such a way that those interventions involving less effort and cost would be administered first when possible, and then the next stage or step of treatment

would be administered if that did not work or was not likely to work. Introducing treatments in a stepped-care fashion raises its own set of research questions. However, stepped care is already being used and studied selectively (e.g., psychoeducation, then an evidence-based treatment) in some studies.[44,75] As with novel models, the task is making stepped care more fully integrated in contemporary treatment research and then application.

In conclusion, I discussed the use of translational research to provide a framework that could help to integrate the type of work we as researchers have been doing to develop and establish treatments as evidence based. The development of EBPs in well-controlled settings reflects enormous success for the "bench" part of our research. Now we are innovating and struggling to get the treatments to clinicians and patient care, the "bedside" part of our work. These are efforts to breakdown and simplify treatment (e.g., modules) or develop individual treatments that can be applied widely (transtreatment) so that it is more feasible for training individuals currently in or being trained for private practice. Yet in all of this we have pretty much neglected the "community" part of the work. In this context, extension to the community refers to the large numbers of individuals in the population who are in need of services but receive no intervention.

Current "bench" based EBPs rely on a model of individual one-on-one therapy administered by a mental health professional. This is a viable model to be sure, but it cannot reach the many people in need. Multiple other models are required that begin with the need to reach the community. Actually, several models are available (see Table 2), but they are not mainstream within the mental health professions (psychiatry, psychology, social work). The challenges of reducing the burdens of mental illness within a country and worldwide will require attending to novel ways of scaling up treatment and collaborating with other disciplines, as well as with governments, third-party payers, and policymakers.

There is a huge need worldwide for interventions that reduce the personal and social burdens of mental illness. With the success of EBPs in highly controlled settings (bench), it important to focus on models of delivering treatment that can have impact beyond the small proportion of individuals who have access to care (bedside). Which of our interventions can be used to make a difference on a large scale (community)? The advances in EBPs have made this gap in our knowledge more salient, and perhaps that will make it more likely to be addressed in the coming years.

Acknowledgments

The author is not aware of any affiliations, funding, or financial holdings that would influence comments or positions taken in this chapter. Preparation of this chapter was not supported by any funding agency. The opinions presented in this chapter are those of the author and should not be construed as the official position of the Agency for Healthcare Research and Quality, the National Institutes of Health, or the U.S. Department of Health and Human Services.

Author's Affiliation

Alan E. Kazdin, PhD, is Sterling Professor of Psychology, Professor of Child Psychiatry, and Director of the Yale Parenting Center, New Haven, CT.

Address correspondence to: Alan E. Kazdin, Department of Psychology, 2 Hillhouse Avenue, Yale University, New Haven, CT 06520-8205; email: alan.kazdin@yale.edu

References

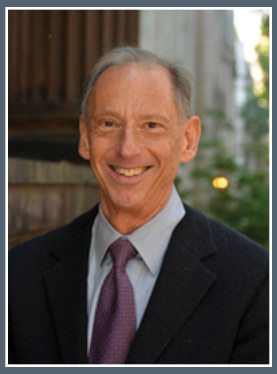
1. American Psychological Association Presidential Task Force on Evidence-based Practice. Evidence-based practice in psychology. Am Psychol 2006; 61:271-85.
2. Chambless D L, Ollendick TH. (2001). Empirically supported psychological interventions: Controversies and evidence. Ann Rev Psychol 52:685–716.
3. National Registry of Evidence-Based Programs and Practices (NREPP). Rockville, MD: Substance Abuse and Mental Health Services Administration; 2014. Available at http://nrepp.samhsa.gov/. Accessed July 23, 2014.
4. Wooton R. (Ed.). Telepsychiatry and e-mental health. London: Royal Society of Medicine Press; 2003.
5. Singh M. Therapists' apps aim to help with mental health issues. Shots – Health News. National Public Radio, March 26, 2014. Available at http://www npr.org/blogs/health/2014/03/26/294374936/therapists-apps-aim-to-help-with-mental-health-issues. Accessed July 23, 2014.
6. L'Abate L. (Ed.). Low-cost approaches to promote physical and mental health: Theory, research, and practice. New York: Springer Science+Business Media; 2007.
7. Andrews G, Cuijpers P, Craske MG, et al. (2010). Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: A meta-analysis. PLoS One 2010; 5(10):e13196.
8. Cartreine JA, Ahern DK, Locke SE. A roadmap to computer-based psychotherapy in the United States. Harv Rev Psychiatry 2010; 18:80-95.
9. Harwood TM, L'Abate L. Self-help in mental health: A critical review. New York: Springer; 2010.
10. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. (5th ed.). Arlington, VA: American Psychiatric Publishing; 2013.
11. World Health Organization (WHO). International classification of diseases-10 (4th ed). Geneva: WHO; 2010.
12. Harvey AG, Watkins E, Mansell W, et al. Cognitive behavioural processes across psychological disorders: A transdiagnostic approach to research and treatment. New York: Oxford University Press; 2004.
13. Kring AM, Sloan, DM. (Eds.). Emotion regulation and psychopathology: A transdiagnostic approach to etiology and treatment. New York: Guilford Press; 2009.
14. Kessler RC, McGonagle KA, Zhao S, et al. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States: Results from the National Comorbidity Survey. Arch Gen Psychiatry 1994; 51:8-9.
15. Wichstrøm L, Berg-Nielsen TS, Angold A, et al. (2012). Prevalence of psychiatric disorders in preschoolers. J Child Psychol Psychiatry 2012; 53:695-705.
16. Fairburn CG, Cooper Z, Shafran R. Cognitive behaviour therapy for eating disorders: A "transdiagnostic" theory and treatment. Behav Res Ther 2003; 41(5):509-28.
17. Green JG, McLaughlin KA, Berglund PA, et al. Childhood adversities and adult psychiatric disorders in the national comorbidity survey replication. I: Associations with first onset of DSM-IV disorders. Arch Gen Psychiatry 2010; 67:113-23.
18. Meyer-Lindenberg A, Tost H. Neural mechanisms of social risk for psychiatric disorders. Nature Neurosci 2012; 15:663-668.
19. Smoller JW, Craddock N, Kendler K, et al. Identification of risk loci with shared effects on five major psychiatric disorders: A genome-wide analysis. Lancet 2013; 381:1371-9.
20. Johnston L, Titov N, Andrews G, et al. Comorbidity and internet-delivered transdiagnostic cognitive behavioural therapy for anxiety disorders. Cogn Behav Ther 2013; 42(3):180-92.
21. Barlow DH, Sauer-Zavala S, Carl JR, et al. The nature, diagnosis, and treatment of neuroticism: Back to the future. Clin Psychol Sci (forthcoming).

22. Caspi A, Houts RM, Belsky DW, et al. The 'p factor': One general psychopathology factor in the structure of psychiatric disorders? Clin Psychol Sci 2014; 2:119-37.
23. Egan SJ, Wade TD, Shafran R. Perfectionism as a transdiagnostic process: A clinical review. Clin Psychol Rev 2011; 31:203-12.
24. McEvoy PM, Mahoney AEJ. To be sure, to be sure: Intolerance of uncertainty mediates symptoms of various anxiety disorders and depression. Behav Ther 2012; 43:533–45.
25. Harvey AG, Lee J, Williams J, et al. Improving outcome of psychosocial treatments by enhancing memory and learning. Perspect Psychol Sci 2014; 9(2):161-79.
26. McEvoy PM, Nathan P, Norton PJ. Efficacy of transdiagnostic treatments: A review of published outcome studies and future research directions. J Cogn Psychother 2009; 23(1):20-33.
27. Reinholt N, Krogh J. Efficacy of transdiagnostic cognitive behaviour therapy for anxiety disorders: A systematic review and meta-analysis of published outcome studies. Cogn Behav Ther (in press).
28. Chorpita BF, Daleiden EL. Building evidence-based systems in children's mental health. Treatment for children and adolescents. In: Weisz JR, Kazdin AE (Eds), Evidence-based psychotherapies for children and adolescents (2nd ed, pp. 482-99). New York: Guilford Press; 2010.
29. Shafran R, Craske M. (Eds). Special issue: Dissemination and implementation of cognitive behavioral therapy. Behav Res Ther 2009; 47:901-99.
30. Weisz JR, Ng MY, Bearman SK. Odd couple? Re-envisioning the relation between science and practice in the dissemination and implementation era. Clin Psychol Sci 2014; 2:58-74.
31. Weisz JR, Chorpita BF, Palinkas LA, et al. Testing standard and modular designs for psychotherapy treating depression, anxiety, and conduct problems in youth: A randomized effectiveness trial. Arch Gen Psychiatry 2012; 69(3):274-82.
32. Kazdin AE, Rabbitt S. Novel models for delivering mental health services and reducing the burdens of mental illness. Clin Psychol Sci 2013; 1:170-91.
33. Kessler R, Aguilar-Gaxiola S, Alonso J, et al. The global burden of mental disorders: An update from the WHO World Mental Health (WMH) surveys. Epidemiol Psichiatr Soc 2009; 18:23–33.
34. Kessler RC, Wang PS. The descriptive epidemiology of commonly occurring mental disorders in the United States. Ann Rev Pub Health 2008; 29:115-29.
35. Druss BG, Hwang I, Petukhova M, et al. Impairment in role functioning in mental and chronic medical disorders in the United States: Results from the National Comorbidity Survey Replication. Mol Psychiatry 2009; 14:728-37.
36. The great push: Investing in mental health. World Health Day 2011. Baltimore, MD: World Federation for Mental Health; 2011.
37. Task-shifting: Global recommendations and guidelines. Geneva: World Health Organization; 2008.
38. Kessler RC, Demler O, Frank RG, et al. Prevalence and treatment of mental disorders, 1990 to 2003. N Engl J Med 2005; 352:2515-23.
39. Kazdin AE. Evidence-based treatment and usual care: Cautions and qualifications. JAMA Psychiatry 2013 70:666-7.
40. Hoge MA, Morris JA, Daniels AS, et al. An action plan for behavioral health workforce development. Washington, DC: U.S. Department of Health and Human Services; 2007.
41. Institute of Medicine. Child and adolescent health and health care quality: Measuring what matters. Washington, DC: National Academies Press; 2011.
42. Institute of Medicine. The mental health and substance use workforce for older adults: In whose hands? Washington, DC: National Academies Press; 2012.
43. Kazdin AE, Blase SL. Rebooting psychotherapy research and practice to reduce the burden of mental illness. Perspect Psychol Sci 2011; 6:21-37.
44. Patel V, Weiss HA, Chowdhary N, et al. Effectiveness of an intervention led by lay health counsellors for depressive and anxiety disorders in primary care in Goa, India (MANAS): A cluster randomised controlled trial. Lancet 2010; 376:2086–95.
45. Christensen CM, Grossman JH, Hwang J. The innovator's prescription: A disruptive solution for health care. New York: McGraw Hill; 2009.
46. Rotheram-Borus MJ, Swendeman D, Chorpita BF. Disruptive innovations for designing and diffusing evidence-based interventions. Am Psychol 2012; 67(6):463-76.
47. Linnan LA, Kim AE, Wasilewski Y, et al. Working with licensed cosmetologists to promote health: Results from the North Carolina BEAUTY and Health pilot study. Prev Med 2001; 33:606-12.
48. Madigan ME, Smith-Wheelock L, Krein SL. Healthy hair starts with a healthy body: Hair stylists as lay health advisors to prevent chronic kidney disease. Prev Chron Dis 2007; 4:A64.
49. Chisholm D, Lund C, Saxena S. Cost of scaling up mental healthcare in low- and middle-income countries. Br J Psychiatry 2007; 191:528-35.
50. World Health Organization. Prevention and control of NCDS: Conference proceedings. First Global Ministerial Conference on Healthy Lifestyles and Noncommunicable Disease Control. April 28-29, 2011; Moscow. Available at http://www.who.int/nmh/

events/moscow_ncds_2011/conference_documents/conference_report.pdf?ua=1. Accessed July 24, 2014.
51. Deslandes A, Moraes H, Ferreira C, et al. Exercise and mental health: Many reasons to move. Neuropsychobiol 2009; 59:191-8.
52. Walsh R. Lifestyle and mental health. Am Psychol 2011; 66:579-92.
53. Baker JR, Moore SM. Blogging as a social tool: A psychosocial examination of the effects of blogging. Cyberpsychol Behav 2008; 11:747-9.
54. Gorini A, Gaggioli A, Vigba C, et al. A second life for ehealth: Prospects for use of 3-D virtual worlds in clinical psychology. J Med Internet Res 2008; 10(3):e21.
55. Singhal A, Rogers EM. Entertainment-education: A communication strategy for social change. Mahwah, NJ: Lawrence Erlbaum; 1999.
56. Singhal A, Cody MJ, Rogers EM, et al. (Eds.). Entertainment-education and social change: History, research, and practice. Mahwah, NJ: Lawrence Erlbaum; 2003.
57. Muñoz RF. Using evidence-based internet interventions to reduce health disparities worldwide. J Med Internet Res 2010; 12:e60.
58. Bluthenthal RN, Jones L, Fackler-Lowrie N, et al. Witness for Wellness: Preliminary findings from a community-academic participatory research mental health initiative. Ethn Dis 2006; 16(1):S1-18, 34.
59. Wells KB, Jones L, Chung B, et al. Community-partnered cluster-randomized comparative effectiveness trial of community engagement and planning or resources for services to address depression disparities. J Gen Intern Med 2013; 28(10):1268-78.
60. Bang AT, Reddy HM, Deshmukh MD, et al. Neonatal and infant mortality in the 10 years (1993-2003) of the Gadchiroli field trial: Effect of home-based neonatal care. J Perinatol 2005; 25:92-107.
61. Greenwood BM, Greenwood AM, Snow RW, et al. The effects of malaria chemoprophylaxis given by traditional birth attendants on the course and outcome of pregnancy. Trans R Soc Trop Med Hyg 1989; 83:589-94.
62. Institute of Medicine. Promoting cardiovascular health in the developing world: A critical challenge to achieve global health. Washington, DC: National Academies Press; 2010.
63. Institute of Medicine. Country-level decision making for control of chronic diseases. Washington, DC: National Academies Press; 2011.
64. United Nations. United Nations Millennium Declaration. Resolution adopted by the General Assembly, 8th Plenary Meeting, Sept 8, 2000. Available at http://www.un.org/millennium/declaration/ares552e.htm. Accessed July 24, 2014.
65. Patel V, Weiss HA, Chowdhary N, et al. Lay health worker led intervention for depressive and anxiety disorders in India: Impact on clinical and disability outcomes over 12 months. Br J Psychiatry 2011; 199:459–66.
66. Balaji M, Chatterjee S, Koschorke M, et al. *The development of a lay health worker delivered collaborative community based intervention for people with schizophrenia in India.* BioMed Central Health Serv Res 2012; 12. Available at www.biomedcentral.com/1472-6963/12/42/. Accessed July 24, 2014.
67. Rahman A, Malik A, Sikander S, et al. Cognitive behaviour therapy-based intervention by community health workers for mothers with depression and their infants in rural Pakistan: A cluster-randomised controlled trial. Lancet 2008; 372:902-9.
68. Bloom DE, Cafiero ET, Jané-Llopis E, et al. The global economic burden of noncommunicable diseases. Geneva: World Economic Forum; 2011. Available at www.weforum.org/EconomicsOfNCD. Accessed July 24, 2014.
69. World Health Organization. Scaling up action against non-communicable diseases: How much will it cost? Geneva: WHO; 2011.
70. Chisholm D, Saxena S. Cost effectiveness of strategies to combat neuropsychiatric conditions in sub-Saharan Africa and South East Asia: Mathematical modelling study. Br Med J 2012; 344:e609.
71. Prince M, Patel V, Saxena S, et al. No health without mental health. Lancet 2007; 370:859–77.
72. World Health Organization. Mental health: Facing the challenges, building solutions. Report from the WHO European Ministerial Conference. Regional Office for Europe, p. 11. Copenhagen: WHO; 2005.
73. Lancet Global Mental Health Group. Scale up services for mental disorders: A call for action. *Lancet 2007; 370:1241-52.*
74. Sharan P, Gallo C, Gureje O, et al. Mental health research priorities in low- and middle-income countries of Africa, Asia, Latin America and the Caribbean. Br J Psychiatry 2009; 195:354-63.
75. Clark DM. Implementing NICE guidelines for the psychological treatment of depression and anxiety disorders: The IAPT experience. Int Rev Psychiatry 2011; 23(4):318-27.
76. Draper C, O'Donohue W. (Eds.) Stepped care and e-health: Practical applications to behavioral disorders. New York: Springer Science + Business Media, LLC; 2011.
77. Haaga DA. Introduction to the special section on stepped care models in psychotherapy. J Consult Clin Psychol 2000; 68(4):547-8.
78. L'Abate L. Clinical psychology and psychotherapy as a science. New York: Springer; 2013.

79. Newman FL, Howard KI. Therapeutic effort, treatment outcome, and national health policy. Am Psychol 1986; 41:181-7.
80. Duman CH, Schlesinger L, Russell, et al. voluntary exercise produces antidepressant and anxiolytic behavioral effects in mice. Brain Res 2008; 1199:148-58.
81. Wolff E, Gaudlitz K, von Lindenberger BL, et al. Exercise and physical activity in mental disorders. Eur Arch Psychiatry Neurol Sci 2011; 261(2):186-91.
82. Miranda J, Bernal G, Lau AS, et al. State of the science on psychosocial interventions for ethnic minorities. Ann Rev Clin Psychol 2005; 1:113-42.
83. Kazdin AE. Evidence-based treatments and delivery of psychological services: Shifting our emphases to increase impact. Psychol Serv 2008; 5:201-15.
84. Bardo MT, Pentz MA. Translational research. In Cooper H. (Ed.) APA handbook of research methods in psychology. Washington, DC: American Psychological Association; 2012. Vol. 3, pp. 553-68.
85. Woolf SW. The meaning of translational research and why it matters. JAMA 2008; 299(2):211-3.
86. Davis M. NMDA receptors and fear extinction: Implications for cognitive behavioral therapy. Dialogues Clin Neurosci 2011; 13(4):463-74.
87. Wampold BE, Budge SL, Laska KM, et al. Evidence-based interventions for depression and anxiety versus intervention-as usual: A meta-analysis of direct comparisons. Clin Psychol Rev 2011; 31:1304–12.
88. Kazdin AE, Whitley MK. Comorbidity, case complexity, and effects of evidence-based treatment for children referred for disruptive behavior. J Consult Clin Psychol 2006; 74:455-67.
89. Henggeler SW. Efficacy studies to large-scale transport: The development and validation of multisystemic therapy programs. Ann Rev Clin Psychol 2011; 7:351-81.
90. Jensen-Doss, AJ, Weisz JR. Syndrome co-occurrence and treatment outcomes in youth mental health clinics. *J Consult Clinl Psychol 2006; 74*:416-25.

Alan E. Kazdin, PhD, is Sterling Professor of Psychology, Professor of Child Psychiatry, and Director of the Yale Parenting Center, a service for children and families. While at Yale, he has been Chairman of the Psychology Department, Director of the Yale Child Study Center at the School of Medicine, and Director of Child Psychiatric Services at Yale-New Haven Hospital. Previously, he was on the faculty of The Pennsylvania State University and the University of Pittsburgh School of Medicine. He has received the Outstanding Research Contribution by an Individual Award and Lifetime Achievement Awards from the Association of Behavioral and Cognitive Therapies, the Outstanding Lifetime Contributions to Psychology Award and Distinguished Scientific Award for the Applications of Psychology from the American Psychological Association, and the James McKeen Cattell Award from the Association for Psychological Science. In 2008, Dr. Kazdin served as president of the American Psychological Association.



Emerging Tools for Studying Population Health

Mathematical and Computational Simulation of the Behavioral and Social Drivers of Population Health

Mark G. Orr, Bryan Lewis, Kathryn Ziemer, and Sallie Keller

Abstract

The use of computational and mathematical simulations is gaining momentum in population health. Specifically, these methods are particularly useful when examining population health issues that implicate dynamic behavioral and social processes. The goal of this chapter is to familiarize the reader with a variety of modeling approaches, across scales and types of phenomena, and to discuss emerging issues and directions. We review some examples of recent work that addresses research questions that are difficult to penetrate precisely because of the complexity they represent: (1) the dynamics of health behavior within the individual, (2) policy effects on the black/white disparities in obesity-related behaviors, and (3) the effects of human behavioral choices on the spread of infectious disease. Then, we address some of the implications of simulation and modeling for public health practice ranging from the obvious (e.g., incorporate simulation into public health practice where appropriate) to the more subtle (e.g., simulation may highlight unforeseen data needs) and offer guidance for potential users of the simulation approach. The directions for future research are many and multifaceted; we emphasize two key challenges in moving forward. The incorporation of "Big Data" into simulations is welcome and inevitable but comes with several manageable but substantial obstacles. Finally, we suggest that the key challenge for moving forward is cultural. How do we build consensus that simulation should be a core methodological approach in population health?

Introduction

Defining Computational/Mathematical Simulation and Its Purpose

A natural confusion arises when one tries to explain computational and mathematical modeling and the reasons why we use it to aid in understanding and leveraging changes in population health. Such confusion probably stems from the ubiquity of the statistical modeling approach in population health, a highly successful venture that has aided greatly in determining which variables are associated with key health-related outcomes. Computational and mathematical modeling has a different but related focus—i.e., to understand the systems, processes, and related dynamics of

population health phenomena. Most often, this is accomplished using computer simulation of some sort, a method for understanding how the dynamics of a system unfold.

The strength of the computational/mathematical modeling approach is its focus on dynamics, feedback loops, and interdependent non-linear processes, all of which are notoriously difficult to estimate[1] with the statistical modeling approaches commonly used in population health. Furthermore, it affords a useful mode for thinking outside of the available data, towards future data, and potentially, about what is not yet known. In short, computational and mathematical modeling offers an alternative and unique perspective in relation to statistical modeling. Thus, the two approaches are well poised to be mutually informative.

With the exception of mathematical models of infectious disease, the use of simulation in population health was relatively novel until recently (for two early exceptions, see Morris and Kretzschmar[2] and Weinstein, Coxson, Williams, et al.[3]). The systems science approach, coming into prominence in the 2000s,[4-7] has helped to spur the recognition that the tools used to study complex systems, including simulation, can be useful for studying population health phenomena, especially in three areas—theory development, intervention/prevention, and policy—all of which use simulation for understanding the implications of an idea, ***X***, specifically, for providing insight into three types of "what-if" scenarios. In theory development, the "what-if" can be translated to: what if ***X*** is a true process (e.g., would the HIV mortality rate increase or decrease, given an increase in antiretroviral therapy if it was known or hypothesized that ART caused an increase in risky behavior of a specific magnitude in a gay community)?[8] In intervention/prevention work, the question becomes: what if we intervene by doing ***X*** (e.g., what outcomes can we expect from different social network-driven intervention strategies to reduce or halt the spread of obesity along social channels)?[9] In policy, the "what-if" considers the effects of policy ***X***—e.g. to what degree would the prevalence of cigarette use change given a ban on menthol cigarettes?[10]

Our primary task here is to probe into the state of the art in simulation modeling that is of direct relevance to population health and that takes into account social and behavioral processes. We will do this by providing examples from the literature of selected cases describing recent work in this area. The examples we review were selected because they represent recent attempts to provide novel insights into difficult-to-understand phenomena—phenomena that call for the representation of complex and dynamic processes.

The first example focuses on individual-level health behavior theory (e.g., Reasoned Action Theory,[11] Health Belief Model,[12] Social Cognitive Theory[13]), borrowing techniques from cognitive science and computational psychology, to provide insights into how people dynamically integrate past experience with current social and environmental contexts to inform behavior. The second example, nearly opposite the first example in scale, addresses ways to leverage what are considered dynamic macrosocial determinants to reduce racial disparities in obesity-related behaviors. The third example explores the importance of understanding dynamic social and behavioral processes with respect to the spread of infectious disease within large segments of the U.S. population. Finally, we will explore the implications for public health practice and directions for future research.

Three Examples

Individual-Level Health Behavior

The closely related fields of health behavior change, health behavior and education, and behavioral medicine share the conviction, rightfully, that theoretically driven intervention and prevention efforts are more successful than those that are not theoretically-driven. The published literature supports this conviction.[14,15] Theory provides organizing frameworks for understanding what to measure, how to measure, and from whom to measure. Despite a deep commitment to theory, the health behavior field has yet to forge a strong bridge to population health, especially in reference to dynamic processes (e.g., the effects of peer influence and the built environment on behavior). These types of processes, although acknowledged in health behavior theory, are not well understood at the level of the individual and thus are not yet well integrated with the population health approach.

A significant barrier to integrating health behavior theory and population health is the dearth of work in behavior change that leverages simulation at the individual level of analysis. The exception is the work we describe next, previously published work of our own[16] that was designed to explain some of the more difficult issues in health behavior theory related to the dynamics of behavior and learning—an issue for which simulation is well-suited if not required.

Our work re-conceptualizes health behavior theory using computational modeling practices developed in the fields of cognitive science and computational psychology (i.e., the general study of the mind as an information processing system). Health behavior theory, which stems largely from psychological principles and constructs is at its essence about information processing; beliefs, learning by imitation, valuation, and so on are fully engaged with the notion of information and information processing.

Specifically, we developed a computational model of Reasoned Action Theory[11] (hereafter RAT), a well-known theory that explains the performance (or not) of a behavior as driven by one's intention to perform the behavior where intention is driven by attitudes, norms, and perceptions of behavioral control, each of which is driven by a set of beliefs related to the behavior in question (i.e., beliefs → attitude, norms, control → intention). The innovation of our model is that it attempts to capture the dynamics of what we call intention formation, a conceptualization of how intention is generated on-the-fly that accounts for: (1) what a person has learned about a behavior (in terms of related beliefs) from past experience via social learning, and (2) the pressures from the more immediate social context also with respect to the same set of beliefs. For example, imagine this situation: a drip coffee aficionado is in conversation at a cocktail party hosted by the Society for Espresso Drinking on the topic of reasons to abandon drip coffee in favor of espresso (two very different types of coffee). Our model attempts to capture the on-the-fly intention formation (to consider switching from drip to espresso) of the drip coffee aficionado given his/her past experiences and the current conversation at the cocktail party. Outside of this more limiting case, consider these other, similar examples directly relevant to population health and health behavior: an adolescent moves to a new school and community; a woman is exposed to tobacco-related point-of-sale advertising; and, finally, a man changes jobs, relocating to a new community and new company.

Figure 1 represents the formal structure of our model—most prominent is the representation of each belief as a coupled set of memory units, one representing positive valence and the other representing negative valence that can vary in activation from not active to fully active. Our model assumes that the memory units are activated or cued by social exposure to relevant beliefs. Once cued, the system settles into a local equilibrium that is dictated by both the presence of the cued beliefs and the set of connection weights between memory units. The set of connection weights, importantly, are derived from past experience and encode a learned, cultural belief structure. The equilibrium state, in the aggregate across beliefs, captures the formation of intention given the cue—i.e., the current and immediate social context. This model structure was used as the basis for the simulation we describe next.

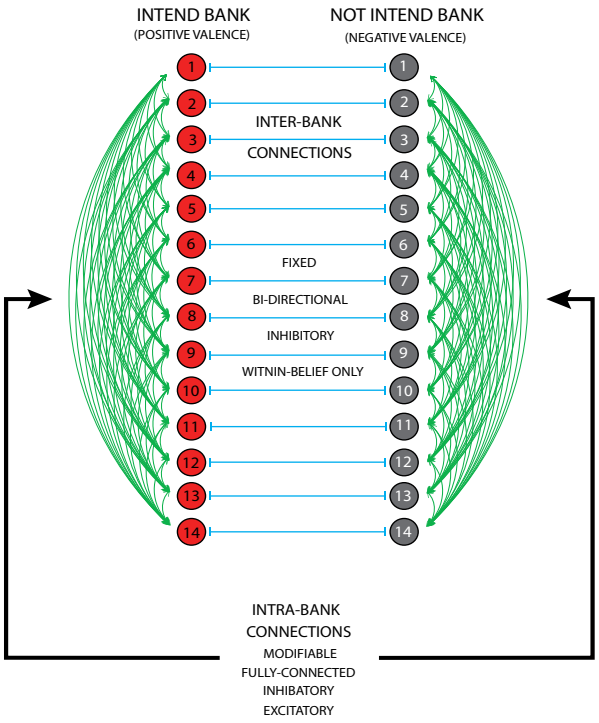


Figure 1. Schematic of the computational model of the Theory of Reasoned Action

Source: Borrowed with permission from Orr MG, Thrush R, Plaut DC. The Theory of Reasoned Action as parallel constraint satisfaction: towards a dynamic computational model of health behavior. Plos One 2013;8:e62409. Used with permission under Creative Commons Attribution (CC BY) license.

We present the main results of a simulation that attempted to capture a person's attitude formation given a shift in social contexts. Specifically, we focused on sexual behavior in adolescent females and a shift in social context from exposure to less positive to exposure to more positive beliefs about sexual intercourse. To do this, we trained our model to learn the belief structure in the context of 10th grade females who had never had sex. (The data to train our model came from a

Reasoned Action Survey conducted in a U.S. school district in the Northwest[17]). Then, we shifted contexts of the model to 12th grade females who had never had sex and 12th grade females who had previously had sex. Thus, this simulation procedure probed the degree to which intention formation was constrained simultaneously by past (being exposed to 10th grade non-experienced females) and current contexts (sexually experienced females). The key finding for this simulation was that the immediate social context was able to influence but not override what was learned from past experience—i.e., both the past and the present were important. Specifically, while exposure to more positive beliefs about sexual intercourse in the immediate social context did in fact increase the intention to have sexual intercourse, this tendency was dampened by past learning from contexts with less positive beliefs about sexual intercourse.

One might be tempted to argue that these results reflect no more than common sense: yes, of course both the past and the present impact health behaviors. This interpretation, however, misses the point. What we have accomplished in this simulation is highly important for the health behavior field—the development of a formal and mechanistic account of how the past and present are simultaneously and seamlessly integrated to generate an intention state. The formalism, called constraint satisfaction, was derived from prior work in cognitive science[18,19] and social psychology[20] (to model personality, attitudes, and cognitive consistency) and thus has a non-negligible degree of generality.

Insights from our simulations were: (1) our model allows for explicit representation of how beliefs might be formed (this is not addressed to any satisfaction in the current health behavior literature); (2) it affords an integration directly with other types of simulations, ones that focus on population health phenomena (e.g., influence spreading on social networks); and (3) it provides very explicit, novel, and testable hypotheses (e.g., about the assumed belief structure and the learning rates of beliefs). We point the reader to another paper of ours that lays out more completely the implications of this approach for health behavior.[21] In this work, we introduced into the literature the term "computational health behavior modeling" in an attempt to drive the development of a sub-discipline within the health behavior field—one that calls for the integration of contemporary health behavior theory and simulation.

Macrosocial Drivers of Disparities in Health Behavior

The macrosocial approach to population health addresses social processes that are potentially amenable to change at the societal scale.[22] Changes in macrosocial factors are uniquely promising because of the potential for widespread and long-standing societal and structural changes that, in turn, may support long-term population-level changes in health. We focus here on the macrosocial drivers of health-related behaviors, an important aspect of non-communicable disease and one that has large potential to benefit from the simulation approach due to the potential for long- and short-term feedback among the many related components of the system.

Figure 2 illustrates the macrosocial approach. Macrosocial processes represent factors such as income and educational distributions, crime rates, political will and policy, and corporate practices that are amenable to change at the societal level. The macro-behavioral interface represents the way

in which macrosocial processes shape peoples' physical and social contexts (e.g., norms; the built environment). The behavioral process level captures individual-level behaviors that are directly related to outcomes (e.g., physical activity and dietary behaviors →obesity). Figure 2 allows us to think about how to change health behaviors related to health outcomes in a way that respects the need to (1) make long-term structural changes, (2) engender contextual changes in individual-level behaviors, and (3) address the population as a whole. A key feature in Figure 2 is the bi-directional arrows, whereby the influences between the macrosocial processes, the macro-behavior interface, and health behavior are mutual and bidirectional, thus illustrating the potentials for dynamic feedback among these levels of analysis.

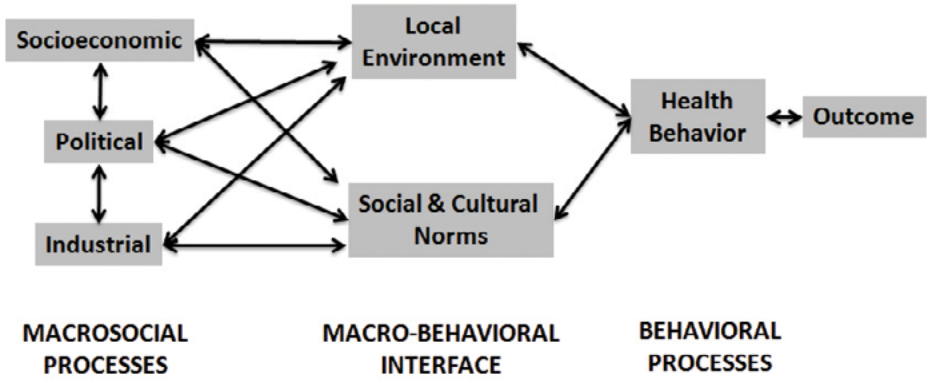


Figure 2. Schematic of the macrosocial approach to understanding population health dynamics

In this section, we present recently published work of our own that used simulation to address a pressing population health issue—racial/ethnic disparities in obesity.[23] Over one-third of the U.S. adult population is obese (35.7 percent),[24] leading to increasing rates of chronic disease (heart disease, stroke, type 2 diabetes, some cancers) and medical costs ($147 billion in 2008).[25] The obesity problem is even more problematic for key racial/ethnic groups. The highest age-adjusted obesity rate is among non-Hispanic blacks (49.5 percent), followed by all Hispanics (39.1 percent), and non-Hispanic whites (34.3 percent).[26]

The basis for the simulation was an agent-based model that was developed within the macrosocial framework to capture behavioral, social, and environmental determinants of diet and physical activity, the primary determinants of obesity. For the simulation presented here, we concentrate on diet-related behavior alone (a measure of the quality of dietary intake as driven by individuals' decisions and choices).

The structure of the model was designed to represent the racial and economic distributions of black and non-Hispanic whites in the 100 largest metropolitan statistical areas in the United States. Each agent resided in a household within a specific neighborhood and underwent a simulated developmental trajectory (birth, schooling, workforce, retirement, death) during which the agent engaged in health-related behaviors (smoking, dietary intake quality, degree of physical activity) to

generate health outcomes (body mass index [BMI] among these). The determination of diet quality for each agent was structured so that neighborhood-level variables (school quality and access to good food stores) had a direct effect on each agent's dietary choices, as did the behavior of other agents via social network influence. Figure 3 shows the structure of the model in relation to dietary behavior.

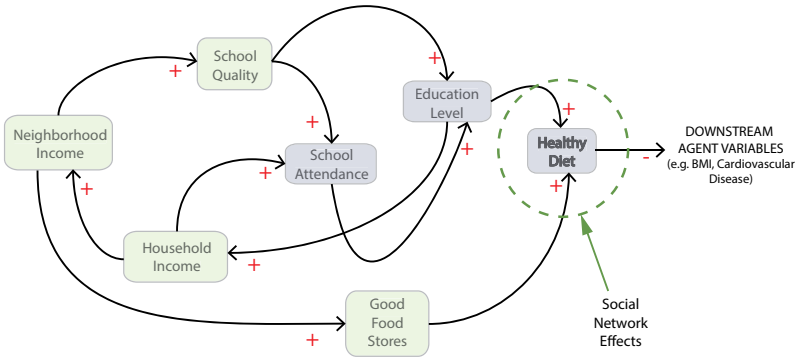


Figure 3. Causal structure of the agent-based model of macrosocial drivers of obesity disparities

Source: Borrowed with permission from Orr MG, Galea S, Riddle M, et al. Reducing racial disparities in obesity: simulating the effects of improved education and social network influence on diet behavior. Ann Epidemiol 2014;24:563-9. Used with permission from Elsevier.

Using this model, we simulated the potential long-term (~2.5 generations in terms of agent lifecycles) effects of a social policy—improving the quality of schools in neighborhoods that suffer from poor school quality—on racial disparities in diet quality. We used the ratio of students to teachers as a proxy for school quality. For the policy manipulation, we simulated targeting neighborhoods in the lowest 20 percent of school quality where we reduced the student-to-teacher ratio by about 60 percent (from about 15:1 to 7:1). It is important to keep in mind that the policy manipulation was designed to not target race directly and was of considerable strength.

The simulations were focused on two questions: Question 1: To what extent does targeting school quality have an effect on racial disparities in dietary quality? Question 2: Are the effects of targeting school quality self-sustaining (what happens when we stop implementation of the policy)?

Regarding question 1, we found that targeting school quality reduced the black/white disparity in dietary quality by about 40 percent. Given the timeframe of the simulation (2.5 generations), we did not see any evidence that the disparity would be reduced further given more time. This probably stemmed, in part, from the structure of our manipulation. By targeting neighborhoods and not race, we effectively targeted about 40 percent of blacks and 15 percent of whites, so both blacks and whites saw an improvement in dietary quality. It is of interest that over time, the school quality manipulation affected the system globally. Both the levels of gross household income and the quality of food stores increased noticeably by the end of the simulation—the former driven by the increase in agents' combined education levels and the latter driven by the former (see Figure 3).

Regarding question 2, because of the presence of social network effects (on dietary behavior) and the potential for system-level effects, we were surprised to find that the effects of the policy manipulation were not self-sustaining. Once the policy was no longer in effect, the initial disparities returned. Although we have yet to pinpoint exactly why the policy was not self-sustaining, this feature of the simulation points to the potential complexity in understanding dynamic processes that include many interdependent agents (in this case people).

In summary, our simulation was supportive of macrosocial policies and in line with the literature on the determinants of population health and racial disparities. The value here, really, is to illustrate the potential of simulation modeling for thinking through the complexity of such issues in population health. We would like to emphasize that these conclusions must only be taken as a suggestion for the potentials of macrosocial, upstream policy effects, and the potential issues to consider, and that the reader should not rely on the precise numerical results. The simulation was designed primarily as a proof-of-concept to show these potentials. We urge readers to consult the original work for details on the model assumptions, data calibration and validation, limitations, and simulation results.

Population-Level Infectious Disease and Individual Behavior

The modeling and simulation of population-level dynamics of the spread of infectious disease have undergone a wholesale increase over the past decade in the number of parameters used. To a large degree, the newer parameters capture aspects that are crucial for understanding and potentially predicting epidemics (e.g., inclusion of social networks and geospatial data). Not surprisingly, some of these parameters attempt to account for somewhat complex social and behavioral processes and policies, such as zoning laws, court-ordered quarantine, and vaccination requirements for school entry. Although such policies can work, they are susceptible to variations in human behavior and social processes. Consider the tension between herd immunity and the so-called anti-vaccination movement in the United States today—the effectiveness of our policies, and thus our herd immunity, is at risk of degradation due largely to social and behavioral processes (e.g., spread of values and attitudes).

The field of computational epidemiology arose in part from the supposition that in order to study relevant policy implications, not only did we need to improve our ability to rapidly simulate the large-scale spread of infectious diseases,[27-31] but we needed to do so while incorporating key social and behavioral processes. Through the use of simulation techniques, we now have support that this supposition was true; incorporating individual-level behavior when attempting to model population-level infectious disease dynamics yields obvious benefit.

Here we describe two simulation studies[32,33] where we seek to explore the potential impact of individual-level behaviors on infectious disease dynamics under the condition of heterogeneity among individuals. These simulation studies represent individuals' behavior as a set of simple rules that alter the risk of exposure to or further transmission of disease. These studies were done with an agent-based network diffusion simulation platform, where the synthetic populations are painstakingly constructed through the fusion of many data sources (U.S. census, time-use, etc.) such that the household structures, geo-spatial distributions of demographics, and time use of the populations are preserved to the census block level. As illustrated in Figure 4, and described in

more detail elsewhere,[34-36] census information at the block group level is further disaggregated to the individual level, while maintaining the joint distributions of demographics most predictive of time use patterns. The time use surveys are then assigned to demographically appropriate households and used in a microsimulation where every member of the population is assigned locations for each of their activities; steps of iterative refinement using capacity information for these locations (e.g., enrollment data for schools) ensures the locations have the appropriate numbers of occupants and preserves the distribution of commuting distances. Once the synthetic society is constructed, discrete event simulation platforms are used to represent the percolation and time course of the processes being modeled. In the case of infectious diseases, the duration of various states of infectiousness and incubation, as well as differential levels of infectiousness based on receiving different treatments, are represented. These simulations result in a single realization of the complex processes, and thus are run many times to generate distributions of possible outcomes. The studies discussed below were conducted on large populations capturing a wide diversity of demographics, specifically synthetic representations of Miami and Seattle (2.1 million and 3.2 million individuals respectively).

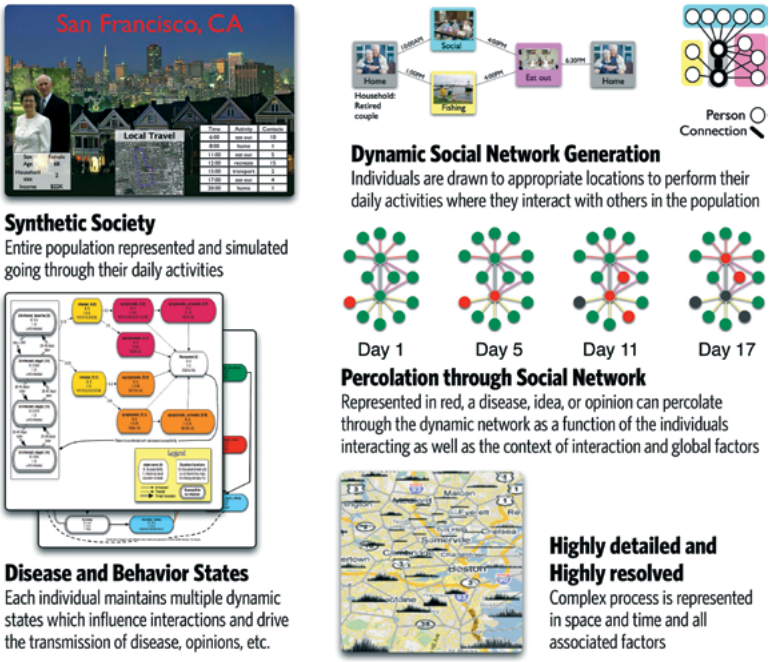


Figure 4. Basic ingredients for a highly-detailed agent-based synthetic population

The first study[33] illustrates the importance of care-taking behaviors within a household on the extent of disease spread across the population. Specifically, it looks at how a household alters its contact patterns when a single household member is symptomatic with influenza. Two extremes of responses are explored for comparison's sake. The first "single care-taker" scenario has the symptomatic individual remove themselves from contact with all other members of the household except one person who cares for the sick individual and thus is at risk for infection. The opposite extreme, is that all members behave the same regardless of the symptoms of any of their members and are thus in contact with other household members while engaged in their home activity. To

study the impact of these two behaviors, simulations were run for outbreaks of moderate to high levels of infection (cumulative attack rates of ~15 percent, and ~25 percent respectively) and in two cities with different demographic profiles (Miami: larger household sizes and older population; Seattle: smaller households and younger population). The "single care-taking" behavior was enough to reduce infection rates 10-88 percent, with the greatest reductions occurring for epidemics most similar to an average influenza season. These reductions were analyzed across household sizes, which showed the reduction to be larger in larger households, further reinforcing the importance of the behavior within the household. Certainly the degree to which these behaviors are truly followed and how they are distributed in the population rests somewhere between the extremes analyzed; however, this study demonstrates that these behaviors are an important element of disease transmission.

The second study[32] compared the cost-effectiveness of an individually motivated behavior against a more top-down, driven behavior. This was motivated by the need to evaluate policy choices between subsidies designed to enable more choice and more directed interventions requiring closer surveillance. Two behaviors were studied as they were dynamically applied during the course of an outbreak of influenza. The individually motivated behavior is based on a person's desire to seek care, in this case get a vaccine or a course of antiviral prophylaxis, to reduce their chance of infection. In the simulation, this motivation is queued when the number of symptomatic people in an individual's social network exceeds a certain threshold, so that the individual's local social context becomes the driving force for mitigating the outbreak. This was compared to a more top-down approach that offers these interventions to sections of the population (census blocks or households with children in a particular school), when the numbers of people in that section exceeds a threshold. Care was taken to make these thresholds comparable and to explore the sensitivity of adherence with the offers. Even under the extreme conditions where adherence was absolute in the top-down branch, the number of cases averted per vaccine or course of antiviral prophylaxis was highest in the individually motivated care-seeking case. This finding was robust across a range of thresholds, diagnosing rates, and outbreak sizes. In many of these cases the number of cases averted per course is two orders of magnitude larger. This demonstrates the tremendous power of enabling individuals to take actions based on their particular environment rather than broadly applying actions to groups of individuals, even when precise aggregate knowledge about these groups is known. This highlights the profound effectiveness that individual-level behaviors can have in mitigating future cases of infectious diseases, especially when individuals are enabled to take effective action.

In summary, infectious disease dynamics at the population level are absolutely driven by individual-level behaviors. The studies described here illustrate how critically important it is to take these behaviors into account when simulating and argue for inclusion of individual-level behaviors in all simulation-based analyses of infectious disease spread. While the implementation of individual-level behaviors in these studies is still based on relatively simple rules, by taking the context of the individual into account, these simulations are a step in the direction of including individual-level behaviors into large-scale simulations of infectious disease spread.

Implications for Public Health Practice

Hopefully, we have convinced the reader that the simulation approach has much to offer in terms of understanding population health phenomena, especially those that involve social and behavioral processes. As described above, the principal appeal of simulation is that it allows for thinking in terms of "what-if" scenarios related to theory development, intervention/prevention, and policy. So, the implication is simple: Public health practice, at least when social and behavioral constructs are implied, should incorporate "what-if" simulations whenever feasible. In fact, we want to push people to think of new and creative ways to ask "what-if?" Beyond this rather broad assertion, however, are the details. Under what conditions should we ask "what-if"? Are there classes of questions that are more amenable to "what-if" scenarios? And so on.

Our collective experiences provide some insights into these questions. First and foremost, developing a simulation is fundamentally about developing formal algorithms to study the time-dependent processes thought to underlie a phenomenon. So, when developing a "what-if" scenario, keep in mind that a set of formal algorithms will represent the "what-if." Although this might initially seem to constrain the potential questions you can pose (e.g., How do I make an algorithm for how self-efficacy changes?), this is not necessarily the case. A shift in how you see a particular problem or phenomenon may yield the development of feasible algorithms. Second, for many phenomena of interest, multiple levels of analysis will be represented in a simulation. Much of the difficulty will come from defining algorithms that address the interplay between levels of analysis (e.g., the interplay between a person's attitude and his/her built environment) because this is probably where there is a dearth of theoretical and empirical work. Third, there are different classes of algorithms; some are strictly probabilistic (e.g., given X days of exposure to others smoking, my chance of smoking will increase by Y), some are more mechanistic, and some are both. Being mindful of these types of distinctions can aid in the development of algorithms and help to sort out the relation between the data and theory that might serve to ground the algorithms.

Finally, we suggest strongly that one should get involved in a direct way with building and using simulations, either through collaboration with experienced modelers or by self-guided exploration with dedicated software.[a] Furthermore, there are useful textbooks that introduce the simulation approach for behavioral and social phenomena.[37,38]

Two other important but subtle implications remain for consideration. First, simulations have the potential to develop new social and behavioral theory or expand on existing theory (e.g., the individual-level health behavior simulation presented above) and thus may offer an opportunity for public health practitioners to apply new theoretical advances. In other words, public health practitioners should be concerned with both the potential of using simulations for thinking "what-if" and the integration of new theory arising from simulation work into practice. Second, the development of the algorithms that drive a simulation often draw a stark picture of what we don't know. This can be tremendously fruitful for both the science and practice of public health.

[a] The free software platform Netlogo serves this purpose very well. See https://ccl.northwestern.edu/netlogo/.

Directions for Future Research

The simulation modeling approach in population health is a method that weaves together several threads in applied mathematics, computer science, and complex systems. It represents some theoretical first principles on the nature of complexity and its potential characteristic signatures, as well as theoretical constructs from topical fields such as genetics, psychology, sociology, economics, public policy, and social epidemiology. When incorporated with social and behavioral sciences, the number of possibilities becomes rather large. In this light, we feel presumptuous to provide specific directions for future research. So, instead, we focus the directions for future research on the very important issue of how to increase the extent to which the simulation approach is useful for the understanding of and intervention in population health phenomena that implicate social and behavioral processes. We have identified two central challenges in moving forward: (1) incorporation of "Big Data" to be used for empirical grounding of the simulations and (2) integration of simulation modeling into mainstream population health. We address each in turn.

Challenges in Empirical Grounding with the Advent of "Big Data"

In the context of population health, empirical grounding refers to developing confidence that the assumptions of a simulation model represent the processes that occur in the real world. This issue is not new, not in population health modeling or in any other fields that use computational or mathematical simulation. At this point in history, the consideration of issues at the intersection of empirical grounding, population health, and the social and behavioral sciences must seriously consider the role of the so-called Big Data revolution. The availability of massively unstructured data on human behavior and social process is increasing beyond precedent.

Generally, there are two principal sources of data to consider: administrative data and digital data. Administrative data are data collected for the administration of an organization or a program. Examples of these include Internal Revenue Service data for individuals and businesses, Social Security earnings records, Medicare and Medicaid health utilization data, and credit card transactions. Of particular interest for population health simulations is the use of data from electronic health records (EHRs) used by large, organized health systems, such as Kaiser Permanente, and by Federal entities, such as the Centers for Disease Control and Prevention. Digital data are data being reported for purposes other than statistical or administrative use. These data come from all variety of sensors, including mobile and wearable devices, and through the use of the Internet (e.g., Twitter). All of this has the potential to provide extremely large volumes of data in near real-time.

These two types of data offer possibilities for studying behavior and social drivers of population health at a finer level of geographic and/or demographic resolution and in more frequent time intervals, as well as studying human interactions at a societal scale, with rich spatial and temporal dynamics. In contrast to designed data collection, the data volume per unit cost is cheap. Crowd sourcing organizations, such as Amazon Mechanical Turk or Jana.com, exist and offer inexpensive venues to collect data. The clear cost driver is no longer the data collection but rather the development and execution of the data analytics, a significant game-changer in how to think about social and bio-informatics and the interface with simulation modeling.

Naturally, there are issues, both known and unknown, with these new data sources. They have unknown quality, population representativeness, and statistical properties. These factors may be knowable, but they simply have not been well studied. Also, the data come with little to no documentation about coverage, representativeness, bias, and longitudinal gaps in the data. These data may present time continuity problems, e.g., companies may merge or change focus. Finally, as exemplified by EHRs, are the challenges of interoperability and privacy concerns. Many of the data formats of EHRs lack interoperability, which in turn, hinders effective exchange or integration of health information. None of these issues are insurmountable; thus, we should look forward to their integration with simulation approaches.

Big Data, given its potential allure, needs to be considered carefully. The domains of interest for population health simulations often represent multiple levels of analysis and, thus, multiple disciplinary theories and notions of processes. Historically, data collection in the social and behavioral sciences has been purposeful and largely driven by theory. Given Big Data, we must be careful to avoid the trap of deriving simulation modeling assumptions that are driven by new data sources without new and well-tested theory. A related issue is that the empirical grounding of any assumption about a process should be considered as conditional on other relevant processes, e.g., people's individual decisionmaking strategies are embedded in neighborhoods and social contexts that interface with policy and industry concerns dynamically over time. Big Data, given its unstructured format and high-temporal signature should excel in this respect.

As a final consideration, it is instructive to consider work in other disciplines that face similar hurdles when attempting to build confidence in model assumptions. For example, consider the Across Trophic Level System Simulation (ATLSS) family of agent-based models used to understand ecological issues in South Florida (e.g., the long-term effects on population dynamics of white-tailed deer and the Florida panther in relation to water management policy).[39] The degree to which many of the ATLSS assumptions are empirically grounded is unprecedented in terms of what we have accomplished in population health simulations, especially in relation to assumptions that govern the behavior and dynamics of model entities—e.g., data from radio-tagged panthers over time and space with high resolution to develop the agent-rules for habitat usage. We mention this contrast simply to generate discussion: Does population health need such temporal, spatial and process resolution? And, if yes, will Big Data provide for this need?

Integration into the Mainstream

The technical skills required for development and implementation of simulations are not typically found among researchers, educators, review panels, editors, and others in population health. Over time, there is potential for this to change if the next generation gains more exposure to simulations through pioneering programs like the National Institutes of Health (NIH) Institute on Systems Science and Health[b] and the newly developed core curriculum at the Columbia University Mailman School of Public Health.

The largest barrier for reaching this potential is cultural and normative—will we agree as a field that simulation should become a central approach in population health? And, will others

[b] See http://obssr.od.nih.gov/training_and_education/issh/.

(policymakers and the public) see population health as a discipline that can answer dynamic process questions? In order to remove this barrier, we need clarity in our collective expectations of what simulations actually do well, despite their limitations. Two issues are critically important.

First, the types of simulation models used in population health excel as qualitative tools for understanding alternative futures, situations, contexts, and importantly, unanticipated levers of change. These models generally are not designed for predicting precisely what will happen and to what degree. At first glance, the notion that computationally-explicit highly-quantitative models should really be used for qualitative understanding appears to be a glaring contradiction. However, this is simply not the case—it is well appreciated even in the most quantitative of sciences (e.g., physics) that many computational and mathematical models are meant to provide qualitative insights.

Second, models will contain a set of assumptions that are not empirically grounded—to include those that represent hypothesized but not yet fully understood processes. Mimetic social influence is a case in point, whereby we have evidence that it exists but the details of how it works with respect to other model assumptions are highly speculative in most simulation models. This limitation reflects the nature of simulation modeling. By definition, a simulation model must implement explicit assumptions about all hypothesized processes in the model. Otherwise, the simulation would not run, so to speak. When considering this issue, it is important to realize that one of the primary outcomes of simulation modeling is the uncovering of gaps in both theory and data.

John Sterman's central thesis on the use of simulation in public health is apropos: simulation provides a useful approach for learning in a complex world.[40] Some phenomena are complex and difficult to understand and thus call for the simulation approach. Learning is not gained from precise predictions and certainly does not require all assumptions to be known. It simply requires agreement as a field that we need new and sometimes more appropriate tools for understanding the core issues in population health.[41]

Acknowledgments

Collectively, the work presented above was funded in part by the following: Defense Threat Reduction Agency Comprehensive National Incident Management System Contract HDTRA1-11-D-0016-0001; National Institutes of Health Grants 5U01GM070694-11, 1R01GM109718, and 1R21HD067570-01; Grant No. 60466 from the Robert Wood Johnson Foundation; The Network on Inequalities, Complex Systems, and Health (HHSN276200800013C); Columbia University Mailman School of Public Health Epidemiology Merit Fellowship to Mark Orr; and an Institute for Integrative Health TIIH Scholar Award to George A. Kaplan. We thank Nathan Osgood, Ronald Mintz, Dylan Knowles, David Plaut, and members of the Social and Decision Analytics Lab and the Network Dynamics and Simulation Science Laboratory (NDSSL) at Virginia Tech for their suggestions and comments. The opinions presented herein are those of the authors and do not necessarily reflect the official position of the Agency for Healthcare Research and Quality, the National Institutes of Health, or the U.S. Department of Health and Human Services.

Authors' Affiliations

Mark G. Orr, PhD, is Associate Research Professor at Virginia Tech; Bryan Lewis, PhD, MPH, is an Assistant Research Professor at the Virginia Bioinformatics Institute at Virginia Tech; Kathryn Ziemer, PhD, is a Postdoctoral Associate at the Social and Decision Analytics Laboratory, Virginia Bioinfomatics Institute, Virginia Tech; and Sallie Keller, PhD, is Professor of Statistics and Director of the Social and Decision Analytics Laboratory, Virginia Bioinformatics Institute, Virginia Tech.

Address correspondence to: Mark Orr, Research Associate Professor, Social and Decision Analytics Laboratory, Virginia Tech-National Capital Region, 900 North Glebe Road, Arlington, VA 22203; email morr9@vbi.vt.edu

References

1. Senge PM. Statistical estimation of feedback models. Simulation 1977;28:177-84.
2. Morris M, Kretzschmar M. Concurrent partnerships and transmission dynamics in networks. Soc Net 1995;17:299-318.
3. Weinstein MC, Coxson PG, Williams LW, et al. Forecasting coronary heart disease incidence, mortality, and cost: the Coronary Heart Disease Policy Model. Am J Public Health 1987;77:1417-26.
4. Complex systems approaches to population health. May 2007 conference; University of Michigan. Available at http://sitemaker.umich.edu/complexsystemspopulationhealth/home. Accessed May 5, 2015.
5. Systems thinking (special issue). Am J Public Health 2006 Mar;96(3).
6. Annual Institute on Systems Science and Health. Bethesda, MD: Office of Behavioral and Social Sciences Research, National Institutes of Health. Available at http://obssr.od nih.gov/training_and_education/issh/. Accessed May 5, 2015.
7. Calhoun JG, Ramiah K, Weist EM, et al. Development of a core competency model for the master of public health degree. Am J Public Health 2008;98:1598-607.
8. Blower SM, Gershengorn HB, Grant RM. A tale of two futures: HIV and antiretroviral therapy in San Francisco. Science 2000;287:650-4.
9. Bahr DB, Browning RC, Wyatt HR, et al. Exploiting social networks to mitigate the obesity epidemic. Obesity 2009;17:723-8.
10. Levy DT, Pearson JL, Villanti AC, et al. Modeling the future effects of a menthol ban on smoking prevalence and smoking—attributable deaths in the United States. Am J Public Health 2011;101:1236-40.
11. Fishbein M, Ajzen I. Predicting and changing behavior: the reasoned action approach. New York, NY: Taylor and Francis, Psychology Press; 2010.
12. Janz N, Becker M. The health belief model: a decade later. Health Educ Q 1984;11:1-47.
13. Bandura A. Social foundations of thought and action: a social cognitive theory. Englewood Cliffs, NJ: Prentice Hall; 1986.
14. Weinstein N. Testing four competing theories of health—protective behavior. Health Psychol 1993;12:324-33.
15. Painter JE, Borba CPC, Hynes M, et al. The use of theory in health behavior research from 2000 to 2005: a systematic review. Ann Behav Med 2008;35:358-62.
16. Orr MG, Thrush R, Plaut DC. The Theory of Reasoned Action as parallel constraint satisfaction: towards a dynamic computational model of health behavior. Plos One 2013;8:e62409.
17. Gillmore MR, Archibald ME, Morrison DM, et al. Teen sexual behavior: applicability of the theory of reasoned action. J Marriage Fam 2002;64:885-97.
18. Rumelhart DE, McClelland JL, Group PR, eds. Parallel distributed processing: explorations in the microstructure of cognition. Vol 2: Psychological and biological models. Cambridge, MA: MIT Press; 1986.
19. McClelland JL, Rumelhart DE, Group PR, eds. Parallel distributed processing: Explorations in the microstructure of cognition. Vol 1: Foundations. Cambridge, MA: MIT Press; 1986.
20. Read SJ, Miller LC, eds. Connectionist models of social reasoning and social behavior. Mahwah, NJ: Lawrence Erlbaum Associates; 1998.
21. Orr MG, Plaut DC. Complex systems and health behavior change: insights from cognitive science. Am J Health Behav 2014;38:404-13.

22. Galea S, ed. Macrosocial determinants of population health. New York: Springer; 2007.
23. Orr MG, Galea S, Riddle M, et al. Reducing racial disparities in obesity: simulating the effects of improved education and social network influence on diet behavior. Ann Epidemiol 2014;24:563-9.
24. Ogden CL, Carroll MD, Kit BK, et al. Prevalence of obesity in the United States, 2009-2010. Hyattsville, MD: National Center for Health Statistics; 2012.
25. Finkelstein EA, Trogdon JG, Cohen JW, et al. Annual medical spending attributable to obesity: payer- and service-specific estimates. Health Aff 2009;28:w822-31.
26. Flegal KM, Carroll MD, Kit BK, et al. Prevalence of obesity and trends in the distribution of body mass index among US adults, 1999-2010. JAMA 2012;307:491-7.
27. Chao DL, Matrajt L, Basta NE, et al. Planning for the control of pandemic influenza A (H1N1) in Los Angeles County and the United States. Am J Epidemiol 2011;173:1121-30.
28. Eisenberg JNS, Scott JC, Porco T. Integrating disease control strategies: balancing water sanitation and hygiene interventions to reduce diarrheal disease burden. Am J Public Health 2007;97:846-52.
29. Goldstein E, Apolloni A, Lewis B, et al. Distribution of vaccine/antivirals and the 'least spread line' in a stratified population. J R Soc Interface 2010;7:755-64.
30. Halloran ME, Ferguson NM, Eubank S, et al. Modeling targeted layered containment of an influenza pandemic in the United States. Proc Natl Acad Sci 2008;105:4639-44.
31. Kumar S, Grefenstette JJ, Galloway D, et al. Policies to reduce influenza in the workplace: Impact assessments using an agent-based model. Am J Public Health 2013;103:1406-11.
32. Marathe A, Lewis B, Barrett C, et al. Comparing effectiveness of top-down and bottom-up strategies in containing influenza. PLoS One 2011;6:e25149.
33. Marathe A, Lewis B, Chen J, et al. Sensitivity of household transmission to household contact structure and size. PLoS One 2011;6:e22461.
34. Barrett CL, Bisset KR, Eubank SG, et al. EpiSimdemics: an efficient algorithm for simulating the spread of infectious disease over large realistic social networks. Proceedings of the 2008 ACM/IEEE Conference on Supercomputing; 2008: IEEE Press. p. 37.
35. Bisset KR, Chen J, Feng X, et al. EpiFast: a fast algorithm for large scale realistic epidemic simulations on distributed memory systems. Proceedings of the 23rd International Conference on Supercomputing; 2009: ACM. p. 430-9.
36. Eubank S, Guclu H, Kumar VS, et al. Modelling disease outbreaks in realistic urban social networks. Nature 2004;429:180-4.
37. Sterman JD. Business dynamics: systems thinking and modeling for a complex world. New York: McGraw-Hill; 2000.
38. Miller JH, Page SE. Complex adaptive systems: an introduction to complex models of social life. Princeton: Princeton University Press; 2007.
39. DeAngelis DL, Gross LJ, Huston MA, et al. Landscape modeling for Everglades ecosystem restoration. Ecosystems 1998;1:64-75.
40. Sterman JD. Learning from evidence in a complex world. Am J Public Health 2006;96:505-14.
41. Diez Roux A. Complex systems thinking and current impasses in health disparities research. Am J Public Health 2011;101:1627-34.

Mark G. Orr, PhD, is an Associate Research Professor at Virginia Tech. He was originally trained as a cognitive psychologist at the University of Illinois at Chicago. Dr. Orr received augmentation to this training with postdoctoral fellowships in computational modeling (Carnegie Mellon), neuroscience (Albert Einstein College of Medicine), and epidemiology/complex systems (Columbia University). Over the past decade, he has become heavily involved in understanding dynamic processes and drivers of risky behavior and decisionmaking, primarily in a public health context, at the individual and population levels.

Bryan Lewis, PhD, MPH, is an Assistant Research Professor at the Virginia Bioinformatics Institute at Virginia Tech. His work focuses on the use of highly-detailed computer simulations of human society to support public health decisionmaking and policies.

Kathryn Ziemer, PhD, is a Postdoctoral Associate at the Social and Decision Analytics Laboratory at Virginia Tech. Her research focuses on social cognitive theory, attitude formation and change, decisionmaking, and emotion regulation. Dr. Ziemer received her PhD in Counseling Psychology from the University of Maryland in 2014. Her dissertation involved creating a novel intervention using concepts from social cognitive theory and emotion regulation theory to improve the psychological and physical outcomes of individuals with chronic pain. More recently, she has focused on novel methods for measuring attitudes and the diffusion of attitudes across social networks.

Sallie Keller, Ph.D., is Professor of Statistics and Director of the Social and Decision Analytics Laboratory at the Virginia Bioinformatics Institute, Virginia Tech. Formerly, she was Professor of Statistics and Academic Vice-President and Provost at the University of Waterloo, Director of the IDA Science and Technology Policy Institute, and Professor of Statistics and the William and Stephanie Sick Dean of Engineering at Rice University. She has served as a member of the National Academy of Sciences Board on Mathematical Sciences and Their Applications, chaired the Committee on Applied and Theoretical Statistics, and currently is a member of the Committee on National Statistics. She is a national associate of the National Academy of Sciences, fellow of the American Association for the Advancement of Science, elected member of the International Statistics Institute, and member of the JASON advisory group. Dr. Keller is a fellow and past president of the American Statistical Association.

Understanding the Relationship Between Education and Health: A Review of the Evidence and an Examination of Community Perspectives

Emily B. Zimmerman, Steven H. Woolf, and Amber Haley

Abstract

Education is critical to social and economic development and has a profound impact on population health. We review evidence for the health benefits associated with education in the context of a socioecological model of health. The health benefits of education accrue at the individual level (e.g., skill development and access to resources); the community level (e.g., the health-related characteristics of the environments in which people live); and the larger social/cultural context (e.g., social policies, residential segregation, and unequal access to educational resources). All of these upstream factors may contribute to health outcomes, while factors such as ability to navigate the health care system, educational disparities in personal health behaviors, and exposure to chronic stress act as more proximate factors. It is also important to consider the impact of health on educational attainment and the conditions that occur throughout the life course that can impact both health and education, such as early childhood experiences. After exploring the literature linking health and education, we describe a project to engage residents of a low-income, urban community in a process of creating causal models to try to identify new links between education and health and help refine our understanding of the complex phenomena that shape this relationship. We asked community researchers to map out the pathways linking education and health in an effort to explore the possibility that people outside of academia might be able to help refine our understanding of complex phenomena by positing factors and relationships from their lived experience.

Introduction

It is now widely recognized that health outcomes are deeply influenced by a variety of social factors outside of health care. The dramatic differences in morbidity, mortality, and risk factors that researchers have documented within and between countries are patterned after classic social determinants of health, such as education and income,[1,2] as well as place-based characteristics of the physical and social environment in which people live and the macrostructural policies that shape them.

A 2013 report from the National Research Council and Institute of Medicine cited these socioecological factors, along with unhealthy behaviors and deficiencies in the health care system,

as leading explanations for the "health disadvantage" of the United States. In a comparison of 17 high-income countries, age-adjusted all-cause mortality rates for 2008 ranged from 378.0 per 100,000 in Australia to 504.9 in the United States. The report documented a pervasive pattern of health disadvantages across diverse categories of illness and injury that existed across age groups, sexes, racial and ethnic groups, and social class.[3]

Recent attention has focused on the substantial health disparities that exist within the United States, where life expectancy varies at the State level by 7.0 years for males and 6.7 years for females,[3] but mortality and life expectancy vary even more substantially across smaller geographic areas such as counties[4,5] and census tracts. In many U.S. cities, life expectancy can vary by as much as 25 years across neighborhoods.[6] The same dramatic geographic disparities can be seen for other outcomes, such as infant mortality, obesity, and the prevalence of diabetes and other chronic diseases.

Of the various social determinants of health that explain health disparities by geography or demographic characteristics (e.g., age, gender, race-ethnicity), the literature has always pointed prominently to education. Research based on decades of experience in the developing world has identified educational status (especially of the mother) as a major predictor of health outcomes, and economic trends in the industrialized world have intensified the relationship between education and health. In the United States, the gradient in health outcomes by educational attainment has steepened over the last four decades[7,8] in all regions of the United States,[9] producing a larger gap in health status between Americans with high and low education. Among white Americans without a high school diploma, especially women,[10] life expectancy has decreased since the 1990s, whereas it has increased for others.[8] Death rates are declining among the most educated Americans, accompanied by steady or increasing death rates among the least educated.[11] The statistics comparing the health of Americans based on education are striking:

- At age 25, U.S. adults without a high school diploma can expect to die 9 years sooner than college graduates.[12]
- According to one study, college graduates with only a Bachelor's degree were 26 percent more likely to die during a 5-year study followup period than those with a professional degree. Americans with less than a high school education were almost twice as likely to die in the next 5 years compared to those with a professional degree.[13]
- Among whites with less than 12 years of education, life expectancy at age 25 fell by more than 3 years for men and by more than 5 years for women between 1990 and 2008.[8]
- By 2011, the prevalence of diabetes had reached 15 percent for adults without a high school education, compared with 7 percent for college graduates.[14]

What accounts for the growing health disadvantages that exist among people with lower educational attainment? Is it what they learn in school, such as how to live a healthy lifestyle, or the socioeconomic advantages that come from an education? Or is the cross-sectional association between education and health more complex, involving nuanced contextual covariables in our society that provide a fuller back story?

This chapter explores the relationship between education and health from the perspective of the peer-reviewed literature and that of community members, engaged through a research exercise, to

blend insights from lived experience with the empirical data accumulated from scholarly research. Unpacking the reasons for the connection between education and health is not just an exercise in scientific inquiry, it is also essential to setting policy priorities. As increasing attention is focused on the need to address social inequity in order to address health inequities, understanding the links between broad upstream factors such as education and health outcomes becomes a critical challenge. Awareness of the importance of education might help drive investment in education and improvements in education and educational policy.

Conceptual Framework

An overarching theoretical framework for the impact of social determinants on health is provided by an ecological model in which individuals and their behavior are embedded, across the lifespan, within a framework of nested institutional contexts (Figure 1).[15] The individual and his or her characteristics are situated within and affected by the family and household, the community and its institutions (e.g., school, workplace, civil institutions), and policies of the larger society. Each level brings access to opportunities, as well as constraints on actions and opportunities. Furthermore, these levels interact with one another, such that family resources, for example, may mediate or moderate the resources available within the community. Social scientists widely agree that unequal social status creates unequal access to resources and rewards. Social structure, as embodied in social position, structures individual behaviors and values and therefore affects many of the mediators in the relationship between education and health.

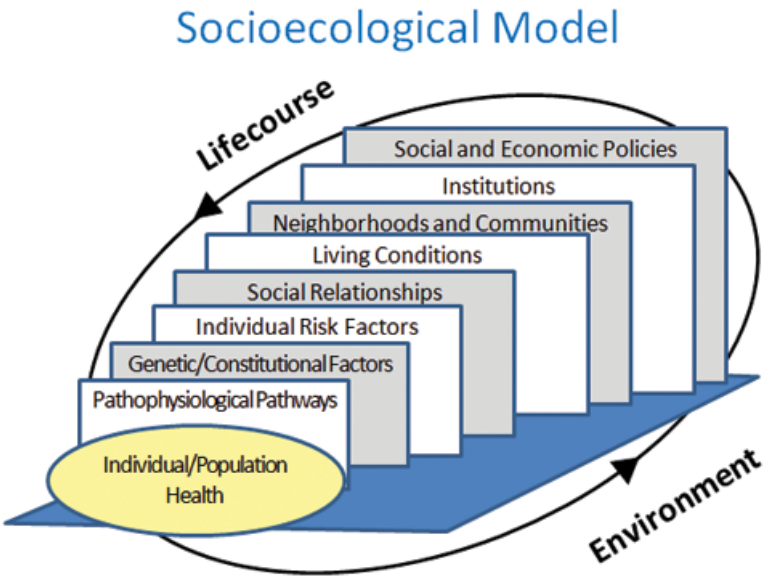


Figure 1. The Socioecological Model

Source: Kaplan GA, Everson SA, Lynch JW. The contribution of social and behavioral research to an understanding of the distribution of disease: a multilevel approach. In Smedley BD, Syme SL (eds), Promoting health: intervention strategies from social and behavioral research. Washington, DC: National Academies Press; 2000. Used with permission.

Note: Figure depicts a multilevel approach to epidemiology.

Education is one of the key filtering mechanisms that situate individuals within particular ecological contexts. Education is a driving force at each ecological level, from our choice of partner to our social position in the status hierarchy. The ecological model can therefore provide a context for the numerous ways in which education is linked to our life experiences, including health outcomes. It also provides a framework for understanding the ways in which educational outcomes themselves are conditioned on the many social and environmental contexts in which we live and how these, in turn, interact with our individual endowments and experiences.

Within this rich contextual framework, educational attainment (the number of years of schooling completed) is important but is far from the whole story. Educational attainment is often a key indicator in research studies, not least because it is often measured and recorded; life expectancy is compared by educational attainment because it is the only information about education recorded on death certificates. Besides obvious measures of the quality of education such as proficiency scores and understanding of mathematics, reading, science, and other core content, other dimensions of education are clearly important in the ecological context as well; cognitive development, character development, knowledge, critical thinking, and problem solving are a few examples.

Additionally, the relationship between years of education and health is not a purely linear function. As part of a literature attempting to clarify the functional form of the relationship between education and health, Montez et al. have documented a negative relationship between years of education and mortality risk for attainment less than high school graduation, a steep decline for high school graduates (with reduction of risk five times greater than attributable to other years of education), and a continued yet steeper negative relationship for additional years of schooling (Figure 2).[16] The drop at high school graduation points to the importance of obtaining credentials in addition to other benefits of educational attainment.

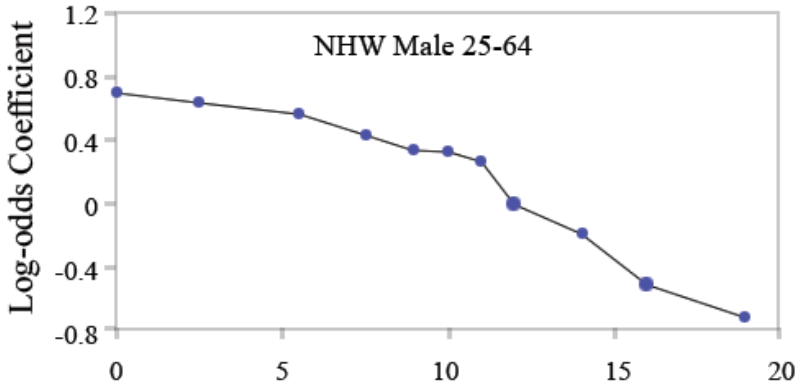


Figure 2. Log-odds coefficients for semi-nonparametric levels of educational attainment (functional form 1) by race-gender-age

Source: Montez JK, Hummer RA, Hayward MD. Educational attainment and adult mortality in the United States: a systematic assessment of functional form. Demography 2012;45:315–36. Used with permission.

In order to present a nuanced picture of the relationship between education and health, this chapter is presented in two parts. First, we review the health benefits associated with education, focusing on the primary mechanisms, both distal and proximate, by which education may be considered a driving force in health outcomes. We take a socioecological approach by presenting these concepts in a hierarchy, moving from the level of the person to the community/institution and then the larger social/policy context. Next, we turn to issues of causality that can make it difficult to draw conclusions about the relationship between education and health. These include reverse causality and selection, in which education may actually be impacted by ill health, and confounding, where both education and health are affected by some other causal factor(s) that may also provide important clues about the root causes of poor education and poor health.

Finally, this chapter moves beyond abstract academic models to discuss alternate ways of understanding and prioritizing these mechanisms. We look at preliminary results from a project to garner a "view from the inner city" based on the lived experiences of residents of a disadvantaged neighborhood and how their insights may highlight, broaden, or reinterpret our understanding of the mechanisms presented earlier in the chapter. Our goal is not to settle the question of which are the most important mechanisms by which education and health are related, but rather to call attention to the value of engaging people within communities in enabling researchers and policymakers to better understand and operationalize the importance of education in everyday life and the meaning of empirical evidence from the literature. Our work is part of a larger trend in community-based participatory research (CBPR) that is invigorating a dialogue that incorporates community engagement into the important discussions surrounding social and health inequalities.[17]

Readers are cautioned that this chapter touches on a diverse spectrum of factors—all linked to education—that vary from urban design to psychosocial characteristics, access to health care, air pollution, and economic policy. These very diverse domains are each the subject of large literatures that cannot be systematically catalogued in this space. Rather than offering a systematic review, our goal is to draw attention to these factors as part of the education-health relationship and to cite representative sources where readers can explore these topics in more detail; we encourage this research because the quality of evidence linking these factors to health outcomes is uneven and in some cases speculative. Education is linked to established health determinants supported by extensive evidence, such as tobacco use and poverty, but also to factors with less developed evidence, such as allostatic load and social cohesion. Research on methods for improving educational outcomes and learning is not catalogued here due to space constraints but is of vital importance. Finally, the individual elements of the socioecological model exist in a context, and disciplinary and transdisciplinary research is highly relevant in understanding the interplay of contextual factors in a complex systems relationship.[18,19]

Health Benefits Associated with Education

Among the most obvious explanations for the association between education and health is that education itself produces benefits that later predispose the recipient to better health outcomes. We may think of these returns from education, such as higher earnings, as subsequent "downstream"

benefits of education (later in the chapter we will discuss "upstream" factors that may influence both education and health throughout the life course, especially before children ever reach school age). Following the socioecological framework presented in the introduction, we describe a range of potential downstream impacts of education on health, starting with the ways individuals experience health benefits from education, but then going on to discuss the health-related community (or place-based) characteristics that often surround people with high or low education, and closing with the larger role of social context and public policy.

Impact at the Individual Level

Education can impart a variety of benefits that improve the health trajectory of the recipient. Below we discuss its role in enhancing non-cognitive and cognitive skills and access to economic resources, and we highlight the impacts of these on health behaviors and health care usage. Although this section focuses specifically on the health benefits of education, we do so in full knowledge that education is impacted by health, development, and a host of personal, community, and contextual factors.

Education Impacts a Range of Skills

Education contributes to human capital by developing a range of skills and traits, such as cognitive skills, problem solving ability, learned effectiveness, and personal control.[20] These various forms of human capital may all mediate the relationship between education and health. Personality traits (also known as "soft" or non-cognitive skills) are associated with success in later life, including employment and health. The 'Big Five' personality factors include conscientiousness, openness to experience, extraversion, agreeableness, and neuroticism/emotional stability.[21] Roberts et al. postulate three pathways whereby personality traits may impact mortality: through disease processes (e.g., response to stress), health-related behaviors, and reactions to illness. They suggest that the strength of association between the 'Big Five' personality traits and mortality is comparable to that of IQ and stronger than socioeconomic status.[22] Although enduring, these skills are also mutable, and research indicates that educational interventions to strengthen these skills can be important, especially among children in disadvantaged areas, who may find it more difficult to refine these skills at home and in their social environments.

Personal control, also described in the literature in terms of locus of control, personal efficacy, personal autonomy, self-directedness, mastery, and instrumentalism,[23] is another soft skill associated with educational attainment. According to Ross and Wu (p. 723), "Because education develops one's ability to gather and interpret information and to solve problems on many levels, it increases one's potential to control events and outcomes in life. Moreover, through education one encounters and solves problems that are progressively more difficult, complex, and subtle, which builds problem-solving skills and confidence in the ability to solve problems."[23]

Personal control can impact individuals' attitudes and behaviors, potentially including health behaviors. Furthermore, an individual's sense of mastery and control may mediate stress, possibly by facilitating better coping mechanisms. Lack of personal control, on the other hand, may provoke physiological responses, leading to suppression of the immune system.[23]

Box 1. Impact of Education on the Ability to Navigate Health Care

Achieving positive health outcomes in today's health care environment requires a variety of factors to come together that may be affected by educational attainment and a combination of soft and hard skills. Patients benefit from the ability to understand their health needs, follow or read instructions, advocate for themselves and their families, and communicate effectively with health providers. A systematic review of health literacy and health outcomes found that individuals with lower health literacy had poorer health-related knowledge and comprehension,ability to demonstrate taking medications properly, and ability to interpret medication labels and health messages. They also had increased hospitalizations and emergency care, decreased preventive care, and, among the elderly, poorer overall health status and higher mortality.[26] For example, low literacy and low levels of other basic skills such as listening and numeracy have been associated with greater difficulty in asthma care in adults.

In a review of the impact of patient socioeconomic status on patient-physician communication, Willems et al.[28] concluded that communication is influenced in part by patients' communicative ability and style, which depend largely on education and other personal attributes. Education contributes to more active communication, such as expressiveness and asking questions. In response, physicians tend to communicate less to patients who seem less educated and to provide care that is more directive and less participatory.

In addition to its impact on soft skills, education has the potential to impart skills in reading, mathematics, and science/health literacy that could contribute to an individual's health. Learners of English as a second language are helped to overcome language barriers that can interfere with understanding of health needs. Education may also improve a range of other skills, such as cognitive ability, literacy, reaction time, and problem solving. Pathways from these skills to health outcomes may be indirect, via attainment of better socioeconomic circumstances or behavior, but they may also apply directly in clarifying the increasingly complex choices individuals face in understanding health priorities and medical care needs. Skills such as higher cognitive ability and health literacy may also lead directly to improved health outcomes because of an enhanced "ability to comprehend and execute complex treatment regimens," and better disease self-management.[24] A strong education may be important in both navigating health care (see Box 1) and making choices about lifestyle and personal health behaviors (see Box 2). Cutler and Lleras-Muney report that increased cognitive ability resulting from education contributes significantly to the education gradient in health behaviors.[25]

Education Increases Economic and Social Resources

A large part of the impact of education on health flows through the attainment of economic resources, such as earnings and wealth, as well social resources such as access to social networks and support.[36] Adults with more education are less likely to experience unemployment and economic hardship and will have greater access to a variety of important material, financial, and social resources (see Box 3). Link and Phelan (p. 87) point out that the specific mechanisms linking socioeconomic status (SES) to health have changed over time but that SES remains a fundamental social cause of disease because it involves "access to resources that can be used to avoid risks or to minimize the consequences of disease once it occurs."[1]

Box 2. Impact of Education on Personal Health Behaviors

Adults with higher levels of education are less likely to engage in risky behaviors, such as smoking and drinking, and are more likely to have healthy behaviors related to diet and exercise. Data from the National Survey on Drug Use and Health (NSDUH) indicate that in 2009-10, 35 percent of adults who did not graduate high school were smokers, compared to 30 percent of high school graduates and 13 percent of college graduates.[29] The impact of education on health behaviors likely stems from education's impact on skills as well as socioeconomic status. Examining competing explanations for the education gradient in health behaviors, Cutler and Lleras-Muney find evidence for the importance of resources, cognitive ability (especially how one processes information), and social integration.[25]

Education offers opportunities to learn more about health and health risks, both in the form of health education in the school curriculum and also by giving individuals the health literacy to draw on, later in life, and absorb messages about important lifestyle choices to prevent or manage diseases. For example, people with more education are more likely to have healthy diets and exercise regularly. Analysis of several waves of data from the National Health and Nutrition Examination Survey (NHANES) found that intake of specific nutrients (e.g., vitamins A and C, potassium, calcium), as well as overall diet quality, are associated with education.[30] In addition, Behavioral Risk Factor Surveillance System (BRFSS) data for 2010 indicate that only 61 percent of adults with less than a high school education and 68 percent of high school graduates said that they exercised in the past 30 days, compared to 85 percent of college graduates.[31] It must be noted, however, that not all behavioral risk factors are higher among those with the lowest educational attainment. BRFSS data for 2011 indicate that the prevalence of binge drinking increases with higher levels of education.[32]

Finally, adults with higher levels of education tend to have lower exposure to stress related to economic deprivation or relative deprivation,[33] and may therefore be less inclined than those with lower levels of education to adopt unhealthy coping behaviors for stress. Individuals with more education tend to have greater socioeconomic resources for a healthy lifestyle and a greater relative ability to live and work in environments with the resources and built designs for healthy living.[34,35]

Economic Resources

Adults with a higher education—especially in today's knowledge economy—have conspicuous advantages in gaining employment and finding desirable jobs (Figure 3). Advanced degrees give workers an advantage in obtaining rewarding jobs that offer not only higher salaries and job satisfaction but other health-related benefits such as health insurance coverage. For example, adults with health insurance in the United States use more physician services and have better health outcomes compared to uninsured or inconsistently insured adults.[37-39] Worksite health promotion programs and policies that protect occupational safety also play a role. An inadequate education markedly increases the risk of unemployment. In 2012, unemployment was 12.4 percent among adults who did not graduate high school, compared to 8.3 percent among adults with a high school diploma and 4.5 percent among college graduates.[40] A body of evidence links unemployment to adverse health outcomes. For example, a higher percentage of employed persons reported in 2010 that they were in excellent or very good health (62.7 percent) than did individuals who were unemployed for less than 1 year (49.2 percent) or unemployed for more than 1 year (39.7 percent). The unemployed also reported more physically and mentally unhealthy days in the past 30 days.[41]

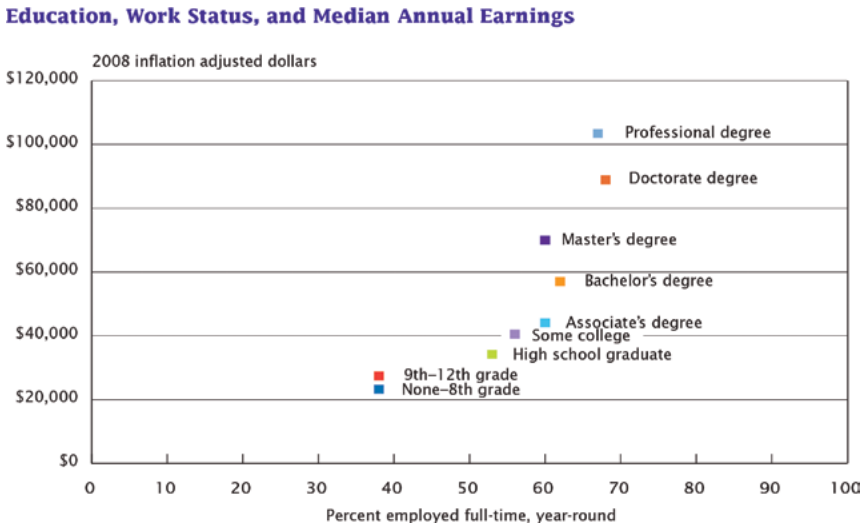


Figure 3. Education, work status, and median annual earnings

Source: Education and Synthetic Work-Life Earnings Estimates. American Community Survey Reports, United States Census Bureau. September 2011. Used with permission.

The income and wealth that come from a good education are leading predictors of health status,[2,42] and accumulated financial strain has been shown to impact health above and beyond the effects of income and wealth.[43] In today's society, economic resources are inextricably linked to education. In 2012, the median wage for college graduates was more than twice that of high school dropouts and more than one and a half times that of high school graduates.[40] Weekly earnings are dramatically higher for Americans with a college or advanced degree. A higher education has an even greater effect on lifetime earnings (Figure 4), a pattern that is true for men and women, for blacks and whites, and for Hispanics and non-Hispanics. According to 2006-2008 data, the lifetime earnings of a Hispanic male are $870,275 for those with less than a 9th grade education but $2,777,200 for those with a doctoral degree. The corresponding lifetime earnings for a non-Hispanic white male are $1,056,523 and $3,403,123.[44]

The economic vulnerability that can arise from an inadequate education can affect health through a cascade effect on the ability to acquire resources that are important to health (e.g., food, stable housing, transportation, insurance, and health care).[45] People with low income are more likely to be uninsured and to be vulnerable to the rising costs of health care, which insurance carriers are increasingly shifting to patients through higher copayments, deductibles, and premiums. In 2012, one-fourth (24.9 percent) of people in households with an annual income less than $25,000 had no health insurance coverage, compared to 21.4 percent of people in households with incomes ranging from $25,000 to $49,999; 15.0 percent in households with income ranging from $50,000 to $74,999; and 7.9 percent with incomes of $75,000 or more.[46]

Figure 4. Median synthetic work-life earnings by education, race/ethnicity, and gender: full-time, year-round workers

Source: Reprinted with permission of the Center on Society and Health, Virginia Commonwealth University.

Individuals with the higher incomes that accompany education have more resources to purchase healthy foods, afford the time and expenses associated with regular physical activity, have easy transportation to health care facilities or work locations, and afford health care expenses. According to 2010 BRFSS data, 27 percent of adults with less than a high school education reported not being able to see a physician due to cost, compared to 18 percent and 8 percent of high school and college graduates, respectively.[47] Accordingly, the costs of a healthy lifestyle pose more of a barrier for people with less education. The health implications of these financial barriers to health care are well documented: the uninsured are less likely to receive preventive care or help with disease management,[48] and they have a higher risk of mortality.[49]

Box 3. Stress and Allostatic Load

Allostatic load results in an individual's inability to adapt to long-term stress, leading to chronic illness. Individuals with lower levels of educational attainment are at greater risk of exposure to stress, such as chronic occupational stress or unemployment, and they may be less likely to have buffers that reduce the impact of stress (e.g., social support, sense of control or mastery over life, and high self-esteem).[55] Effects of stressors vary depending on factors such as genetic makeup, development, early experiences, the availability of coping mechanisms, and responses to threats.[56]

A growing body of research is documenting that life changes, traumas, chronic strain, and discrimination—all of which can accompany an inadequate education—can be harmful to both physical and psychological health. Chronic stressors can be related to a wide variety of circumstances, such as social roles, interpersonal conflict, and the environment or living conditions. Stressful events may interact with the experience of chronic stress to affect outcomes, and these stressors are, in turn, influenced by one's personal traits and values and mediated by factors such as coping mechanisms and social support.[57] For those confronting life without a good education, individual stressors can accumulate over time and may, in turn, heighten exposure to further stressors.

The biological consequences of stress and allostatic load are increasingly clear, as are their effect on cognition. For example, a longitudinal study of high functioning older adults found associations between baseline measures of allostatic load and cognitive function, physical performance, and the incidence of cardiovascular disease during the study period.[58] A 4.5-year followup study of the same subjects found increased risk of mortality among individuals with higher baseline allostatic load scores as well as among those whose score increased.[59] The combination of high perceived stress and risky health behaviors has been found to be associated with increased mortality among individuals of low socioeconomic status.[60]

Social Resources

Educational attainment is associated with greater social support, including social networks that provide financial, psychological, and emotional support. Social support includes networks of communication and reciprocity. Individuals in a social network can relay information, define norms for behavior, and act as modeling agents. Those with higher levels of education may also have higher levels of involvement with civic groups and organizations. Conversely, low social support (i.e., not participating in organizations, having few friends, being unmarried, or having lower quality relationships) is associated with higher mortality rates and poor mental health.[50,51] The social integration that often accompanies education has been linked to health outcomes in a causal chain that begins with the macro-social and ends with psychobiological processes.[52] Berkman et al.[52] propose several mechanisms through which social integration affects health: social support, social influence, social engagement/attachment, and access to goods and resources. Social connection can be an

important buffer to the negative health consequences of health stressors. Marriage imparts benefits in longevity, but weaker network ties can also have important health effects, such as the effects of peers on behavior.[53] The effect of social networks on smoking cessation is a well-known example.[54]

Impact at the Community Level

Individuals with education benefit not only from the resources that schooling brings to them and their families but also from health-related characteristics of the environments in which they tend to live, work, and study. Although there are many methodological challenges in estimating community-level effects on individuals,[61,62] communities appear to confer a range of benefits or risks that can impact health. In the midst of growing recognition that "place matters" to health, many studies have tried to estimate neighborhood effects on outcomes such as child/youth educational attainment, behavioral/well-being outcomes, or health status and mortality. For example, Ross and Mirowsky[63] used multilevel analysis of survey data from Illinois to address the question of whether community SES impacts health above and beyond the contributions of individual SES. They found that individual-level indicators of SES explained most of the variation in physical functioning (about 60 percent), but that neighborhood-level measures had a significant influence as well. Given the wide range of methodologies and data sources utilized, findings are not uniform among such studies, but there is general agreement that a relatively modest neighborhood effect exists independent of individual and family-level factors such as education or income.[61,64,65] Effects that appear to occur at the neighborhood level may represent aggregated individual characteristics (compositional effects), neighborhood variability (contextual effects), or local manifestations of larger scale processes (e.g., higher-level planning or regulatory decisions).[66] Furthermore, it is important to recognize the dynamic interaction that occurs between the individual and the environment[67] and conceptions of space as "relational geographies."[68]

At one level, community characteristics matter because access to resources that are important to health is contingent on community-level resources and institutions. Macintyre and Ellaway categorize these as physical features, services, sociocultural features, reputation, and availability of healthy environments at home, work, and play.[69] Theories about the mechanisms by which social environments affect the health of individuals also focus on community characteristics such as social disorganization, social control, social capital, and collective efficacy.[70] Kawachi et al. note that communities with higher social capital tend to be more resilient in the face of disasters and are better able to employ informal control mechanisms to prevent crime.[71]

People with low education tend to live in certain communities that, through a combination of resources and characteristics, expose individuals to varying levels of risk versus safety (e.g., crime, unemployment, poverty, and exposure to physical hazards) and provide different levels of resources (e.g., food supply, green space, economic resources, and health care). One notable resource that differs among communities is the quality of education itself. Low-income neighborhoods often have fewer good schools, not least because public schools tend to be poorly resourced by low property taxes and cannot offer attractive teacher salaries or properly maintain buildings, supplies, and school safety. Adverse community factors can compound the difficulty that children face in obtaining a good education while also compromising their health trajectory.

Below we touch on several additional community characteristics that have been linked to health outcomes and tend to vary with the level of education of the population. These characteristics include

food access, spaces and facilities for physical activity, access to health care, community economic resources, crime and violence, and environmental exposure to toxins.

Food Access

Unhealthy eating habits are linked to numerous acute and chronic health problems such as diabetes, hypertension, obesity, heart disease, and stroke, as well as higher mortality rates, but access to healthier foods tends to be limited in neighborhoods with lower median incomes and lower levels of educational attainment. In one study, access to healthier food outlets (defined as at least one healthier food retailer within the census tract or within 1.5 miles of tract boundaries) was 1.4 times less likely in census tracts with fewer college-educated adults (less than 27 percent of the population) as in tracts with a higher proportion of college-educated persons; these differences varied by region and were highest in the South and lowest in the West and Northeast.[72] Conversely, low-SES neighborhoods often have an oversupply of fast food restaurants, convenience stores, bodegas, liquor stores, and other outlets that sell little fresh produce but promote inexpensive calorie-dense foods and unhealthy beverages.

Spaces and Facilities for Physical Activity

People with higher education and income are more likely to live in neighborhoods that provide green space (e.g., parks), sidewalks, and other places to enable residents to walk and cycle to work and shopping, exercise, and play outside. Lower-income neighborhoods and those with higher proportions of non-white residents are also less likely to have commercial exercise facilities.[73] The health benefits of green space have been documented in urban environments, especially for lower income, young, and elderly populations.[74] A longitudinal study in Great Britain found immediate, positive mental health effects of moving to urban areas with more green space.[75]

Access to Health Care

Because of the maldistribution of health care providers in the United States,[76] access to clinicians and facilities tends to be in shortest supply in the rural and low-income areas populated by people with limited education. Thus, apart from whether residents have the health insurance coverage and resources to afford health care, they may struggle to find primary care providers, specialists, and hospitals in their area that provide quality health care services.

Community Economic Resources

The lack of jobs in low-income communities can exacerbate the economic hardship that is common for people with less education. Such individuals are more likely to live in communities with a weak economic base that is unattractive to businesses, employers, and investors and are thereby often caught in a self-perpetuating cycle of economic decline and marginalization.

Crime and Violence

Elevated crime rates in neighborhoods populated by people with low education can impact health through the direct effects of violent crimes on victims, such as trauma and high youth mortality rates. Crime can also affect health indirectly, such as through fear of crime or the cumulative stress of living in unsafe neighborhoods.[77] The high incarceration rates of residents in some low-SES communities can have deleterious effects on social networks, social capital, and social control, further compromising public health and safety.[78] The 2006 and 2007 rounds of the American Community Survey found that,

among young male high school drop-outs, nearly 1 in 10 was institutionalized on a given day in 2006-2007 versus less than 1 of 33 high school graduates.[79]

Environmental Exposure to Toxins

People of color and those with less education are more likely to live in neighborhoods that are near highways, factories, bus depots, power plants, and other sources of air and water pollution. A large body of research on environmental justice has documented the disparate exposure of low-income and minority neighborhoods to hazardous waste, pesticides, and industrial chemicals.[80] This exposure to toxins is perhaps the most undiscriminating place-based characteristic because residents' personal socioeconomic advantages (e.g., education, income) offer no protection against the adverse health consequences of inhalation or ingestion of such toxins.

The Larger Social Context and Social Policy

Health inequities are driven, in large part, by the social context in which people are born, live, and work—that is, the social policies that shape resources, institutions, and laws; the economic system through which material and financial resources are created and distributed; and the social norms that govern interactions. The conditions in which people live—for example, the built environment, public transportation, urban design, crime rates, food deserts, and the location of polluting factories—are determined by macrostructural policies and the cultural values that shape them. Formulation of effective analyses and solutions to problems affecting health must address factors that go beyond the level of the individual and proximal risk factors.[81] These influences have been recognized by organizations concerned with health outcomes locally, nationally, and internationally. The World Health Organization calls for improved living and working conditions, social protection policy supportive of all, reduced inequality, and strengthened governance and civil society.[2] Healthy People 2020 has many policy objectives for health, including improved environmental conditions (e.g., air/water quality and exposure to hazards), violence prevention, poverty reduction, and increased rates of postsecondary education.[82] The Place Matters team in Alameda County, CA has identified five policy areas to impact health outcomes locally: economics, education, criminal justice, housing and land use, and transportation.[83]

Decisions made by society, voters, and policymakers—both within and outside of government—exert deep influences on education itself, as well as on the institutions and resources that populate the socioecological framework linking education and health. For example, in other societies, the adverse health consequences of poverty are often buffered by social services that act to safeguard the health of children, young parents, and other vulnerable groups. Bradley et al. found that while most high-income countries spent more on social services than on health expenditures, the converse was true in the United States. The average ratio of social to health expenditures in OECD countries from 1995 to 2005 was 2.0; the ratio in the United States was 0.91.[84]

Economic policies have a large influence on the employment and wealth-building opportunities of workers and the marketability of an education. Major economic and technological shifts of the last few decades have favored "non-tradable" service jobs in sectors such as government and health care, while manufacturing jobs have moved to less developed countries in large numbers. Remaining jobs in the "tradable" sectors such as technology and finance increasingly require advanced skill sets.[85] These employment trends provide a critical context in the relationship between education and health—those unable to acquire the necessary education to be competitive in an increasingly restrictive job environment are vulnerable to long-term economic hardship.

Educational opportunities, however, are not equally distributed in the United States. Public school funding, largely dependent on local property taxes, varies widely both within and between States. The best funded school systems in the United States have per pupil expenditures almost four times the per pupil expenditures in the lowest spending schools.[86] Although early studies failed to find a strong relationship between school funding amounts and student achievement, some meta-analyses have supported the link between school funding and individual achievement.[87]

Inequalities by education cannot be disentangled from the backdrop of inequalities by gender, race, ethnicity, sexual orientation, and disability and their effects on both risks and opportunities. Figure 5 shows persistent gender and race disparities in earnings.[88] There are cultural as well as material dimensions of inequality (see Box 4), as when cultural status beliefs influence inequality primarily at the social relational level by shaping people's expectations for themselves and others.[89] Societies that impose social status hierarchies based on "categories" of difference solidify and perpetuate differentials in power and control of resources—thus leading to material inequalities. Income inequalities in the United States are significant and have become more pronounced, with wages at the lower or middle of the income distribution stagnating or falling while those at the top continue to rise. Income inequality persisted during the recovery from the Great Recession, during the first 3 years of which 95 percent of income gains accrued to the top 1 percent of earners.[90] The Gini coefficient, which measures income inequality, rose from 0.394 in 1970 to 0.469 in 2010; the share of household income earned by the bottom quintile was 3.3 percent in 2010, compared to 50.2 percent among the top quintile.[91]

Figure 5. Women's earnings as a percentage of men's median usual weekly earnings (full-time wage and salary workers) in current dollars, by race and ethnicity (1980-2010 annual averages)

Source: Women's earnings as a percent of men's in 2010. Bureau of Labor Statistics, U.S. Department of Labor, The Economics Daily, January 10, 2012.

Note: Data shown are from the Current Population Survey (CPS). To learn more, see Women in the Labor Force: A Databook (2011 Edition), BLS Report 1034, December 2011.

The continuing racial residential segregation and increasing economic segregation of urban landscapes affect the life chances of those living in concentrated poverty "irrespective of personal traits, individual motivations, or private achievements" and expose residents, many of whom lack adequate education, to higher levels of social problems.[92] These historical, economic, and cultural factors have also shaped and reinforced the racial division of labor and adverse impact on the low-wage sector.[93]

Box 4. Impact of the Cultural Context on Health Disparities and the Use of Health Care

Cultural influences can be important features of the causal web linking education and health. Experience with discrimination and racism (e.g., perceived discrimination, segregation, institutional discrimination, reduced access to goods and services), which may occur more commonly among people with less education, has a known relationship to stress and stress-related health disparities, as well as to health care seeking, treatment adherence, and risky health behaviors.[94] Mistrust among patients and bias among health care providers can affect the quality of care.[95] For example, a study of 202 African American patients with HIV in a primary care setting found that patients with higher educational attainment reported higher levels of trust, better communication with providers, and higher levels of shared treatment decisions. It also found that health outcomes were related to the belief that the health care provider should integrate culture in HIV treatment and to the perceived quality of provider communication. Trust was related to medical self-care but not to other outcomes.[96] Care is also affected by the cultural competency of providers—that is, their ability to recognize and appropriately respond to key cultural features that affect health care, which may include language, cultural values, patient beliefs, folk illnesses, and provider practices.[97]

Reverse Causality and Selection

The association between education and health may reflect not only the health benefits of education but a selection phenomenon caused by the detrimental effects of illness on educational success. Basch identifies five causal pathways by which health may impact motivation and ability to learn—sensory perceptions, cognition, school connectedness and engagement, absenteeism, and temporary or permanent dropping out.[98] For example, chronic health conditions can impact children's development and educational performance.[99] Such children are more likely to have absences for medical reasons and to be distracted by health concerns. Nonetheless, research evidence demonstrating that poor health has a causal relationship with educational outcomes is incomplete,[100] and findings of the overall effects range from about 1.4 years reduced educational attainment[101] to about half a year,[7] but there are notable exceptions. For example, evidence across countries and time periods demonstrates the harmful effect of low birth weight on education.[100,102] Disease, malnutrition, and prenatal and childhood exposures to toxins can also impact physical and cognitive development and educational achievement.[103]

The extent to which reverse causality contributes to the association between education and health requires further study, but longitudinal data—the most compelling evidence to resolve the controversy—tend to suggest that most of the association is attributable to the downstream benefits of education. Eide and Showalter[102] reviewed studies incorporating a range of methodologies that attempted to examine causal links between education and health outcomes. Studies of natural experiments in the United States (e.g., changes in compulsory school laws) generally found evidence of a causal link with mortality. Twin studies found evidence for causal links between years of schooling and self-reported health, the probability of being overweight (among men but not women),

and the effects of college attendance on preventive health care later in life.[102] Link and Phelan also discussed research attempting to show the direction of causality using quasi-experimental approaches, longitudinal designs, and analyses of risk factors that cannot be attributed to individual illness (e.g., plant closings). They concluded that these studies "demonstrated a substantial causal role for social conditions as causes of illness."[1]

Conditions Throughout the Life Course that Affect Both Health and Education

A third way that education can be linked to health is when education acts as a proxy for factors throughout the life course—most notably in early childhood—that affect both education and health. For example, as noted earlier, the social and economic environment facing individuals and households and the stresses and allostatic load induced by material deprivation can affect success in school (and work) while also inducing biological changes and unhealthy behaviors that can increase the risk of disease. Although this can occur throughout the life course, increasing attention is being placed on the role of these factors on children before they ever reach school age.

Early Childhood Experiences

The education community has long understood the connections between early life experiences and educational success. It is well-established that school readiness is enhanced by positive early childhood conditions—for example, fetal well-being and social-emotional development,[104] family socioeconomic status,[a,100,105,106] neighborhood socioeconomic status,[107,108] and early childhood education[109]—but some of these same exposures also appear to be vital to the health and development of children and their future risk of adopting unhealthy behaviors and initiating adult disease processes.

Below are several examples from the literature of early childhood experiences that influence health:

- Low birth weight affects not only educational outcomes but also health and disability.[110]
- Nurturing relationships beginning at birth, the quality of the home environment, and access to stimulation provide a necessary foundation for children to grow and thrive.[111] One example of this is the importance of child-directed speech during infancy for developing language skills.[112] The effects of stress can be reduced when children have a responsive and supportive caregiver available to help them cope with stress and provide a protective effect.[113]
- Unstable home and community life, such as economic factors, family transitions, housing instability, and school settings, can harm child development and later outcomes spanning education and health.[114] In one study,[b] homelessness and struggles with mortgage payments and foreclosure were predictive of self-rated health, and these combined with other categories (e.g., moved for cost in past 3 years, behind on rent) also predicted mental health problems.[115]
- Family and neighborhood socioeconomic status not only affect education but also predict developmental and health trajectories as children grow and develop.[116,117] The duration and timing of childhood poverty are important. Longitudinal studies indicate that the largest effects of

[a] Children's birth weight, developmental outcomes, health status (e.g., obesity and specific health conditions), disability, and success in school are strongly linked to parents' education and family income and assets.

[b] Data were based on research by the National Poverty Center on the basis of the Michigan Recession and Recovery Study of adults ages 19-64 in southeastern Michigan. The researchers examined the relationship between various forms of housing instability and health, controlling for prior health problems and sociodemographic characteristics.

poverty on child outcomes are during early childhood development, when children experience poverty for multiple years, and when they live in extreme poverty.[118] The timing of poverty during early adolescence is also important for adolescent achievement.[119]

Biological Pathways

A growing body of research suggests that the similar root causes that lead children to poor educational outcomes and poor health outcomes may not operate via separate pathways but may relate to the biology of brain development and the pathological effects of early childhood exposure to stress and adverse childhood events (ACEs). Children in low SES households are more likely to experience multiple stressors that can harm health and development,[120] mediated by chronic stress.[121] These disruptions can thereby shape educational, economic, and health outcomes decades and generations later.[122]

- *Neuroanatomy and neuroplasticity*: Infants and toddlers exposed to toxic stress, social exclusion and bias, persistent poverty, and trauma may experience changes in brain architecture and development that affect cognition, the ability to learn new skills, behavioral and stress regulation, executive function, and the capacity to adapt to future adversity.[123,124]
- *Endocrine disruption*: Early life stressors also appear to cause physiological increases in allostatic load that promote stress-related diseases later in life.[113] Such stressors may, for example, disrupt the hypothalamic-pituitary-adrenal axis of the endocrine system and stimulate overproduction of stress-related hormones that are thought to adversely affect end organs and lead later in life to heart disease and other adult health problems.[125]
- *Immune dysregulation*: The release of interleukins and other immune reactant proteins is thought to create conditions of chronic inflammation that may increase the risk of heart disease and other chronic diseases later in life.[125]
- *Epigenetic changes*: Chronic stress is thought to affect methylation of DNA and cause epigenetic changes that "turn on" expression of genes that may cause cancer and other diseases.[126]

Enhanced understanding of these biological pathways is shedding light on research, first reported in the 1990s, that called attention to the correlation between adult disease rates and a history of childhood exposure to ACEs. In a seminal study on the subject, the Adverse Childhood Experiences Study, Felitti et al. surveyed more than 13,000 adult patients at Kaiser Permanente and asked whether they recalled childhood exposure to seven categories of ACEs: psychological, physical, or sexual abuse; violence against the mother; or living with household members who were substance abusers, mentally ill/suicidal, or had a history of imprisonment. More than half of the adults recalled ACEs as children, and those with greater trauma were more likely to report unhealthy behaviors as adults (e.g., smoking, physical inactivity, alcoholism, drug abuse, multiple sexual partners) and to have a history of depression or a suicide attempt. The researchers reported a dose-response relationship: those who recalled four categories of ACEs had significant odds ratios for adult diseases, including ischemic heart disease (2.2), cancer (1.9), stroke (2.4), chronic lung disease (3.9), and diabetes (1.6).[127]

The ACE study and subsequent studies with similar results relied on retrospective designs that faced the limitation of recall bias (relying on the memory of adults); recollections of ACEs were vulnerable to the criticism that sick adults might have skewed perceptions of their childhood experiences. Nevertheless, prospective studies that documented ACEs contemporaneously during childhood have also documented higher rates of disease when the children were followed into adulthood. The Centers

for Disease Control and Prevention (CDC) maintains a Web site that is cataloguing the burgeoning research on ACEs,[128] and increasing attention is shifting toward strategies for policy and clinical practice to help ameliorate childhood exposure to ACEs and to buffer their adverse biological and psychosocial effects (see Box 5). This work has relevance to understanding of the education-health relationship to the extent that prior exposure to ACEs affects both educational success and health trajectories.

Box 5. Behavioral Responses to Stress

Children exposed to stress may also be predisposed to take up unhealthy behaviors, such as smoking or unhealthy eating, during adolescence, the age when risky behaviors are often first established and then carried into adulthood. This may be an important contextual factor in understanding the higher prevalence of unhealthy behaviors among persons with limited education, especially if toxic stress affects both education and health outcomes. There is some evidence that stress affects areas of the brain associated with reward and addiction.[129] Dysfunctional coping skills and these changes in brain function may draw children to unhealthy behaviors (e.g., smoking, alcohol or drug use, unsafe sex, violence) as adolescents. These risk factors for disease, along with harmful stress-related physiological changes discussed above, not only increase their subsequent risk of illness and injury but also stifle success in school and employment.[45,122,130]

Summary: What Accounts for the Association of Education and Health?

The building evidence that stress and other contextual factors can have effects on both education and health throughout the life course—as in the lasting effects on development, behavior, learning and health of children—adds important insights for understanding the correlation between education and health. As discussed earlier in the chapter, reverse causality plays some role in the association, and a much larger influence comes from the downstream benefits of education (e.g., greater socioeconomic resources and personal skills), but the upstream influence of adverse experiences on the young child also cannot be ignored. The effects of ACEs on the developing brain and on behavior can affect performance in school and explain setbacks in education—but they can also affect health outcomes. Thus, the correlation between reduced education and illness may have as much to do with the seeds of illness that are planted before children ever reach school age than with the consequences of education itself. The children end up with fewer years of education and greater illness, but an important way to improve their health is to address the root causes that expose children to stress in the first place.

Exploring the Lived Experience

The above conclusions spring from the pages of published research and the theoretical models of scholars in social science, economics, and social epidemiology, but an overlooked perspective is the lived experience of those who contend daily with these living conditions. Our research team at the Center on Society and Health has become increasingly interested in eliciting this perspective and blending the more nuanced insights from community members who face conditions on the ground with the more abstract empirical findings published by academia. In the work described in the second part of this chapter, as well as other recent pilot studies,[131] we have demonstrated that this fresh perspective helps transform causal models emanating from the literature to more sophisticated frameworks that incorporate mediators, moderators, and outcomes that are unfamiliar to academics. Although empirical evidence may be lacking to scientifically document the association between these new elements and health outcomes, we believe the insights are powerful tools to help define a research agenda that outlines testable hypotheses that future research can explore.

The recent focus on patient and stakeholder engagement stimulated by the Patient-Centered Outcomes Research Institute (PCORI)[132,133] has merged with the established discipline of CBPR to bring new energy and interest in community engagement in research and greater respect among academia in studying how insights gathered through engagement affect the design and results of studies. With support from our university's Clinical and Translational Science Awards (CTSA) grant, we have been working since 2011 to engage community members in sharing their perspectives about the influence of social determinants of health. Using an approach we had previously tested to engage community members in developing a causal model *de novo* without knowledge of published research findings, we asked residents of a low-income urban community to map out the pathways linking education and health, and we compared the results with the empirical findings discussed above.

Stakeholder Engagement in Modeling Health Outcomes

Background

The research community increasingly seeks to involve stakeholders in health research, both to enhance accountability and to improve the quality of the research, including increased validity, relevance, acceptance, and sustainability.[134-138] Until recently, lay explanations of health and disease have been denied a "place at the etiological table"[139] and have rarely been used to generate new conceptualizations of the link between social conditions, behavior, and health outcomes. The problem with this has been recognized for two decades: "If research in the field of public health is to develop more robust and holistic explanations for patterns of health and illness in contemporary society, then it must utilize and build on lay knowledge—the meanings that health, illness, disability, and risk have for people"[135] (p. 760).

Participatory research methods have become an important framework for including stakeholders in understanding and addressing health disparities.[17,140-143] The principles of CBPR can provide entrée into more meaningful lay engagement in understanding health outcomes. CBPR "aims to make research more democratic, ensure the poor and people of color are not excluded from decisions that impact their lives, and incorporate local knowledge and lived experience into research and action."[144] CBPR partnerships have engaged in diverse topics, interventions, and study designs that have strengthened methodology in areas such as research design, recruitment, and cultural appropriateness.[140]

Community Engagement in Causal Modeling

CBPR efforts aimed at conceptualization and causal modeling have been uncommon. As long ago as the late 1970s, causal modeling by stakeholders, including community participants, was utilized in development projects, particularly on nutrition. For example, in Zaire a participatory causal modeling approach was used in 1987 to address nutritional problems by engaging a multidisciplinary group that included two international nutrition consultants and diverse local participants. The resulting causal model was used in research design, education, intervention, and community development.[145] This participatory causal modeling approach was described by Beghin et al. in a 1988 publication by the World Health Organization.[146] Lefèvre et al. described a focus-group causal modeling approach as a component of a participatory action research project in Bolivia

and Peru and proposed that this method might be useful for comparing perceptions or competing explanations.[147]

More recently, The Dan River Partnership for a Healthy Community, composed of community stakeholders and researchers from the Department of Human Nutrition, Foods and Exercise at Virginia Tech, used the Comprehensive Participatory Planning and Evaluation (CPPE) process within a CBPR framework to focus on obesity in the region.[148] The problem-assessment phase of the project included a causal analysis workshop to explore potential mechanisms and root causes of obesity. "CPPE causal models do not necessarily have to portray a hierarchal structure or infer causation, rather they are meant to uncover the complexity of problems and encourage participants to discuss potential solutions"[148] (p. 49). The models were used to prioritize community interventions. Such exercises in causal modeling, as done in the CPPE process, are meant to build consensus among stakeholders on the factors affecting an identified problem, working backward from problems to root causes, with the goal of identifying appropriate solutions (and potential research hypotheses to study).[149]

Another participatory modeling approach, applied in the field of systems dynamics, is group model building (GMB). GMB "is a participatory method for involving people in a modeling process" that focuses on understanding and solving systems problems. Community-based system dynamics explicitly includes community members in the process.[150] Stave describes using a participatory model building process to involve stakeholders in environmental decisions.[151]

In the study we present here, our specific aim was to explore whether community stakeholders would develop a causal model that added to the pathways and mechanisms already hypothesized in the academic literature (and reviewed earlier in this chapter). Secondarily, we sought to explore whether the lived experiences of participants would elucidate new descriptions and nuances about pathways that are already recognized but are not fully understood. Although the theoretical model and empirical work involved in elaborating the relationship between education and health have evolved since health disparities first garnered wide attention,[152] we believe this participatory approach provides a unique framework for testing and expanding the theoretical model.

Engaging Richmond

The CBPR partnership that conducted this exercise, known as Engaging Richmond, is an ongoing program that involves community researchers[c] who are residents of the East End, a low-income African American neighborhood in Richmond City and faculty and staff of Virginia Commonwealth University. The Engaging Richmond community researchers have received Institutional Review Board (IRB)-certified research training, conducted and analyzed focus group and interview data,[153] successfully recruited participants for research, and disseminated findings in the community.[154,155] The idea of engaging stakeholders in crafting conceptual models arose from the initial successes of the Engaging Richmond team in modeling various health-related outcomes for proposed ideas and a report on the potential connections between food stamp benefits and health.[131]

[c] The community researchers on the CBPR team are residents of Richmond City's East End who have an ongoing role on the research team and have received training in various aspects of the research process. The community researchers who collaborated in the process described here included two men and six women who received training in social determinants of health research and have experience living in communities with low educational attainment and poor health outcomes.

Methods

Members of the CBPR team worked together in a facilitated concept mapping exercise designed to tap into stakeholders' experiences of how education is related to health outcomes. The goal was to develop a conceptual model of the social, behavioral, environmental, biological, and other factors that link education and health and to place their lived experience—and their understanding of the cultural and social context—into an analytical framework. The process was not about attaining consensus but instead tapping stakeholders' experiences to generate new insights and ideas to inform the causal model.

The group's tasks were to list the potential factors influencing the relationship between health and education and, following some training, to sketch a diagrammatic model of how determinants are interrelated on an upstream-downstream (distal-proximate) continuum. The process was facilitated by a faculty member (E.B.Z.) who had worked with the team for more than 2 years. The venue for the meeting was a community center in the residents' neighborhood that was regularly used for team meetings. The facilitator introduced the goal of the exercise as follows:

> "We are here to talk about how education affects health. We want to draw on the experiences of everyone here. The purpose of this exercise is to find new ways of looking at this relationship between education and health, and we are going to focus on various factors that you think might affect the relationship."

In the first part of the exercise, the community researchers followed a series of facilitated steps to individually brainstorm, identify, and record a broad list of factors that they believed might be influential in the relationship between education and health. The facilitator encouraged them to list "everything that comes to your mind that you think might be part of this relationship between education and health… anything you can think of… that impacts how one's own education might affect their health." Participants were encouraged to draw on a range of experiences in thinking through the topics. The community researchers then reviewed a prepared list of many potential factors, grouped into domains (social, behavioral, family/community, physical/mental, demographic, health care, genetic, environmental, and attitudes/beliefs), and were given the opportunity to expand or change their initial list of factors, as well as to eliminate any factors that they did not consider influential. They were then asked to highlight the factors they had selected which they perceived to be most important.

In the next step, the group discussed the factors they had highlighted, indicated which factors to include in the model and which to exclude, and decided how to group factors. Participants provided examples to illustrate why specific factors were important. As they began listing behavioral factors, the facilitator instructed them that, "You need to kind of think through: How does this happen? How does education affect diet?... Whatever it is, you've got to think these things through." Box 6 shows two examples of how this process unfolded. After listing and discussing factors across the various domains, the group agreed on a final list of factors.

Training in conceptual modeling was provided by the facilitator. Although this team had been exposed to causal path diagrams on previous projects,[131] the training was useful to present key terminology and review the purpose and structure of path diagrams.

In the final step, the facilitator worked with the team to sketch a causal path diagram of the factors listed in step one. Team members took turns presenting particular factors from the list and discussing where, in relation to other elements of the diagram, they might be important and elaborating through examples. In the process of sketching the path diagram, the group was asked to decide whether each factor added to the model was exogenous (a variable that influences the value of other variables in the model, but whose own value is determined outside of the model[156]) or a mediator (a variable that lies intermediate between independent causal factors and a final outcome).[157] Sketching the model was informed by encouraging the participants to consider how the factors interrelated (see example in Box 7), which in many cases pointed to the bi-directional linkages that make these relationships so complex. For instance, there was extensive discussion about the many factors that affect educational achievement.

Factors were iteratively added to the model as time permitted, and the group then reviewed the diagrammed relationships, adding or removing arrows between factors to more accurately reflect the participants' sense of the causal pathways.

The modeling session was transcribed, and the data were compared to the key elements identified in the literature for explaining the linkages between education and health (see results section below). The community researchers reviewed and provided feedback on the draft models, as well as the findings presented here.

Box 6. Identifying Relevant Indicators

Example 1: Social Skills and Sleep

Participant: *"I have social skills, sleep habits, exercise."*
Facilitator: *"Social skills - any example?...*
Participant: *"For instance, when I think of social skills I think of the ability to interact with people, the ability to mingle with strangers, you know, go into environments that you are unfamiliar with…* [continues]"
Facilitator: *"And what about sleep habits? How would you describe that in terms of education and health?"*
Participant: *"From my experience, and my friends,' when exam time comes, and the pressure and all of that, those poor sleep habits or not being able to get enough sleep, leads to drug abuse…"*

The discussion continued, and reflected the bi-directional nature of education and health, as many of the examples illustrated how behavioral and health issues impact education.

Example 2: Accessing Information and the Internet

Participant: *"By us having Internet now, people that wouldn't have access to certain information can access it. When I was growing up and we used to have to do things that involved the encyclopedia, we never had the whole volume of the encyclopedia, so I would have to go to the library. But there was no libraries near where I was, so I had to wait 'til I got to school to do my projects. Whereas now my kids can go on the Internet and pull up whatever they need to pull up for anything."*
Facilitator: *"What about health information?"*
Participant: *"That's what I was getting to next. Over the past 3 months I was changing my health, as far as eating vegetables and things like that. A lot of the diets and the juices that I made, I found it on the Internet."*

Results

The community researchers on the CBPR team focused on numerous mediators in the link between education and health, many of which mirror the predominant frameworks in the existing literature. For example, Adler and Stewart[152] have already articulated important components of the causal pathway. Here we focus on residents' insights that added new perspectives or emphasized different aspects of those causal factors, while highlighting certain specific aspects of the experiences of low income and minority groups. We present these in the next section, following the same structure as the first half of this chapter for consistency.

Box 7. Hypothesizing Pathways

Facilitator: *"So, where else is school going to take us, besides just what you learned about health?"*
Participant: *"The workplace."…*
Participant: *"My income is going to help work with my motivation and outlook, because I may be able to go to the gym. Possibly, I may do it; possibly not. My income is going to help my health behavior."*
Participant: *"My word was lifestyle."…* [discussion moves on to the community environment]
Facilitator: *"Does environment affect this* [points to indicator] *or does it go straight to health?"*
Participants: *"It affects your lifestyle too."*

Impact at the Individual Level

The first half of the chapter noted that an important pathway by which education impacts health is through the development of a range of skills and traits, including cognitive skills, problem solving, and diverse personality traits. The community researchers focused on the types of opportunities that help to develop non-cognitive skills, particularly social skills, as well as the reasons why social skills are important to health. They particularly focused on the development of effective social skills as a function not only of formal education but also the informal educational exposure that can occur outside the classroom. Some examples of educational opportunities leading to enhanced soft skills cited by the community researchers included opportunities they had experienced while young to attend art performances or read literature. They also mentioned the importance of community programs, such as summer camps and youth development programs, which provided the opportunity to engage young people in new experiences and interactions:

> "Kids that participate in extra-curricular, the summer camps and things like that, they learn those social skills. They learn the environmental skills. They get exercise. They learn to… their attitudes and personalities tend to be a little better than the kid that stays locked up in the house playing video games. So it's like, it's a positive that goes to it."

Opportunities to develop non-cognitive skills at school, community programs, and even daily activities such as getting to and from school, were discussed within a larger framework that highlighted the many possible repercussions that these skills can offer throughout the life course. For example, they described situations in which social skills are an important precursor to other dispositions and behaviors that are important to good health. Strong social skills lay the foundation for opportunities to embrace new situations and get along with others:

> "When I think of social skills I think of the ability to interact with people, the ability to mingle with strangers, go into environments that you are unfamiliar with."

To gain new information:

> "I'm sure there's probably some preventive measures I could have learned to strengthen or to help with that [medical condition], but at that point in time I wasn't that social, literate person.

And, to reduce conflict:

> "… you have to know how to deal with people. We don't always agree, but we also know how to just disagree and part ways. And everybody doesn't know how to do that. Sometimes they want to argue about it, and it's not even that serious..."

Ultimately, the development of social skills and other non-cognitive skills was linked to a cascade of possible effects throughout life, impacting social networks and isolation, attitudes, ability to obtain and utilize health-related information, personal health behaviors, and the ability to navigate the health care system. This issue is salient to community researchers from low-income, segregated neighborhoods because despite the importance of participating in enriching activities, youth from lower income families are less likely to participate in most contexts, with the exception of tutoring. Participation rates also vary by ethnicity and race, with Latino youth particularly underrepresented.[158] Children who reside in poor urban neighborhoods and isolated rural areas tend to have reduced access to programs and greater barriers to participation.[159]

The academic literature provides evidence that youth participation in organized activities affects educational attainment and achievement, behavioral problems (including substance abuse), and psychosocial competence (e.g., emotions, motivation, initiative, and self-esteem).[159] Literature that relates youth development opportunities to health outcomes is less extensive. A recent review of the impact of Positive Youth Development (PYD) programs failed to find evidence for improved health outcomes for youth with chronic illness due to a lack of rigorous evaluation.[160] Gavin et al. identified PYD programs associated with sexual and reproductive health, but the findings were still relatively weak.[161] Studies tend to show a positive association between alcohol use and sports participation (as least for some types of sports) and a negative association with illicit drug use.[162] These studies, which tend to focus on sports or formal youth programs, examine some health outcome measures but do not focus on the mechanism by which such opportunities may ultimately impact health. The community researchers point to possibilities such as reduction of anxiety, stress, isolation, and conflict and access to new forms of information and new opportunities.

As noted in Part I, education can impact health through its effects on personal health behaviors, including engagement in risky behaviors, opportunities to learn about health, and availability of resources to make healthy choices. The community researchers described the potential impact of a range of factors (e.g., knowledge, health beliefs, and mental status) on personal health behaviors. They also discussed the potential influence of traits and attitudes on health behaviors and how they may be affected directly by formal and informal education.

Through the modeling exercise, the community researchers noted a number of ways that attitudes can impact health, but much is unknown about whether these attitudes are impacted by education and how important the attitudes are to health outcomes. They provided a number of examples about how such attitudes could have an effect on health behaviors, including setting priorities, facilitating

or hindering access to information, and ability or willingness to seek help. Attitudes they felt might impact health behaviors included materialism, hostility, anger, and pessimism and willingness to change.

Materialism, or the importance attached to material possessions, was perceived by the community researchers as a barrier to effective decisionmaking and the setting of healthy priorities, especially for young people. They noted that materialism has "warped reality for a lot of people" and can have negative effects on resource allocation. Although likely an underexplored causal link in the education/health literature, materialism has been linked to subjective well-being, self-esteem, and stress[163,164] and risky behavior.[165]

Hostility, anger, and pessimism were other attitudes identified as potentially important. Although there are many possible mechanisms whereby hostility may impact health,[166] these community researchers focused on mistrust and its effects on receiving needed information or help:

> "The hostility comes when a lot of times you talk to people and they think you're talking against them or belittling them, and really you're just trying to get them to go or just trying to educate them."

As discussed previously, the cognitive and non-cognitive skills developed through education can also impact individuals' ability to navigate the health care process.[26] This topic came up during the community researchers' model development as well. They noted that education can improve access to quality health care by enhancing communication skills and the ability to advocate for quality care. They added that challenges in diverse skill domains may mean that individuals with less education do not benefit as much from the information that is available:

> "The information is there. You see a lot of pamphlets getting dust on them, and they also have little things that they have around the community. Barely anyone shows up other than the service providers and who's with them. And I'm just saying it's like the information's there, it's a matter of going to get the information and participating and just being involved enough to find out about what is out there, what is going on."

Or, they are less able to deal with the complexity of the health system:

> "Bureaucracy of applying for health care, and not understanding all that whole co-pay, how it's gonna affect your paycheck, when you apply for health care, when you have employment."

Referrals for specialty and followup care seemed to be particularly difficult to navigate without the communication skills and cognitive skills necessary to engage in the interaction with health professionals:

> "…you don't know exactly why you're being sent to another doctor, because it wasn't worded so that you could relay that information when you were making your own appointment, if you needed to make that appointment yourself."

Disparities in health care quality and access are well documented by socioeconomic status, race and ethnicity, and even while health care overall may improve, reducing disparities has proven to be quite difficult.[167] Attention to disparities in skills, communication, and access to resources (and how those play out in lived experience) that have their roots in educational disparities may prove a promising route to reducing otherwise intractable disparities in access, quality, and outcomes. This brief exercise has highlighted a few of these.

Another important set of factors at the individual level, discussed earlier in this chapter, includes access to economic and social resources. The community researchers, echoing the fundamental importance of the pathway between education and health via employment, discussed multiple pathways by which employment may impact health, including exposure to work-related stress, effects on motivation and outlook, ability to build social networks, and economic impact on the environment where one lives.

An important pathway runs from lower educational attainment to lower-status occupations and employment-related stress. The community researchers added nuance about the stresses of a poor education related to job insecurity, long work hours, work/family conflicts, and conflicts with co-workers.

> "A father in a company misses a whole lot of plays, a whole lot of educational programs, a whole lot of PTA meetings."

> "[work-related stress] depends on your job. Depends on what you see and what you encounter that can lead to those sleepless nights or whatever..."

At a more fundamental level, the community researchers noted that the income resulting from one's education can affect motivation, outlook, and lifestyle, which in turn may affect health behaviors.

> "My income is going to help work with my motivation and outlook, because I may be able to afford to go to the gym. Possibly, I may do it. Possibly not. My income is going to help my health behavior."

Social networks and peer groups play an important role in health.[53] The community researchers linked the development of social skills to effects throughout the life course on social integration and isolation. Many studies of the effects of social isolation on health focus on the elderly, whereas the community researchers felt that people who experience social marginalization due to behavior or various other reasons may suffer isolation that leads to ill health. In one example, their causal model connects lack of education to stress and anxiety, which may cause social isolation. They described the potentially negative impacts of social isolation, such as stress, impaired communication with others, and inability to solicit help.

> "Living apart from others to the point where you can't even get the help you need because you're so isolated. People don't know how to communicate with you because you keep yourself so isolated."

Impact at the Community Level

The community researchers mentioned the role of place-based determinants of health, such as access to healthy food outlets and the greater risk of exposure to toxins and environmental risks in disadvantaged neighborhoods that are populated by people with limited education. Many of the deficiencies they noted in access to and the quality of health care transcended individual-level resources and abilities and related to the service environment in the community, such as the availability of treatments, appropriateness of care, coordination of care, cultural competency, and barriers to health care. Lack of access to services such as transportation has a significant effect on residents' ability to access opportunities, including health care:

> "That referral that you might have also, and I'm continuously saying, are you listening to me? I don't have transportation. And they say you can go through the insurance to get this set up... I call transportation it's like, I don't fall short. I'm there. My transportation is 2 hours late."

Unequal treatment was also a concern, including inadequate availability of preventive care. The community researchers felt that their community was less likely to receive the type of preventive health information that would be more accessible in the more affluent communities populated by people with higher education.

> "...when you look at most things that are being done in our community, it's always from that intervention side. Very little prevention is being offered to us. How do we prevent? Very little."

> "In certain places, there is certain information that they will give to this group of people that they wouldn't give this group of people. So they will know how to prevent high blood pressure, as our information would be more so what to do after you get it. Cause you're going to get it."

Finally, children and adults in disadvantaged communities may be more likely to experience chronic stress or trauma, and community researchers were concerned about the appropriateness of their diagnostic evaluations and treatments:

> "What would it be as far as misdiagnosis, as far as it could be a learning disability but it could be something else that's preventing the child from being able to function in the classroom or preventing the adult from being able to function at work. And it could be a health problem or it could be a learning problem. It could be environmental. It could be literacy, or whatever. But if you don't have the [resources] …, it could be you don't have the right kind of insurance to be able to find this. You don't have the right kind of doctors available to your call. You could have that doctor that's just doing enough to get you in and get you out."

The problem of childhood trauma and its relationship to conditions such as attention deficit hyperactivity disorder (ADHD) appears in the clinical literature,[168] but insights about this relationship from parents and service providers in communities particularly affected by high levels of trauma exposure point to sources of concern and may help identify, through further research, areas of intervention.

Contextual Factors

The participants highlighted the intersections between access to health care (and other necessary social supports) and public policy. Policy decisions contribute to gaps in health insurance coverage for the underserved, and the participants discussed how this contributes to health complications.

> "When people have to have major surgeries and stuff done and not having Medicaid or health insurance. And it can be life threatening and they are scared to go and get their self checked out, just checked out when they know there's something major going on with them, because they don't have health insurance or Medicaid."

Their responses underscored the ways in which individuals with lower educational attainment, low skill levels, or poor mental health would be disadvantaged by the bureaucracy and documentation

required to access social welfare programs. Individuals with limited education and their families are more vulnerable due to the burdens placed on them by bureaucratic structures and regulations.[95] The group pointed out that often individuals who might otherwise qualify for services and supports (e.g., Medicaid or school programs) could miss out because they cannot keep up with paperwork and rules.

One participant discussed a local program for the uninsured:

> "I think about the [program] and how that lasts for a year and then you have to reapply. I get that, but that can be a hardship too. Why should you have to apply every year? First of all, they don't send you a notification asking you to reapply. So when that date comes around, oftentimes you forget. And so then you realize, oh, I don't have insurance anymore. You know, if you don't have certain documents to prove that you're financially eligible they won't accept you. So there are barriers to things that are designed for you."

In addition, populations that are disadvantaged by an inadequate education are more likely to rely on public services that may fall short of expectations because society has invested insufficient resources. For example, public transportation may be inadequate, forcing patients who lack transportation alternatives to rely on medical transportation services that may not be trustworthy. Public services are subject to budget cuts, and restrictive welfare programs may inadequately cover the needy, leading to further disadvantage.

An overarching theme in the discussion that transcended the specific elements was a narrative of exclusion. Throughout the process, the team members made links to contextual factors that, more often than not, seemed to progressively diminish the chances that individuals with little education, poor skills, and few economic resources could achieve positive health outcomes. They described a tableau of contextual factors—ranging from failing schools to complex bureaucratic structures and 'top down' decisionmaking—that distance individuals from success in education and health but are not explicitly mentioned in published causal models. The risks associated with failing schools, under-resourced communities, and unequal access to quality health care are intensified when individuals with limited education and income face the additional challenges of fewer social skills and social networks, restricted access to information and the ability to use it, limited ability to advocate for quality care, and increased exposure to stress. The link between social exclusion and health has been recognized[169] but is not often explicitly included in the education/health model. Participatory modeling may serve to draw some attention toward the societal factors that are often overlooked in media and academic accounts of health outcomes and the recommendations and interventions subsequently developed to address disparities.[170,171]

Throughout the exercise, the community researchers framed the connections between education and health not just as a causal path traversed by individuals, but as one whose shape and character were dependent upon the larger social context. The resounding impact of race, class, gender, and age discrimination was the backdrop for discussions of educational opportunity, workplace experiences, health care, and policy.

Discussion of Engagement Exercise

The process described in this section presents an approach that extends prior, predominantly practical, applications of participatory modeling (e.g., prioritizing community interventions) to a role in advancing theory and scholarly inquiry. It explores the possibility that people outside of

academia may be able to help refine our understanding of complex phenomena by positing factors and relationships less familiar to investigators who do not share their life circumstances. None of the observations described here are meant to stand as evidence, but they are intended to illustrate how the process (1) may provide the bases for hypotheses that can be further explored, or (2) provide deeper understanding of how the highlighted relationships may operate and why they may be important mediators or moderators of health disparities.

This small pilot has many limitations. The insights come from a limited sample of participants from one neighborhood of a southern city. Other findings would undoubtedly emerge with greater diversity and a larger number of participants. In any setting, delving into the broad expanse of variables that occupy the relationship between education and health—a web of influences noteworthy not only for its breadth but for the bi-directionality and endogeneity of the many factors involved—is not a simple task. Others may wish to continue gathering community perspectives on upstream social determinants by breaking this complex model into smaller components. Despite its limited scale and the complexity of the topic, the community researchers who participated in this exercise demonstrated not only a wealth of insight but an ability to put their personal experiences into context and breathe life into a critically important issue on which their voices are too infrequently heard.

Implications for Practice

The relationships between education and health are relevant to the clinician, beginning with the patient's ability to understand diagnostic information and treatment recommendations but extending to larger issues. Health care professionals, social workers, and other service providers must consider the knowledge and literacy of clients to ensure that instructions and choices are fully understood, ranging from reading prescription bottles to understanding how to file for claims. But the education-health relationship has relevance to practitioners beyond the level of one-on-one care, because their cachet creates leverage to promote efforts in the community to improve educational opportunities and create conditions in early childhood to put youth on a path for socioeconomic success and better health. Physicians and other health care professionals can speak to the health benefits of community investments that expand opportunities for preschool and primary/secondary education. However, this chapter has also emphasized that the links between education and health are influenced by policy decisions outside of schools, including neighborhood conditions ranging from sidewalks to street violence, food security, reliable housing, job training, and safety net programs for the disadvantaged. Better grades and higher graduation rates are vital goals, but meaningful effects on population health require an integrated plan for upstream and downstream determinants.

Implications for Research

As noted earlier, the factors surrounding the relationship between education and health are the subject of research in different disciplines that are of uneven quality, and closing the many holes in the evidence is a research priority. Chief among these is the reliance on cross-sectional and ecological evidence that does not adequately tease apart issues of endogeneity and leaves many unanswered questions about causal pathways. The research challenges are inherently transdisciplinary, requiring the integration of traditional population health sciences (e.g., epidemiology) with social and political science, education research, and the use of mixed methods to blend quantitative and qualitative insights. Standards of evidence used for clinical effectiveness are not always applicable to these topics.[172] A particular need exists to bridge the divide between research in population health and

education and to share work across the silos to achieve more integrated research paradigms. Future research should also address the role of contextual factors surrounding the individual elements of the socioecological model, and their collective behavior as complex systems, through simulation modeling and other modern tools for predictive analytics.[173] Finally, the research agenda must address the information needs of policymakers, stakeholders, and change agents who are positioned to make improvements in education and health.[174]

A criticism of social epidemiology and other efforts to identify social determinants of health has been a focus on establishing correlations between social factors and health, with comparatively little attention to the mechanisms through which these factors impact health outcomes, and in turn, inattention to promising leverage points for interventions or policy change.[175] While we reiterate that this pilot serves to illustrate the potential of participatory processes in extending our understanding of these mechanisms rather than providing an empirical base, by scanning the input from the community researchers one might begin to see an emergent list of possible leverage points, from greater access to youth development opportunities to changing bureaucratic processes that make participation in public benefit programs difficult for people with low levels of education or other challenges.

Our approach emphasizes respect and parity in patient/community engagement, an orientation that is gaining ascendance among health services researchers who increasingly recognize the insights and innovations in the design of studies and interpretation of results that become possible when those affected by interventions are invited to participate as research partners—not as human subjects but as coinvestigators. An approach in which patients and community members are treated as coequal partners in the research enterprise creates opportunities for scholarship that are forfeited by more traditional, and sometimes patronizing, engagement methods that examine data through a lens shaped by academics based on theoretical models rather than incorporating the perspective of those who live amid the conditions under study. The respect afforded by the new approach has the added, and important, advantage of helping to build trust with a lay public and especially with marginalized minority communities that were dubious about the sincerity of researchers who sought their input or the ability of professional investigators to truly understand real-world conditions. Communities that have historically experienced condescension rather than respect welcome such collaborative approaches.

Acknowledgments

We gratefully acknowledge the hard work and thoughtful insights of the Engaging Richmond CBPR team: Chanel Bea, Brenda Kenney, Chimere Miles, Toni Mitchell, Valeria Burrell Muhammed, Marco Thomas, Albert Walker, and Breena Wingo. We also acknowledge the important contributions of Sarah Simon. This work was conducted as part of the Education and Health Initiative, funded by the Robert Wood Johnson Foundation under grant No. 70227; our community engagement efforts were supported by CTSA award No. UL1TR000058 from the National Center for Advancing Translational Sciences and a CTSA supplement (UL1RR031990). The contents of this chapter are solely the responsibility of the authors and do not necessarily represent official views of the National Center for Advancing Translational Sciences, the National Institutes of Health, the Agency for Healthcare Research and Quality, or the U.S. Department of Health and Human Services.

With permission of all parties, parts of this publication were previously published by the Institute of Medicine: Zimmerman E, Woolf SH. Understanding the relationship between education and health. Discussion Paper. Washington, DC: Institute of Medicine; 2014.

Authors' Affiliations

Center on Society and Health, Virginia Commonwealth University, Richmond, VA.

Address correspondence to: Emily Zimmerman, PhD, Center on Society and Health, Virginia Commonwealth University, 830 Main St, Suite 5035, PO Box 980212, Richmond, VA 23298; email ezimmerman@vcu.edu

References

1. Link BG, Phelan J. Social conditions as fundamental causes of disease. J Health Soc Behav 1995;36:80–94.
2. Closing the gap in a generation: Health equity through action on the social determinants of health. Geneva, Switzerland: World Health Organization, Commission on Social Determinants of Health; 2008. Available at http://www.who.int/social_determinants/thecommission/finalreport/en/. Accessed November 8, 2014.
3. National Research Council, Institute of Medicine. U.S. health in international perspective: shorter lives, poorer health. Washington, DC: National Academies Press; 2013.
4. County health rankings, 2013. Madison, WI: University of Wisconsin Population Health Institute; 2013. Available at www.countyhealthrankings.org. Accessed November 8, 2014.
5. Kulkarni SC, Levin-Rector A, Ezzati M, et al. Falling behind: life expectancy in U.S. counties from 2000 to 2007 in an international context. Popul Health Metr 2011;9(1):16.
6. Evans BF, Zimmerman E, Woolf SH, et al. Social determinants of health and crime in post-Katrina Orleans Parish. Richmond, VA: Virginia Commonwealth University, Center on Human Needs; 2012.
7. Goldman D, Smith JP. The increasing value of education to health. Soc Sci Med 2011;72:1728–37.
8. Olshansky SJ, Atonucci T, Berkman L, et al. Differences in life expectancy due to race and educational differences are widening, and many may not catch up. Health Aff 2012;31:1803-13.
9. Montez JK, Berkman LF. Trends in the educational gradient of mortality among U.S. adults aged 45 to 84 years: Bringing regional context into the explanation. Am J Public Health 2014;104(1):e82–e90.
10. Montez JK, Zajacova A. Trends in mortality risk by education level and cause of death among U.S. white women from 1986 to 2006. Am J Public Health 2013;103(3):473-9.
11. Jemal A, Ward E, Anderson RN, et al. Widening of socioeconomic inequalities in U.S. death rates, 1993–2001. PLoS ONE 2008;3(5):1–8.
12. Health, United States, 2011: with special feature on socioeconomic status and health. Hyattsville, MD: National Center for Health Statistics; 2012.
13. Ross CE, Masters RK, Hummer RA. Education and the gender gaps in health and mortality. Demography 2012 Nov;49(4):1157-83.
14. Schiller JS, Lucas JW, Peregoy JA. Summary health statistics for U.S. adults: National Health Interview Survey, 2011. Vital Health Stat 2012;10(256).
15. Institute of Medicine. Promoting health: intervention strategies from social and behavioral research. Washington, DC: National Academies Press; 2000.
16. Montez JK, Hummer RA, Hayward MD. Educational attainment and adult mortality in the United States: A systematic assessment of functional form. Demography 2012;45:315-36.
17. Israel BA, Eng E, Schulz AJ, et al. Methods in community-based participatory research for health. San Francisco, CA: Jossey-Bass; 2005.
18. Galea S, Riddle M, Kaplan GA. Causal thinking and complex system approaches in epidemiology. Int J Epidemiol 2010;39:97-106.
19. Diez Roux AV. Complex systems thinking and current impasses in health disparities research. Am J Public Health 2011;101:1627-34.
20. Mirowsky J, Ross CE. Education, learned effectiveness and health. London Rev Educ 2005;3(3):205–20.
21. Heckman JJ, Kautz T. Hard evidence on soft skills. Labour Econ 2012;19:451–64.
22. Roberts BW, Kuncel NR, Shiner R, et al. The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for

predicting important life outcomes. Perspect Psychol Sci 2007;2(4):313–345.
23. Ross CE, Wu CL. The links between education and health. Am Sociol Rev 1995;60(5):719–45.
24. Maitra S. Can patient self-management explain the health gradient? Goldman and Smith's "Can patient self-management help explain the SES health gradient?" (2002) revisited. Soc Sci Med 2010;70:802–12.
25. Cutler DM., Lleras-Muney A. Understanding differences in health behaviors by education. J Health Econ 2010;29(1):1–28.
26. Berkman, ND, Sheridan SL, Donahue KE, et al. Low health literacy and health outcomes: An updated systematic review. Ann Intern Med 2011;155(2):97–107.
27. Rosas-Salazar C, Apter AJ, Canino G, et al. Health literacy and asthma. J Allergy Clin Immunol 2012;129(4):935–42.
28. Willems S, De Maesschalck S, Deveugele M, et al. Socio-economic status of the patient and doctor-patient communication: does it make a difference? Patient Educ Counsel 2005;56:139–46.
29. Garrett BE, Dube SR, Winder C. MMWR 2013 Nov 22;33;62(3):81-4
30. Kant AK, Graubard BI. Secular trends in the association of socio-economic position with self-reported dietary attributes and biomarkers in the US population: National Health and Nutrition Examination Survey (NHANES) 1971-1975 to NHANES 1999-2002. Pub Health Nutr 2007;10(2):158-67.
31. Behavioral Risk Factor Surveillance System. WEAT: Web-Enabled Analysis Tool. Atlanta, GA: Centers for Disease Control and Prevention, Division of Population Health. Available at http://nccd.cdc.gov/s_broker/WEATSQL.exe/weat/index hsql. Accesssed May 13, 2015.
32. Kanny D, Liu Y, Brewer RD, et al. MMWR 2013 Nov 22;62(3):77-80.
33. Pampel FC, Krueger PM, Denney JT. Socioeconomic disparities in health behaviors. Annu Rev Sociol 2010;36:349–70.
34. Estabrooks PA, Lee RE, Gyurcsik NC. Resources for physical activity participation: does availability and accessibility differ by neighborhood socioeconomic status? Ann Behav Med 2003;25(2):100–4.
35. Brownell KD, Kersh R, Ludwig DS, et al. Personal responsibility and obesity: a constructive approach to a controversial issue. Health Aff 2010;29(3):379–87.
36. Skilled for life? Key findings from the Survey of Adult Skills. Paris, France: Organisation for Economic Co-Operation and Development (OECD); 2013. Available at http://skills.oecd.org/documents/SkillsOutlook_2013_KeyFindings.pdf. Accessed November 11, 2014.
37. National Research Council. America's uninsured crisis: consequences for health and health care. Washington, DC: National Academies Press; 2009.
38. Freeman JD, Kadiyala S, Bell JF, et al. The causal effect of health insurance on utilization and outcomes in adults: a systematic review of U.S. studies. Med Care 46(10):1023-32.
39. Hadley J. Sicker and poorer—the consequences of being uninsured: a review of the research on the relationship between health insurance, medical care use, health, work, and income. Med Care Res Rev 60(suppl 2):S3-75.
40. Employment projections. Washington, DC: United States Department of Labor, Bureau of Labor Statistics; 2013. Available at http://www.bls.gov/emp/ep_chart_001 htm. Accessed November 11, 2014.
41. Athar HM, Chang MH, Hahn RA, et al. Unemployment—United States, 2006 and 2010. CDC health disparities and inequalities report—United States. MMWR 62(3):27-32. Available at http://www.cdc.gov/mmwr/pdf/other/su6203.pdf. Accessed May 14, 2015.
42. Braveman PA, Cubbin C, Egerter S, et al. Socioeconomic disparities in health in the United States: What the patterns tell us. Am J Public Health 2010;100(Suppl 1):S186–S96.
43. Shippee TP, Wilkinson LR, Ferraro KF. Accumulated financial strain and women's health over three decades. J Gerontol B Psychol Sci Soc Sci 2012;67(5):585–94.
44. Julian TA, Kominski RA. Education and synthetic work-life earnings estimates. American Community Service Reports (ACS-14). Washington, DC: U.S. Census Bureau; 2011.
45. Braveman P, Egerter S, Barclay C. Issue brief series: exploring the social determinants of health: income, wealth, and health. Princeton, NJ: Robert Wood Johnson Foundation; 2011.
46. Denavas-Walt C, Proctor BD, Smith JC. Income, poverty, and health insurance coverage in the United States. Current Population Reports. Washington, DC: U.S. Census Bureau; 2013. Available at http://www.census.gov/hhes/www/hlthins/. Accessed May 18, 2015.
47. Behavioral Risk Factor Surveillance Syustem, 2010. Atlanta, GA: Centers for Disease Control and Prevention, office of Surveillance, Epidemiology, and Laboratory Services; 2014. Available at http://www.cdc.gov/brfss/data_tools htm. Accessed May 18, 2015.
48. National healthcare quality report 2012. Rockville, MD: Agency for Healthcare Research and Quality; 2013. Available at http://www.ahrq.gov/research/findings/nhqrdr/nhqr12/2012nhqr.pdf. Accessed November 11, 2014.
49. Institute of Medicine. Hidden costs, value lost: uninsurance in America. Washington, DC: National Academies Press; 2003.
50. Seeman TE. Social ties and health: the benefits of social integration. Ann Epidemiol 1996;6(5):442–51.
51. Kaplan GA, Wilson TW, Cohen RD, et al. Social functioning and overall mortality: prospective evidence

from the Kuopio Ischemic Heart Disease Risk Factor Study. Epidemiology 1994;5(5):495–500.
52. Berkman LF, Glass T, Brissette I, et al. From social integration to health: Durkheim in the new millennium. Soc Sci Med 2000;51:843–57.
53. Smith KP, Christakis NA. Social networks and health. Ann Rev Sociol 2008;34:405–29.
54. Christakis NA, Fowler JH. The collective dynamics of smoking in a large social network. N Engl J Med 2008;358:2249–2258.
55. Thoits PA. Stress and health: major findings and policy implications. J Health Soc Behav 2010;51(Suppl):S41–S53.
56. McEwen BS, Stellar E. Stress and the individual: mechanisms leading to disease. Arch Intern Med 1993;153(18):2093–101.
57. Pearlin LI. The sociological study of stress. J Health Soc Behav 1989;30(3):241–56.
58. Seeman TE, Singer BH, Rowe JW, et al. Price of adaptation: allostatic load and its health consequences. Arch Intern Med 1997;157:2259–68.
59. Karlamangla AS, Burton HS, Seeman TE. Reduction in allostatic load in older adults is associated with lower all-cause mortality risk: MacArthur studies of successful aging. Psychosom Med 2006;68:500–7.
60. Krueger PM, Chang V. Being poor and coping with stress: health behaviors and the risk of death. Am J Public Health 2008;98:889–96.
61. Kawachi I, Berkman LF. Neighborhoods and health. New York: Oxford University Press; 2003.
62. Kawachi I, Subramanian SV. Neighbourhood influences on health. J Epidemiol Commun Health 2007;61(1):3–4.
63. Ross CE, Mirowsky J. Neighborhood socioeconomic status and health: context or composition? City Commun 2008;7(2):163-79.
64. Leventhal T, Brooks-Gunn J. The neighborhood they live in: the effects of neighborhood residence on child and adolescent outcomes. Psychol Bull 2000;126(2):309–37.
65. Steptoe A, Feldman PJ. Neighborhood problems as sources of chronic stress: development of a measure of neighborhood problems and associations with socioeconomic status and health. Ann Behav Med 2001;23:177–85.
66. Shankardass K, Dunn JR. How goes the neighbourhood? Rethinking neighbourhoods and health research in social epidemiology. In O'Campo P, Dun JR (Eds), Rethinking social epidemiology (p. 137-56). Dordrecht: Springer; 2011.
67. Rhodes T, Wagner K, Strathdee SA, et al. Structural violence and structural vulnerability within the risk environment: theoretical and methodological perspectives for a social epidemiology of HIV risk among injection drug users and sex workers. In O'Campo P, Dun JR (Eds), Rethinking social epidemiology (p. 205-30). Dordrecht: Springer; 2011.
68. Cummins S, Curtis S, Diez-Roux AV, et al. Understanding and representing "place" in health research: a relational approach. Soc Sci Med 2007;65:1825–38.
69. Macintyre S, Ellaway A. Neighborhoods and health: an overview. In Kawachi I, Berkman LF (Eds), Neighborhoods and health (p. 20-42). New York: Oxford University Press; 2003.
70. Sampson RJ. Neighborhood-level context and health: lessons from sociology. In Kawachi I, Berkman LF (Eds), Neighborhoods and health (p. 132-46). New York: Oxford University Press; 2003.
71. Kawachi I, Takao S, Subramanian SV (Eds). Global perspectives on social capital and health. New York: Springer; 2013.
72. Grimm KA, Moore LV, Scanlon KS. Access to health food retailers—Unites States, 2011. CDC health disparities and inequalities report—United States, 2013. MMWR 2013;62(3):20–6. Available at http://www.cdc.gov/mmwr/pdf/other/su6203.pdf. Accessed November 11, 2014.
73. Powell LM, Slater S, Chaloupka FJ, et al. Availability of physical activity-related facilities and neighborhood demographic and socioeconomic characteristics: A national study. Am J Public Health 2006;96:1676–80.
74. Maas J, Verheij RA, Groenewegan PP, et al. Green space, urbanity, and health: how strong is the relation? J Epidemiol Community Health 2006;60:587–92.
75. Alcock I, White MP, Wheeler BW, et al. Longitudinal effects on mental health of moving to greener and less green urban areas. Env Sci Technol 2014;8(2):1247–55.
76. Council on Graduate Medical Education tenth report: physician distribution and health care challenges in rural and inner-city areas. Rockville, MD: Public Health Service, Health Resources and Services Administration; 1998. Available at http://www hrsa.gov/advisorycommittees/bhpradvisory/cogme/Reports/tenthreport.pdf. Accessed February 14, 2014.
77. Stafford M, Chandola T, Marmot M. Association between fear of crime and mental health and physical functioning. Am J Public Health 2007;97:2076–81.
78. Clear TR. Impact of incarceration on community public safety and public health. In Greifinger RV, Bick JA, Goldenson J (Eds), Public health behind bars: from prisons to communities (p. 13-24). New York: Springer; 2007. Pp. 13–24.
79. Sum A, Khatiwada I, McLaughlin J, et al. The consequences of dropping out of high school: joblessness and jailing for high school dropouts and the high cost for taxpayers. Boston: Center for Labor Market Studies, Northeastern University; 2009.
80. Bullard RD, Johnson GS, Torres AO. Environmental health and racial equity in the United States: building environmentally just, sustainable, and livable communities. Washington, DC: American Public Health Association; 2011.

81. O'Campo P, Dunn JR (Eds). Rethinking social epidemiology. Dordrecht: Springer; 2011.
82. Office of Disease Prevention and Health Promotion. Website: healthypeople.gov. Available at http://www.healthypeople.gov/2020/topicsobjectives2020/default.aspx. Accessed November 12, 2014.
83. Alameda County Public Health Department. Place matters (factsheet). Available at www.acphd.org/media/114473/placematters_overview.pdf. Accessed November 12, 2014.
84. Bradley EH, Elkins BR, Herrin J, et al. Health and social services expenditures: associations with health outcomes. BMJ Qual Saf 2011;20:826–31.
85. Spence M, Hlatshwayo S. The evolving structure of the American economy and the employment challenge. New York: Council on Foreign Relations; 2011. Available at http://www.cfr.org/industrial-policy/evolving-structure-american-economy-employment-challenge/p24366. Accessed November 12, 2014).
86. Public education finances: 2011 (G11-ASPEF). Washington, DC: U.S. Census Bureau; 2013.
87. Greenwald R, Hedges LV, Laine RD. The effect of school resources on student achievement. Rev Educ Res 1996;66(3):361–96.
88. TED: The economics daily. Washington, DC: U.S. Department of Labor, Bureau of Labor Statistics; 2012. Available at http://www.bls.gov/opub/ted/2012/ted_20120110 htm. Accessed November 12, 2014.
89. Ridgeway CL. Why status matters for inequality. Am Sociol Rev 2014;79(1):1–16.
90. Saez E. Striking it richer: the evolution of top incomes in the United States (updated with 2012 preliminary estimates). Berkeley, CA: UCLA, Berkeley; 2013.
91. Selected measures of household income dispersion: 1967 to 2010. Table A3. Washington, DC: U.S. Census Bureau. Available at https://www.census.gov/hhes/www/income/data/historical/inequality/IE-1.pdf. Accessed May 18, 2015.
92. Massey DS, Denton NA. American apartheid: segregation and the making of the underclass. Cambridge, MA: Harvard University Press; 1993.
93. Wilson JW. The truly disadvantaged: the inner city, the underclass, and public policy. Chicago: University of Chicago Press; 1987.
94. Williams DR, Mohammed SA. Discrimination and racial disparities in health: evidence and needed research. J Behav Med 2009;32(1):20–47.
95. Smedley BD, Stith AY, Nelson AR (Eds). Unequal treatment: confronting racial and ethnic disparities in health care. Washington, DC: National Academies Press; 2003.
96. Gaston GB. African-Americans' perceptions of health care provider cultural competence that promote HIV medical selfcare and antiretroviral medication adherence. AIDS Care 2013;25(9):1159–65. Available at http://www.tandfonline.com/doi/abs/10.1080/09540121.2012.752783#preview. Accessed November 12, 2014.
97. Brotaneka JM, Seeley CE, Flores G. The importance of cultural competency in general pediatrics. Curr Opin Pediatr 2008;20:711–8.
98. Basch CE. Healthier students are better learners: a missing link in school reforms to close the achievement gap. J School Health 2011;81(10):593–598.
99. Taras H, Potts-Datema W. Chronic health conditions may also impact children's development and educational performance. J School Health 2005;75(7):255–66.
100. Currie J. Healthy, wealthy, and wise: socioeconomic status, poor health in childhood, and human capital development. J Econ Lit 2009;47(1):87–122.
101. Gan L, Gong G. Estimating interdependence between health and education in a dynamic model (NBER working paper no. 12830). Cambridge, MA: National Bureau of Economic Research; 2007. Available at http://www nber.org/papers/w12830. Accessed November 12, 2014.
102. Eide RE, Showalter MH. Estimating the relation between health and education: what do we know and what do we need to know? Econ Educ Rev 2011;30(5):778–791.
103. Pridmore P. Impact of health on education access and achievement: a cross-national review of the research evidence. Create Pathways to Access Research Monograph No. 2. London: University of London, Institute of Education; 2007. Available at http://files.eric.ed.gov/fulltext/ED508614.pdf. Accessed November 12, 2014.
104. Denhem SA. Social-emotional competence as support for school readiness: what is it and how do we assess it? Early Educ Dev 2006;17(1):57–89.
105. Williams Shanks TR, Robinson C. Assets, economic opportunity and toxic stress: a framework for understanding child and educational outcomes. Econ Educ Rev 33:154-70.
106. Chapman DA, Scott KG, Stanton-Cahpman TL. Public health approach to the study of mental retardation. Am J Ment Retard 113(2):102-16.
107. bmMayer SE, Jencks C. Growing up in poor neighborhoods: how much does it matter? Science 1989;243(4897):1441–5.
108. Jencks C, Mayer SE. The social consequences of growing up in a poor neighborhood. In Inner city poverty in the United States. Washington, DC: National Academies Press; 1990.
109. Barnett WS, Belfield CR. Early childhood development and social mobility. Future Child 2006;16(2):73–98.
110. Avchen RN, Scott G, Mason CA. Birth weight and school-age disabilities: a population-based study. Am J Epidemiol 2001;154(10):895–901.
111. Heckman JJ. Skill formation and the economics of investing in disadvantaged children. Science 2006;312:1900–2.
112. Weisleder A, Fernald A. Early language experience strengthens processing and builds vocabulary. Psychol Sci 2013;24(11):2143–52.

113. Shonkoff JP, Garner AS. The lifelong effects of early childhood toxic stress. Pediatrics 2012;129(1):e232–46.
114. Sandstrom H., Huerta S. The negative effects of instability on child development: a research synthesis. Low-income working families discussion paper 3. Washington, DC: Urban Institute; 2013.
115. Burgard S, Seefeldt K, Zelner S. Housing instability and health: findings from the Michigan Recession and Recovery Study. Population Studies Center research report 12-749. Ann Arbor, MI: University of Michigan Institute for Social Research; 2012. Available at http://www.psc.isr.umich.edu/pubs/pdf/rr12-749.pdf. Accessed November 12, 2014.
116. Duncan GJ, Brooks-Gunn J, Klebanov PK. Economic deprivation and early childhood development. Child Dev 1994;65:296–318.
117. Case A, Lubotsky D, Paxson C. Economic status and health in childhood: the origins of the gradient. Am Econ Rev 2002;92:1308–34.
118. Brooks-Gunn J, Duncan GJ. The effects of poverty on children. Future Child 1997;7:55–71.
119. Guo G. The timing of influences of cumulative poverty on children's cognitive ability and achievement. Soc Forces 1998;77(1):257–87.
120. Evans GW, Kim P. Multiple risk exposure as a potential explanatory mechanism for the socioeconomic status-health gradient. Ann NY Acad Sci 2010;186(1):174–89.
121. Evans GW, Brooks-Gunn J, Klebanov PK. Stressing out the poor: chronic physiological stress and the income-achievement gap. Community Investments 2011;23(2):22–27.
122. Shonkoff JP, Phillips DA (Eds). From neurons to neighborhoods: the science of early child development. Washington, DC: National Academies Press; 2000.
123. Gottesman II, Hanson DR. Human development: biological and genetic processes. Annu Rev Psychol 2005;56:263-86.
124. Hackman DA, Farah MJ, Meaney MJ. Socioeconomic status and the brain: mechanistic insights from human and animal research. Nat Rev Neurosci 2010;11: 651-9.
125. McEwen BS. Brain on stress: how the social environment gets under the skin. Proc Natl Acad Sci 2012;109(2):17180-5.
126. Zhang TY, Meaney MJ. Epigenetics and the environmental regulation of the genome and its function. Ann Rev Psychol 2010;61:439-66.
127. Felitti VJ, Anda RF, Nordenberg D, et al. Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults: the Adverse Childhood Experiences (ACE) study. Am J Prev Med 1998;14(4):245–58.
128. Injury prevention and control: Division of Violence Prevention (Web site). Available at http://www.cdc.gov/violenceprevention/acestudy/index.html. Accessed November 13, 2014.
129. Shonkoff JP, Boyce WT, McEwen BS. Neuroscience, molecular biology, and the childhood roots of health disparities: building a new framework for health promotion and disease prevention. JAMA 2009;301(21): 2252-9.
130. Mistry KB, Minkovitz CS, Riley AW, et al. A new framework for childhood health promotion: the role of policies and programs in building capacity and foundations of early childhood health. Am J Public Health 2012;102:1688-96.
131. Woolf SH, Braveman P, Evans BF. The health implications of reduced food stamp eligibility: a rapid-cycle background report. Richmond, VA: Center on Human Needs, Virginia Commonwealth University; 2013. Available at http://societyhealth.vcu.edu/media/society-health/VCUSNAPIncomeHealth.pdf. Accessed November 17, 2014.
132. Selby JV, Beal AC, Frank L. The Patient-Centered Outcomes Research Institute (PCORI) national priorities for research and initial research agenda. JAMA 2012 Apr 18;307(15):1583-4.
133. Hickam D, Totten A, Berg A, et al. (Eds). The PCORI methodology report. PCORI Methodology Committee. November 2013. http://www.pcori.org/assets/2013/11/PCORI-Methodology-Report.pdf. Accessed November 17, 2014.
134. Entwistle VA, Renfrew MJ, Yearley S, et al. Lay perspectives: advantages for health research. Br Med J 1998;316:(7129):463-6.
135. Winder E. The value+ toolkit. European Patient's Forum. Available at http://www.eu-patient.eu/globalassets/projects/valueplus/value-toolkit.pdf. Accessed November 17, 2014.
136. Caron-Flinterman JF, Broerse JEW, Bunders JFG. The experiential knowledge of patients: a new resource for biomedical research? Soc Sci Med 2005;60:2575–84.
137. Goodare H, Lockwood S, HG chairs BREAST UK. Involving patients in clinical research: improves the quality of research. Br Med J 1999;319:724-5.
138. Domecq Garces JP. Eliciting patient perspective in patient-centered outcomes research: a meta narrative systematic review. A report prepared for the Patient-Centered Outcomes Research Institute. Rochester, NY: Mayo Clinic; 2012.
139. Popay J, Williams G. Public health research and lay knowledge. Soc Sci Med 1996;42(5):759-68. Page 760.
140. Simonds VW, Wallerstein N, Duran B, et al. Community-based participatory research: its role in future cancer research and public health practice. Prev Chronic Dis 2013;10:e78. DOI: http://dx.doi.org/10.5888/pcd10.120205. Accessed March 11, 2014.
141. Wallerstein NB, Duran B. Using community-based participatory research to address health disparities. Health Promot Pract 2006;7(3):312-23.
142. Israel BA, Schulz AJ, Parker EA, et al. Community-based participatory research: policy recommendations for promoting a partnership approach in health research. Educ Health 2001;14(2):182-97.
143. Tandon SD, Phillips K, Bordeaux BC, et al. A vision for progress in community health partnerships. Prog Community Health Partnersh 2007; 1(1):11-30.

144. Corburn J. Community knowledge in environmental health science: co-producing policy expertise. Environ Sci Policy 2007;10:150–61. Page 151.
145. Tonglet R, Mudosa M, Badashonderana M, et al. The causal model approach to nutritional problems: an effective tool for research and action at the local level. Bull WHO 1992;70(6):715-23.
146. Beghin I, Cap M, Dujardin B. A guide to nutritional assessment. Geneva: World Health Organization; 1988. http://apps.who.int/iris/handle/10665/37419. Accessed November 17, 2014.
147. Lefèvre P, de Suremain C-E, Rubín de Celis E, et al. (2004). Combining causal model and focus group discussions experiences learned from a socio-anthropological research on the differing perceptions of caretakers and health professionals on children's health (Bolivia/Peru). Qual Rep 2004;9(1):1-17. Available at http://www.nova.edu/ssss/QR/QR9-1/suremain.pdf. Accessed November 17, 2014.
148. Zoellner J, Motley M, Wilkinson ME, et al. Engaging the Dan River Region to reduce obesity: application of the comprehensive participatory planning and evaluation process. Fam Community Health 2012;35(1):44-56.
149. Lefèvre P, Kolsteren P, De Wael M, et al. Comprehensive participatory planning and evaluation. Antwerp, Belgium: Tropical Medicine; 2000. Available at www.ifad.org/pub/bsf/cppe/cppe.pdf. Accessed November 17, 2014.
150. Hovmand P. Community based system dynamics. New York: Springer; 2014.
151. Stave KA. Using system dynamics to improve public participation in environmental decisions. System Dynamics Rev 2002;18(2):139–67.
152. Adler NE, Stewart J. Health disparities across the lifespan: meaning, methods, and mechanisms. Ann NY Acad Sci 2010;1186:5-23.
153. Zimmerman E, Bea C, Thomas M, et al. Community-based participatory research: involving residents in qualitative coding. Presentation at the 2013 Eastern Sociological Society annual meeting, Boston, MA.
154. Smith T. Project engages East End residents in discussions on improving health. Richmond Times-Dispatch January 27, 2014. Available at http://www.timesdispatch.com/news/local/project-engages-east-end-residents-in-discussions-on-improving-health/article_4609781a-e07f-58aa-8083-3cd5292e901d html. Accessed November 17, 2014.
155. Smith T. Richmond's East End has stark health challenges. Richmond Times-Dispatch, January 27, 2013. Available at http://www.timesdispatch.com/entertainment-life/health/richmond-s-east-end-has-stark-health-challenges/article_dcaba3b1-be4f-5b44-8361-d03527529112 html. Accessed November 17, 2014.
156. Wang CL. Predetermined variable. In: Lewis-Beck MS, Bryman AE, Futing Liao T (Eds), The SAGE encyclopedia of social science research methods, vol. 1.Thousand Oaks, CA: Sage Publications; 2004. Page 849.
157. Olson WK. Mediating variable. In Lewis-Beck MS, Bryman A, Futing Liao T (Eds). The Sage encyclopedia of social science research methods. Thousand Oaks, CA: Sage Publications; 2004.
158. Wimer C, Bouffard SM, Caronongan P, et al. What are kids getting into these days: demographic differences in youth out-of-school time participation. Harvard Family Research Project. Cambridge, MA: Harvard Graduate School of Education; 2006. Available at http://www.hfrp.org/content/download/1074/48577/file/full_report_demographic_diff.pdf. Accessed November 17, 2014.
159. Mahoney JL, Larson RW, Eccles JS, et al. Organized activities as development contexts for children and adolescents. In Mahoney JL, Larson RW, Eccles JS (Eds), Organized activities as contexts of development (pp. 3-22). Mahwah, NJ: Lawrence Erlbaum Associates; 2005.
160. Maslow GR, Chung RJ. Systematic review of positive youth development programs for adolescents with chronic illness. Pediatrics 2013;131:e1605-18.
161. Gavin LE, Catalano RF, David-Ferdon C, et al. A review of youth development programs that promote adolescent sexual and reproductive health. J Adolesc Health 2010;46;S75-91.
162. Kwan M, Bobko S, Faulkner G, et al. Sport participation and alcohol and illicit drug use in adolescents and young adults: a systematic review of longitudinal studies. Addict Behav 2014;39:497-506.
163. Brouskeli V, Loumakou M. Materialism, stress and health behaviors among future educators. J Educ Train Stud 2014;2(2):145-50.
164. Kasser T. The high price of materialism. Cambridge, MA: MIT Press; 2002.
165. Auerbach RP, McWhinnie CM, Goldfinger M, et al. The cost of materialism in a collectivistic culture: predicting risky behavior engagement in Chinese adolescents. J Clin Child Adolesc Psychol 2010;39(1):117–27.
166. Smith W, Christensen AJ. Hostility, health, and social contexts. In Friedman HS (Ed), Hostility, coping, & health (pp. 33-48). Washington, DC: American Psychological Association; 1992.
167. 2012 National healthcare disparities report. Rockville, MD: Agency for Healthcare Research and Quality; 2013.
168. Szymanski K, Sapanski L, Conway F. Trauma and ADHD – association or diagnostic confusion? A clinical perspective. J Infant Child Adolesc Psychother 2011;10(1):51-9.
169. Shaw M, Dorling D, Smith GD. Poverty, social exclusion, and minorities. In Marmot M, Wilkinson RG (Eds), Social determinants of health. Oxford: Oxford University Press; 1999. Pages 211-39.
170. Lofters A, O'Campo P. Differences that matter. In O'Campo P, Dunn JR (Eds), Rethinking social epidemiology. Dordrecht: Springer; 2011. Pages 93-109.
171. Kim AE, Kumanyika S, Shive D, et al. Coverage and framing of racial and ethnic health disparities in U.S. newspapers, 1996–2005. Am J Public Health 2010;100:S224–31.
172. Braveman PA, Egerter SA, Woolf SH, et al. When do we know enough to recommend action on the social determinants of health? Am J Prev Med 2011;40(1 Suppl 1):S58-66.

173. Maglio PP, Sepulveda MJ, Mabry PL. Mainstreaming modeling and simulation to accelerate public health innovation. Am J Public Health 2014;104:1181-6.
174. Woolf SH, Purnell JQ, Simon S, et al. Research translation in population health: the role of engagement and communication with decision makers, stakeholders, and the community. Ann Rev Public Health, in press.
175. Mowat D, Chambers C. Producing more relevant evidence: applying a social epidemiology research agenda to public health practice. In O'Campo P, Dunn JR (Eds), Rethinking social epidemiology. Dordrecht: Springer; 2011. Pages 305-25.

Emily Zimmerman, PhD, is an Associate Professor in the Department of Family Medicine and Population Health, Division of Epidemiology, Virginia Commonwealth University and a senior researcher at VCU's Center on Society and Health. Her current work focuses on research about the social determinants of health and projects to raise awareness about social determinants among disparate audiences. Dr. Zimmerman helped found Engaging Richmond, a community-university partnership, which has worked to link community priorities with research and action since 2011. Through her work on stakeholder engagement, she developed a methodology for involving diverse stakeholders in the process of formulating and prioritizing health research questions. The methodology is currently being implemented and evaluated in two demonstration sites in Virginia.

Steven H. Woolf, MD, MPH, is Director of the Virginia Commonwealth University Center on Society and Health and Professor of Family Medicine and Population Health at VCU. He is board certified in family medicine and in preventive medicine and public health. Dr. Woolf has focused on promoting the most effective health care services and on advocating the importance of health promotion and disease prevention and the need to address the social determinants of health. He has emphasized outreach to policymakers, the public, and the media to raise awareness about the factors outside of health care that shape health outcomes. Dr. Woolf was elected to the Institute of Medicine in 2001.

Amber Haley holds masters-level training in epidemiology and serves as research faculty in the Virginia Commonwealth University Department of Family Medicine and Population Health, Division of Epidemiology and as the Associate Director of Community Engagement for the Center for Clinical and Translational Research. Her research focuses on the development of community-driven research agendas, strategic partnership to address community identified priorities, and the translation of social determinants of health research into useful evidence for various stakeholders.

Aligning Medical Education with the Nation's Health Priorities: Innovations in Physician Training in Behavioral and Social Sciences

Jason M. Satterfield and Patricia A. Carney

Abstract

Although some see today as the "golden age" of medical education, the impressive advances in educational science and pedagogy have not created a physician workforce suited to meet the needs of society. Medical school admissions still do not reflect the diversity of patients served, and specialty disciplines (e.g., radiology, dermatology) are over-subscribed, while primary care is underselected by medical graduates. Despite policy changes such as the prevention mandate of the Affordable Care Act, medical school curricula spend comparatively little time on behavioral science and often fail to adequately prepare learners for practicing health care in complex, interprofessional teams. Recent movements to incorporate more workplace learning, quality improvement, and systems science hold promise for future trainees, as do advances in behavioral science and medical education. With the addition of a substantial behavioral science component to the Medical College Admissions Test (MCAT 2015), trainees are likely to arrive with a higher level of preparation and readiness to quickly master and clinically translate social and behavioral science constructs if given the right learning environment and supportive institutional cultures. A more behaviorally and socially sophisticated graduate may be better prepared to deliver high quality, culturally competent care while maintaining a sense of professional and personal resilience in the face of an evolving and challenged health care system. A critical re-examination of each medical school's "social contract" to meet societal needs paired with the impressive advances in medical science hold great promise to enrich and improve the physician workforce of the future.

Introduction

Undergraduate medical education has been undergoing a much needed transformation to better prepare learners to provide care in complex health systems and to address anticipated workforce needs. Currently, 141 allopathic and 30 osteopathic U.S. medical schools are training over 82,000 students; more than 18,000 new physicians graduated in 2013. These numbers will steadily rise as medical school class sizes grow and new medical schools are created.[1] Curricular structure and pedagogy have also been evolving with more schools challenging the classic Flexnerian structure, where the first 2 years are spent learning basic science and the last two in clinical practice apprenticeships.[2,3] Didactic lectures are being replaced with "learner-centered small groups" and "flipped classrooms" where didactic content is presented online, and classroom time is focused on interactive case-based exercises.[4] Technology has further transformed medical education with the

introduction of Massive Open Online Courses (MOOCs), iBooks, simulations, and other interactive, technology-based learning tools.[5,6] In fact, so much has improved from a pedagogical perspective that our current era has been called the "golden age of medical education."[7]

Despite these advances, our vast, expensive, and complex medical education system has failed to produce a physician workforce capable of addressing the leading determinants of health or meeting the health care needs of our communities. The current physician workforce is too small, too specialized, unprepared to work in interprofessional teams, and ill equipped to manage chronic diseases or robustly promote prevention.[7-10]

Despite spending more on health care than any other nation, the United States ranks 27th in life expectancy and 25th in infant mortality.[11,12] As a recent Institute of Medicine (IOM)/National Research Council report showed, obesity, diabetes, sexually transmitted infections, teen pregnancy, drug use, and motor vehicle accidents contribute to both chronic and acute conditions and are highly influenced by complex social, behavioral, and political factors. In addition to shortcomings in medical education, lack of reliable access to health care, under-emphasis on primary care and prevention, income inequality, and policy shortcomings all contribute to a failure to protect the health of the public.[13]

What is currently needed of undergraduate medical education is to create a 21st century physician workforce that is able to improve the health of patients and communities.[7,9,14] Medical schools should additionally refocus research priorities on areas that will alter public health and patient outcomes, rather than focusing primarily on areas that have little if any utility for the majority of the population. The mission of medical education should represent a balance between workforce preparation and research suited to the needs of our communities and the Nation overall.

In this chapter, we examine challenges and recent innovations that are directly influencing the physician workforce. We address three critical questions regarding the current state of medical education: (1) Are we training the right people (workforce composition)? (2) Are we teaching the right kind of content and competencies (workforce knowledge and skills)? (3) Are we teaching using the best techniques (workforce training methods)? We propose that behavioral and social sciences (BSS) offer important insights into both the causes and potential solutions for each of these questions. Notable educational and key BSS innovations are described in ways that highlight important transformations yet to come and point to next steps in training the physicians our communities need and deserve.

Question 1: Are We Training the Right People?

Public investment in physician training should motivate medical schools to be held more accountable to societal needs.[9,15-18] However, while the number of medical schools and medical student graduates has risen considerably, theoretically to meet potential workforce deficiencies, this expansion does not appear to match the evolving health needs of the U.S. population. More, not fewer, primary care physicians are needed to provide care in underserved areas and to newly insured populations, and the characteristics of the physician workforce should more closely match the diversity of the patients and communities being served.[19,20]

To better characterize this issue, Mullan and colleagues[9] undertook a qualitative study designed to create a "Social Mission Score" for each U.S. medical school using the percent of graduates going

into primary care, percent working in health professional shortage areas, and percent of minority students enrolled. Overall, scores indicated disappointing performance, particularly from many esteemed and well-resourced medical schools. Public medical schools outperformed private institutions, with the Morehouse School of Medicine receiving top honors. Surprisingly, the amount of National Institutes of Health (NIH) research dollars received was inversely correlated with social mission scores. While notable exceptions were found, where both research and social mission were impressively achieved (e.g. University of Washington, University of Minnesota), transformation in medical schools' missions is needed.

A closer look at shifting population demographics highlights the changes needed within our physician workforce. In 2006, racial and ethnic minorities accounted for 28 percent of the U.S. population, yet only 15 percent of medical students and 8 percent of practicing physicians represented these ethnic/minority groups.[21,22] While the number of underrepresented minorities who were first time enrollees in U.S. medical schools increased by 20 percent from 2006 to 2012, overall enrollment has grown by only 12.4 percent, representing modest gains in minority representation.[1] A disproportionate number of Caucasian and Asian medical students continue to be enrolled in medical schools, while the number of African American male students has declined.[1]

Racial and ethnic diversity are not the only areas of concern. In 2010, 12.9 percent of the population was born outside the United States,[23] and 19.4 percent of the population spoke a language other than English at home.[24] Currently, there are no national data available on physicians' non-English language skills.[25] Although mandated, professional medical interpreters often are either underutilized or not available.[26, 27] While matching patient-physician demographics and language may not ensure quality medical care, both dyadic and population discordances have implications for impactful therapeutic relationships, communication, and patient motivation to gain access to care.[28,29] Moreover, demographic discordances provide a compelling argument for medical student learners to understand diversity, race, gender, and other sociocultural factors in addition to core skills in communication, spoken language, and literacy.

Increasing trainee diversity benefits patients and enhances trainees' preparedness to work with minority populations.[30] "Pipeline programs" designed to attract more underrepresented students into medical school or training for other health professions have become increasingly common. For example, the University of California, Los Angeles (UCLA) PREP program offers 7 weeks of free training to better prepare disadvantaged students for medical or dental school, including MCAT or DAT preparation, coaching on application development, mentorship in career choice, and clinical preceptorships. Currently, 81 percent of PREP participants are admitted to a health professional school. However, in a State with 46.7 percent racial and/or ethnic minorities, UCLA typically enrolls fewer than 30 percent under-represented minority medical students.[31,32] A better understanding is needed regarding existing social, developmental, and environmental factors that influence admissions and enrollment to medical school so that diversity recruitment efforts can be improved.[33-35] Similarly, recruitment and retention of minority medical school faculty, which likely plays a role in minority student applications, has been challenging with programs to recruit minority faculty having limited effectiveness.[36,37] The percentage of underrepresented minority faculty increased from 6.8 percent (95 percent CI, 6.7 percent-7.0 percent) in 2000 to only 8.0 percent (95 percent CI, 7.8 percent-8.2 percent) in 2010. Only 36 of 124 medical schools (29 percent) had a minority faculty development program in place in 2010.[36]

Innovative programs, such as the Association of American Medical Colleges' (AAMC's) Holistic Review Project,[38] which started in 2007, provides a flexible process of assessing applicants' capabilities using a balanced review of experiences, attributes, academic metrics, and how the applicant might contribute value as a trainee and physician. The project provides admissions-related tools and resources that medical schools can use to create and sustain diversity, while staying abreast of larger national and political circumstances that influence minority recruitment (e.g. the U.S. Supreme Court's Fisher decision[a]).

In addition to lacking adequate diversity, medical schools are currently failing to produce enough primary care providers (PCPs), especially those with a desire to work in underserved areas or with stigmatized populations. The National Association of Community Health Centers estimates that one in five individuals currently lack adequate primary care because of physician shortages, which are severe in rural communities.[39-41] Federally supported community health centers similarly report recruitment and retention failures, despite such incentives as loan repayment programs.[42,43] Future estimates predict a nationwide shortage of 40,000 PCPs by 2025.[44]

Overall, only 26 percent of medical graduates report they intend to work with underserved populations, and even fewer intend to work in health professional shortage areas (HPSAs).[41] In a study of graduating internal medical residents in California between 2000 and 2003, only 4.7 percent intended to practice in HPSAs, 15.1 percent intended to practice in underserved areas, and 25.2 percent intended to practice in public hospitals.[45] Among those who currently work in underserved areas, minority physicians are disproportionately represented.[46,47]

Although the AAMC encouraged medical schools to increase their class sizes by up to 30 percent,[48] the proportion of students who go into primary care ranges widely; and without additional residency training slots, it is unlikely that increasing the number of medical students will solve the primary care shortage. In addition, the current generation of medical students are more likely to select "lifestyle" specialties such as dermatology, ophthalmology, anesthesia, and radiology, due to more manageable time demands, workload, and more attractive compensation compared to careers in primary care.[49-51] Moreover, medicine has developed specialties that focus more on stage of care (e.g., hospital medicine, palliative care, emergency medicine) rather than comprehensive primary care.[7] Lastly, other health professionals, such as nurse practitioners and physician assistants, could offer substantial support in primary care delivery even in traditionally underserved areas.[52-55] Currently, most medical students receive limited training in how to work as part of interdisciplinary teams, and physicians have been slow to embrace less physician-centric models of care, despite the present well-known shortcomings of today's U.S health care system.[56,57]

Longitudinal integrated clinical clerkships represent a new and rapidly growing approach to improve continuity of care while promoting greater patient and learner-centeredness.[58] In these models, students typically receive a greater proportion of outpatient training and develop longer relationships with faculty role models who are often primary care providers.[59] With greater patient continuity and enriched student-patient relationships, students may be more likely to experience the rewards of primary care and the importance of caring for those most in need.

[a] See syllabus for Fisher vs. University of Texas at Austin, et al. Available at http://www.supremecourt.gov/opinions/12pdf/11-345_l5gm.pdf.

Question 2: Are We Teaching the Right Content and Competencies?

Both the health care needs of the public and systems of care delivery have shifted dramatically over the past century. Medicine has moved from a model of independent private practice focused on the treatment of acute and/or infectious disease to large, team-based organizations heavily involved in the management of chronic diseases.[15] Despite these changes, the U.S. medical education enterprise still focuses primarily on the creation of the "personally expert sovereign physician."[7]

Medical educators have been placed in a seemingly impossible situation. Medical curricula are expected to include foundational material as well as recent scientific advances in biomedicine, while still expanding learners' skill sets to include interprofessional communication, teamwork, systems-based care, quality improvement, population health, informatics, and meta-cognitive skills. Graduates are now expected to become expert clinicians, leaders, and researchers without expanding the traditional 4 years required for the MD/DO degree. We argue that advances in cognitive science (e.g. how we learn), behavioral psychology (e.g. motivation and behavior change), and social science (e.g. understanding communication, group dynamics, and social determinants of health) offer efficient and effective ways to enrich medical curricula and improve clinical practice.

To address the dramatic increase in curricular content demands, educators must restrict material to fundamentals in knowledge and skills that every medical student needs to know to successfully enter residency training in their chosen field. The Accreditation Council for Graduate Medical Education (ACGME) offers guidance in organizing the fundamentals into six broad competency areas: medical knowledge, patient care, interpersonal and communication skills, professionalism, practice-based learning and improvement, and systems-based practice.[60,61] Although initially designed for medical residents, many medical schools have incorporated these competencies into undergraduate medical education. While medical knowledge and patient care fit traditional medical school content, the remaining four competencies have proven challenging, especially in the domains of cultural competence, interprofessional communication and teamwork, and systems-based practice.[62-64] Fortunately, a number of medical schools are deeply engaged in curriculum revisions, and a number of structural, training innovations are being tested (e.g. the American Medical Association [AMA]-supported "Accelerating Change in Medical Education" project). Moreover, recent curricular content guides and development tools are readily available in the foundational sciences.[65-67] While a number of stakeholder issues and competing demands must be considered when selecting curricular content, we suggest that training physicians to improve individual and population health should be paramount. Here, we offer three examples of how medical school content is being improved.

Example 1: Behavioral and Social Sciences (BSS)

Approximately 50 percent of U.S. morbidity and mortality has been directly linked to behavioral and social factors,[68-70] yet BSS content in medical schools remains widely variable and often underrepresented.[71,72] In 2004, the IOM released a report with specific recommendations for the revision of undergraduate medical education designed to address social and behavioral factors affecting the health of the Nation.[72] This report recommended

that medical schools offer a 4-year, integrated BSS curriculum addressing six domains: mind–body interactions in health and disease; patient behavior; the physician's role and behavior; physician–patient interactions; social and cultural issues in health care, and health policy and economics. The report identified 26 BSS topics (Table 1) and called for Federal support to enrich BSS curricula. This content list was later supported by faculty surveys, expert opinion, and a comprehensive textbook review[67] and used to construct a BSS content "matrix" to be used by curriculum developers and front-line educators alike.[65]

Table 1. IOM BSS Report (2004): Recommended content areas

Domain	High Priority Topics	Medium Priority Topics
Mind-Body Interactions in Health and Disease	• Biological mediators between psychological and social factors and health • Psychological, social, and behavioral factors in chronic disease • Psychological and social aspects of human development that influence disease and illness • Psychosocial aspects of pain	• Psychosocial, biological, and management issues in somatization • Interaction among illness, family dynamics, and culture
Patient Behavior	• Health risk behaviors • Principles of behavior change • Impact of psychosocial stressors and psychiatric disorders on manifestations of other illnesses and health behavior	
Physician Role and Behavior	• Ethical guidelines for professional behavior • Personal values, attitudes, and biases as they influence patient care	• Physician well-being • Social accountability and responsibility • Work in health care teams and organizations • Use of and linkage with community resources to enhance patient care
Physician-Patient Interactions	• Basic communication skills • Complex communication skills	• Context of patient's social and economic situation, capacity for self-care, and ability to participate in shared decisionmaking • Management of difficult or problematic physician-patient interactions
Social and Cultural Issues in Health Care	• Impact of social inequalities in health care and the social factors that are determinants of health outcomes • Cultural competency	• Role of complementary and alternative medicine
Health Policy and Economics	• Overview of U.S. health care system • Economic incentives affecting patients' health-related behaviors • Costs, cost-effectiveness, and physician responses to financial incentives	• Variations in care

In response to the IOM report,[72] the National Institutes of Health (NIH) and the Office of Behavioral and Social Science Research awarded K07 curriculum development and leadership grants to nine medical schools in 2005 to develop, pilot test, and disseminate BSS curricula. This "BSS Consortium for Medical Education" trained approximately 6,100 medical students and published 135 BSS papers.[73] The BSS Consortium was subsequently expanded to 16 schools and broadened its purview to include interprofessional training, implementation science, and training for medical residents and faculty with the potential of reaching more than 12,000 learners.[74] Table 2 lists both the original K07 schools and the expanded consortium schools, along with examples of educational innovations. A number of other medical schools have published their own models for successfully integrating and improving BSS content.[16,71,75,76] Specific examples of BSS content and competency innovations are provided in Table 2.

Table 2. Participating medical schools in the BSS consortium for medical education

K07 & R25 Partnered Medical Schools	Principal Investigators	Behavioral and Social Science Innovations in Medical Education
Albert Einstein College of Medicine Warren Alpert Medical School of Brown University	Paul R. Marantz, MD, MPH; Professor of Clinical Epidemiology & Population Health; Associate Dean for Clinical Research Education Phil Gruppuso, MD, Professor of Pediatrics and Paul George, MD, Assistant Professor of Family Medicine	• Establish a curriculum development collegium between Einstein and Brown through a series of meetings alternating between institutions. • Develop and implement a comprehensive, 4-year curriculum in population health. • Enhance the existing curriculum in communication skills and professionalism at both schools, consistent with the 2004 IOM report on behavioral and social science education. • Foster health-related scholarship in the behavioral and social sciences.
Baylor College of Medicine Texas A&M Health Science Center	Beth Nelson, MD, Associate Professor of Medicine, Senior Associate Dean of Medical Education and Anne Gill, DrPH, MS, RN, Associate Professor of Pediatrics and Medical Ethics, Department of Pediatrics, Section of Academic General Pediatrics Lori Graham, PhD, Assistant Professor, Internal Medicine and Courtney West, PhD, Assistant Professor, Humanities in Medicine	• Enhance curricular activities to transform core student-teacher relationships through ongoing evaluation and quality improvement. • Advance the transformation of the hidden curriculum regarding relationship-centered care and communication in clinical settings through integration of interprofessional education (IPE) by transforming relationships between medical students and other health professionals. • Expand integration of the principles and values of relationship-centered care into both classroom/community-based and clinical/simulation curricular models of IPE. • Promote and disseminate relationship-centered educational development and research, regionally and nationally. • Continue collaborations and national dissemination efforts.
Columbia University College of Physicians and Surgeons Weill- Cornell Medical College	Rita Charon, MD, PhD Professor of Clinical Medicine Susan Ball, MD, MPH, MS, Assistant Director, Center for Special Studies at New York Presbyterian Hospital	• Increase the amount of teaching offered to students on topics that include the doctor-patient relationship, ethics, professionalism, understanding patients' lived experiences, and how to reflect on their own practice. • Increase the teaching skill of physicians and other medical school faculty in an intensive Faculty Scholarship Core. • Examine the institutional culture of the teaching hospital where students train, where Columbia and Cornell share clinical placements at New York-Presbyterian Hospital.

Table 2. Participating medical schools in the BSS consortium for medical education (continued)

K07 & R25 Partnered Medical Schools	Principal Investigators	Behavioral and Social Science Innovations in Medical Education
David Geffen School of Medicine at the University of California, Los Angeles University of California, San Diego	Margaret Stuber, MD, Assistant Dean of Student Affairs for Career Development and Well-Being; Jane and Marc Nathanson Professor of Psychiatry Karen Garmen, EdD, Assistant Clinical Professor and Shawn Harrity, MD, Primary Care Physician, Clinical Professor of Medicine	• Improve the teaching of the behavioral and social sciences (BSS) by creating 11 problem-based learning (PBL) or team-based learning (TBL) cases, developing five online interactive learning programs to teach faculty about active learning formats, developing an electronic lecture series about rapidly evolving topics in behavioral and social sciences, and coaching faculty to facilitate behavioral and social science learning. • Increase the number of students conducting or participating in social and behavioral science research by providing six scholarships for students presenting at national conferences, sponsoring a course in mindful awareness, and developing six summer research opportunities in BSS, including work with integrative medicine. • Improve attitudes towards interprofessional education among faculty and future health care providers by collaborating with other health professional schools to create interactive learning using simulation, small group reflective learning, and PBL. • Develop students' knowledge and attitudes about integrative medicine by interacting with practitioners and students in acupuncture, chiropractic/massage, and energy therapies.
Indiana University School of Medicine University of Missouri School of Medicine	Debra K. Litzelman, MD, MA Associate Dean for Research in Medical Education; Professor of Medicine Linda Headrick, MD, Professor of Medicine, Senior Associate Dean for Education	• Enhance medical student BSS curriculum with learning experiences designed to enable the development of a patient/relationship-centered approach to health care that facilitates the effective application of BSS knowledge in clinical interactions. • Enable a range of faculty development opportunities created and implemented to ensure faculty at both institutions have the capacity to model, facilitate student BSS reflections, professional development, and mentorship. • Create a program for institutional review and quality improvement specific to BSS aspects of patient care.
Oregon Health & Science University University of Texas Health Sciences Center at San Antonio	Fran Biogioli, MD, Director, Predoctoral Education; Associate Professor of Family Medicine Jim Tysinger, PhD, School of Medicine, Family and Community Medicine	• Develop a comprehensive BSS EHR/HIT curriculum that includes a simulated EHR (SIM-EHR) training database containing virtual patient cases adaptable to different specialties and levels of medical education (medical student, resident, practicing faculty). • Selectively enhance the overall BSS curriculum at both medical schools. • Evaluate the effectiveness of both the general BSS curriculum and BSS EHR/HIT curriculum. • Export BSS EHR/HIT curricular innovations and research findings to a larger national audience. • Foster BSS careers by recruiting faculty for the interdisciplinary development, teaching, evaluation, and dissemination of curriculum resulting from the BSS EHR/HIT project.

Table 2. Participating medical schools in the BSS consortium for medical education (continued)

K07 & R25 Partnered Medical Schools	Principal Investigators	Behavioral and Social Science Innovations in Medical Education
University of California School of Medicine, San Francisco Stanford University School of Medicine	Jason M. Satterfield, PhD, Director, Behavioral Medicine; Professor of Clinical Medicine Stephanie Harman, MD, Clinical Assistant Professor, Medicine and Rebecca Blankenburg, MD, MPH, Clinical Assistant Professor, Pediatrics	• Promote the translation of core BSS knowledge into clinical skills taught to all third-year UCSF and Stanford medical students by improving existing BSS content, adding meta-cognitive skills instruction, and better preparing medical residents and faculty who train third-year medical students. • Improve resident and faculty BSS teaching through (1) BSS Teaching Consultation program, (2) BSS Clinical Rounds, and (3) BSS Teaching Improvement Program/Teaching Observation Program (BSS TIP/TOP). • Explore new and innovative BSS assessment tools, such as ward ethnographies, calibrated peer review, standardized patients, and portfolios. • Pilot new curricula around patient engagement and shared decisionmaking taught in a learner-centered, inpatient environment.
University of North Carolina School of Medicine Wake Forest University School of Medicine	Barry Saunders, MD, MA, PhD, Associate Professor of Social Medicine and Sue Estroff, PhD, Professor of Social Medicine Nancy King, JD, Professor, Department of Social Sciences and Health Policy, Division of Public Health Sciences	• Expand faculty development opportunities to prepare our faculty to teach in new venues of social and behavioral science. • Draw upon Wake Forest University School of Medicine's expertise in small group clinical teaching, while providing them with expertise in social sciences.
University of Wisconsin School of Medicine and Public Health [K07 only, 2005-2010]	Patrick McBride, MD, MPH, Professor, Departments of Medicine and Family Medicine; Associate Dean for Students	• Integrate new content on cultural competency into core courses across the continuum of medical school training. • Develop a comprehensive plan to foster the expression of professional values and behaviors by medical students across all years of education. • Integrate new public health content into the first-, second-, and fourth-year medical student curriculum within existing courses. • Expand summer research opportunities in behavioral and social sciences. • Assess the institutional impact of curricular changes.

Note: BSS = behavioral and social sciences; EHR = electronic health record; HIT = health information technology; IOM = Institute of Medicine; IPE – interprofessional education; PBL = problem-based learning; SIM = simulated; TBL = team-based learning; TIP/TOP = teaching consultation program

The Baylor College of Medicine implemented a "Best Intentions Workshop," a session for third year medical students, delivered to more than 400 students over 3 years. This workshop session used the Implicit Associations Test (IAT)[77] as a trigger for small group-based reflection about physician implicit bias (how doctors may unknowingly perpetuate health care disparities), as well as strategies for managing potential biases. Students with lower self-awareness about personal bias prior to the session showed significant increases in self-awareness after the session ($p < 0.001$). Students also showed significant ($p<0.01$) increases in agreement that the IAT was effective in increasing awareness of

personal bias and for generating small group discussion about it. Sixty-seven percent of students undertaking this program identified alternate strategies for managing bias toward patients at post-session, and distribution of the strategies changed significantly from pre-session to post-session as well ($p < 0.01$).[78,79]

The University of Wisconsin implemented a required four-semester Cultural Perspectives Curriculum with approximately 8 hours of large group and 4 hours of small group instruction across four semesters of the first and second year. Topics included health and illness perspectives of patients and physicians, culture of medicine, and health literacy. Assignments helped students to understand their own health and illness perspectives and biases. In addition, students completed reflective writing exercises, using an electronic portfolio, on an artifact of cultural competence. A required third year Core Day "Skills to Impact Health Disparities" was implemented to teach skills related to cultural competence. A new interdisciplinary experiential course was offered, "Intercultural Communication in Health Care," which included extensive discussion with patients and health care teams and attracted students from medicine, nursing, social work, physician assistant, and physical therapy programs.[b]

Although examples of BSS innovations are readily available, global changes in BSS content across schools remains uncertain, and obstacles to change persist. In particular, faculty and learner attitudes may impede further curricular changes. Astin and colleagues found that about one-third of respondents believed that addressing psychosocial factors yielded minimal or no improvement in patient care. Most believed their past psychosocial training was ineffective, and few desired additional training. Low self-efficacy, limited knowledge, insufficient time, and low reimbursement were all perceived as reasons why psychosocial factors are seen as less important relative to other biomedical areas of health and illness. Additional reasons for lack of receptiveness include ineffective teaching methods, poor timing, or failure to demonstrate clinical relevance.[80,81] Ongoing projects from the BSS Consortium for Medical Education and others are addressing both attitudinal as well as structural barriers that limit BSS content and its translation into clinical care—that is, a concerted effort to teach the "right" content based on the health and health care needs of the public.

Example 2: Public Health

Despite the impressive advances in biomedical science and other emerging fields, such as personalized genomic medicine, the largest gains in life expectancy and improved quality of life are attributable to social and environmental interventions.[82,83] In the past century, life expectancy has increased by 30 years, and infant mortality has dropped by 95 percent. For 25 of these years, these improvements are attributable to changes in social and physical environments, such as water and food safety, tobacco interventions, and housing conditions.[84-87] While individual behavior change counseling is and should be taught,[88] the roles social networks and broader environmental influences play on health behavior are currently underappreciated by physicians.[89,90] A solid grounding in the population sciences, epidemiology, and public health should be a core element of training. Although evidence-based medicine and epidemiology are now standard elements of every medical student's education, ways to change the conditions of daily life to make them conducive to healthy behaviors — what has been called "citizen-centered health promotion"— should be highlighted in the medical school curriculum.[91]

[b] See Diversity Matters, School of Medicine and Public Health, University of Wisconsin-Madison at http://www med.wisc.edu/diversity/our-vision/38255.

The Center for Community Health and Service Learning at the Morehouse School of Medicine provides an interesting example of intensive, meaningful public health training and community engagement.[92,93] In this required, two-semester course, students learn the basic principles of public health, needs assessments, and policy interventions. Students then are assigned to community placements where they participate in needs assessments and the development and testing of interventions. Project examples include a fitness program for adolescent girls at a local middle school, HIV prevention interventions at the YMCA, and third trimester counseling on the benefits of breastfeeding at a community health center. Morehouse School of Medicine was recently ranked number one on the Social Mission Score[9] and consistently graduates physicians who undertake primary care with a focus on the underserved.

Example 3. Systems Science and Quality Improvement

Health care delivery systems have become vast and complex. Independent or even small group practices are increasingly rare, and regardless of practice setting, advanced skills in treatment referral, care coordination, and billing are routinely necessary. In such large and difficult to coordinate systems, substantial errors, suboptimal care, and lack of access occur.

The IOM reports "To Err is Human"[94] and "Crossing the Quality Chasm"[95] were instrumental in highlighting the magnitude and costs of medical errors while moving the focus away from individual incompetence toward system and contextual factors that influence patient safety and health care quality. The 1999 report on medical errors called for the creation of a National Center for Patient Safety within the Agency for Healthcare Research and Quality (AHRQ) and essentially spawned a robust, new field of health care research. Medical centers across the country invested substantial resources in creating a "culture of safety," and ongoing quality improvement projects targeting both micro and macro-level changes are increasingly commonplace.[96] The 2001 quality of care report highlighted the "chasm" between the current health care delivery system and an ideal one. This report called for sweeping redesigns based on six core aims or features of quality: health care should be safe, effective, patient-centered, timely, efficient, and equitable.[95]

Training programs should intentionally integrate quality aims into medical school curricula.[7] Examples of new competencies include quality improvement, change management, micro- and macro-systems thinking, measuring and correcting performance, maximizing interprofessional communication and cooperation, using health information technology and the electronic health record to provide optimal care, and promoting patient safety and both team and personal accountability.[97] A synthesis of these and other potential new competencies (e.g. medical school content) for the modern physician are provided in Figure 1 and will be represented in the new University of California, San Francisco (UCSF) Bridges Curriculum in fall 2015.[7]

To support emerging curricula in quality improvement and patient safety (QI/PS), the Association of American Medical Colleges recently launched the Teaching for Quality (Te4Q) initiative. The Te4Q expert panel report and recent competencies document both support institutional assessments, faculty development, local QI and patient safety initiatives, and provide ongoing mentorship and support.[98] The program's stated goal is "to ensure that every medical school and teaching hospital

in the United States has access to a critical mass of faculty ready, able, and willing to engage in, role model, and lead education in QI/PS and in the reduction of excess health care costs."[98] Thus far, 11 academic medical centers have participated in the Te4Q pilot to promote QI/PS with all showing changes in continuing medical education, and two demonstrating improvement in clinical outcomes.[99] Similar initiatives such as the "Retooling for Quality and Safety initiative" seek to integrate QI/PS curricula and interventions into both medical and nursing school curricula.[100]

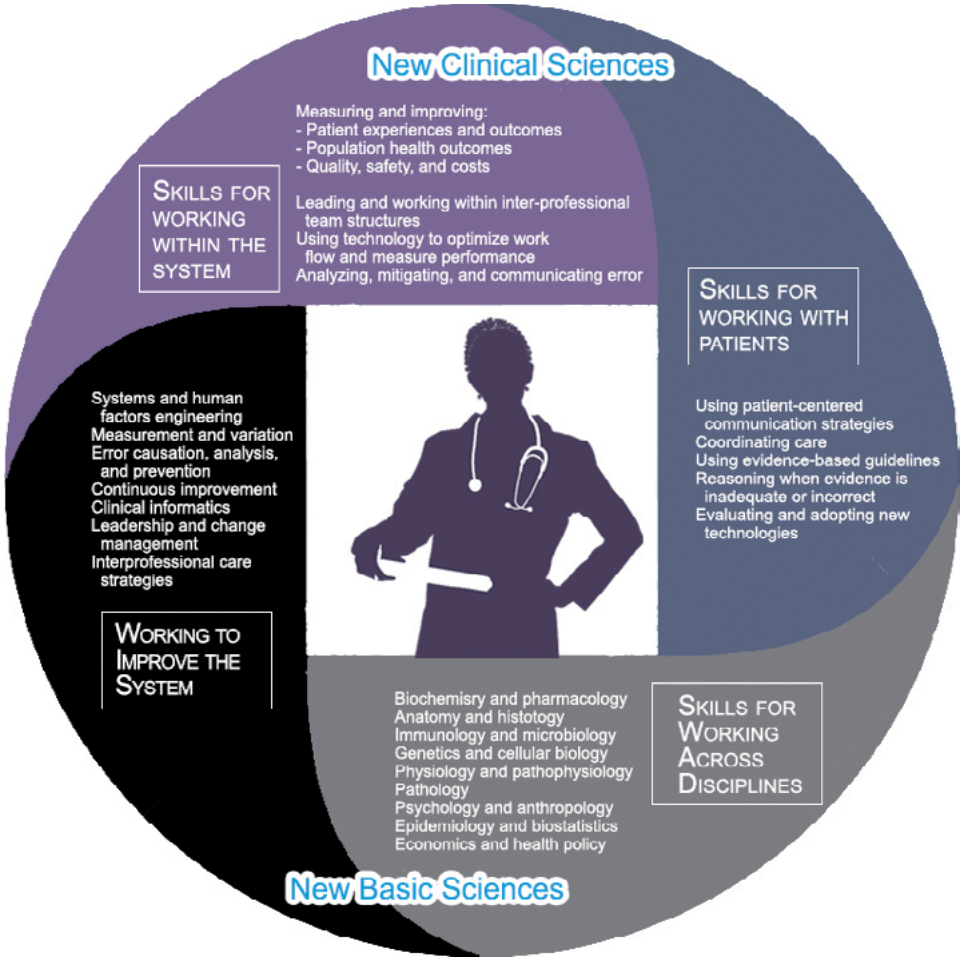


Figure 1. Expanded competencies for the 21st century physician

Question 3: Are We Teaching Using the Best Techniques?

The recent 100-year anniversary of the seminal Flexner report[101] triggered a Carnegie-funded, in-depth review of medical education pedagogy including an update of workforce development theories and evidence-based teaching strategies.[8] Similarly, the American Medical Association, the American Osteopathic Organization, the Association of American Medical Colleges, the Accreditation Council for Graduate Medical Education, the Liaison Committee on Medical Education, and the National Board of Medical Examiners have performed training and/or assessment reviews with the shared goal of improving medical education and clinical competence.[61,102-104] Major conclusions

and recommendations of these efforts include: (1) moving away from traditional didactic lectures and using more engaging and interactive teaching formats (e.g. small group discussion, team-based learning, game theory); (2) standardizing core competencies or content with flexibility in tailoring both depth and time-based aspects of the program based on learner competence or career interests; (3) including early (and frequent) opportunities to integrate and apply new content in clinical, leadership, or research settings; (4) teaching "habits of mind and heart" or meta-skills that transfer across content areas and contexts; (5) providing key professional identity formation activities ("workplace learning"), including immersion experiences with high-functioning interprofessional teams; and (6) providing longitudinal experiences with both mentors and patients that deepen learning and enrich emotional and cognitive skills, such as empathy and ethical/moral judgment. In contrast to the ongoing struggles with medical applicant recruitment and inclusion of curricular content, some have argued that medical school pedagogy has never been better.[7,105]

Although medical schools seem to be continually revising their curricular structures and teaching methods, we question how much progress has really been made. Traditional 2 x 2 medical school curricula still predominate, with a heavy reliance on didactic basic science instruction in the first 2 years and clinical apprenticeships in the last 2 years. Knowledge retention and application of basic science into clinical care have been perpetual problems.[106,107] Moreover, as the body of knowledge in the basic sciences has exponentially grown, it has become virtually impossible to fit "essential" content into a fixed number of curricular hours. Medical students have likened their learning experience to "drinking from a fire hose" and learning through "intellectual bulimia" where they binge on facts that are later purged on examinations and then forgotten. The amount of information to be learned now exceeds a learner's cognitive ability, strongly implicating the need for new learning and information management strategies and requiring consideration of working memory and cognitive load when delivering content.[108]

BSS insights gained from cognitive science and psychology in the realm of "meta-cognitive skills" could be instructive. The skills of critical thinking, elaboration, information management, medical informatics, evaluation of evidence, critical reflection, and independent adaptive or life-long learning will consume curricular time but may ultimately give learners a transferrable skill set that allows them to manage the vast libraries of medical knowledge available at their fingertips.[109-111] Although impressive progress and innovations have been amply demonstrated, many medical schools feel compelled to frontload as much basic science as possible into the first 2 years to improve learner performance on the USMLE Step 1 board exam, typically taken after Year 2 and just before leaners start their clinical rotations in Year 3.[112] Structural, pedagogical, and policy changes will need to evolve hand-in-hand.

Fortunately, examples of more "learner-centered" curricula abound in both undergraduate[4,113] and graduate medical education[114] along with an impressive diversity of pedagogical innovations, such as team-based learning, clinical simulations, and standardized patient assessments.[5,6,115,116,] For example, Indiana University used team-based learning sessions to teach medical ethics around the topics of Informed Consent, Duty to Warn, Futility, and Organ Donation. These highly charged, interactive sessions challenged learners to "solve" complex, real world clinical scenarios within small, competing groups.[117] Medical students at UCSF and Stanford receive a three-part "meta-cognitive skills" course to facilitate the transition from classroom learning to clinical apprenticeship

learning in clerkships. The Albert Einstein College of Medicine and the Warren Alpert Medical School of Brown University developed a series of standardized patient exercises that address ethics and professionalism and use the Implicit Association Test (IAT) to inform discussions and self-recognition of bias.[118]

Although most medical schools have adopted competency-based medical education and assessments,[61,63] the potential for tailored learning has lagged. Ideally, when a learner achieves a specified competency, s/he should be allowed to move on to the next level, abandoning time-based curricular programs in favor of competency-based ones. In addition to having more flexibility in use of time, curricula could also be tailored to the strengths and interests of a particular learner, especially in experiences that might better prepare them for modern day medical practice or assist them in identifying a career path. Some students could opt for infectious disease rotations at the public health department, while others might participate in quality improvement projects to reduce 30-day hospital readmission rates. Such varied but essential "workplace" learning is feasible, increasingly common, and will be featured in a number of new medical school curricula.[119]

One unintended consequence of our current medical pedagogy is the unfortunate reductions in student and resident empathy that occurs over the course of medical training.[120-122] Although motivations for enrolling in medical school vary widely, most students arrive with a genuine altruistic interest and desire to alleviate suffering. Over the course of their training, most learners become burned out, apathetic, and even depressed.[67,123,124] A growing body of literature suggests that attending to emotion and communication skills in medical providers, included as curricular content while additionally used to shape more humane pedagogical approaches, might beneficially impact both patient and provider outcomes and prevent erosion of empathy.[124,125] The skillful detection, understanding, and management of emotions in the self and others may impact learner well-being, satisfaction, and professionalism and may lead to improved patient outcomes.[126-28]

The Columbia University College of Physicians and Surgeons developed an intensive "teach-the-teachers" narrative medicine curriculum to sharpen the clinical, teaching, and reflective skills of faculty, while enriching relationships, deepening empathy, and promoting wellness. Narrative medicine integrates medicine and literary studies to "fortify clinical practice with the narrative competence to recognize, absorb, metabolize, interpret, and be moved by the stories of illness."[129] In the Columbia program, faculty meet regularly to guide and be guided through critical analyses of poetry, short stories, and writing exercises. Results suggest that personal enjoyment and growth occur in addition to meaningful professional transformations and changes in teaching practices.[130-132]

Ultimately, there is little question that medical education pedagogy has made extraordinary improvements in using the science of learning to shape new teaching methods. These pedagogies are not only more engaging but are better suited to teaching complex interrelationships and nuanced skills well known to BSS. These pedagogies include team-based learning, small-group discussions, problem-based learning, reflective writing, mentorship, and simulations where students learn instead of being taught.[133-135] Most incorporate the creation of "learning communities" with longitudinal experiences that allow learners to develop group ownership, and both faculty and students can share both the direction and facilitation of learning activities. This "evidence-based teaching" maximizes effectiveness and encourages ongoing learner engagement and the development of leadership skills.[105,136] Our hope is that recent advances in pedagogy can be applied

to necessary new content and be experienced by more diverse learners interested in improving the health of the public.

As with any workforce development or training program, environmental, economic, and legislative contexts may greatly alter both the type and rate of medical education innovations that occur. For example, the Affordable Care Act potentially will insure millions of new patients or "customers" of medical systems thereby altering the number and proportion of primary care providers that must be trained. Medical education research continues to test and refine innovative teaching methods and assessment tools, although schools struggle with how to satisfy credentialing bodies, prepare learners for licensure tests, maintain student wellness, and empower learners to want to provide patient-centered care. Advances in BSS curricular content abound, but dissemination and implementation have been slow.[10,73] Two potentially large "landscape changes" are described below – the new Medical College Admission Test (MCAT) and the growing acceptance of interprofessional education and non-physician medical team members. Each have the potential to alter and enrich the impressive, ongoing innovations in medical education and should be viewed as opportunities to support medical education in better meeting the needs of the public.

Nearly every medical school applicant takes the MCAT as part of the admissions process. The current MCAT (in use since 1991) includes four sections: physical sciences, verbal reasoning, a writing sample, and biological sciences. In 2015, approximately one-third of a newly revised MCAT will be devoted to behavioral and social sciences questions. The new subtest, "Psychological, Social, and Biological Foundations of Behavior," will measure BSS foundational knowledge in five core content areas: (1) ways that individuals perceive, think about, and react to the world; (2) factors that influence behavior and behavior change; (3) factors that influence how we think about ourselves and others; (4) ways in which culture and social differences influence well-being; and (5) ways in which social stratification affects access to resources and well-being. Core content in the BSS and other new sections of the MCAT was selected through expert panel consensus and faculty surveys developed and administered by the MR5 (MCAT 5th Revision Committee) beginning in 2008.[137] BSS representation included members from the NIH-supported BSS Consortium for Medical Education who were instrumental in identifying "what every medical student needs to know."[138]

Although MCAT changes may not alter the composition of medical school applicants or those admitted, there are important global implications. First, pre-medical preparation will most likely be expanded to include training in psychology, sociology, and public health in addition to the traditional foundational sciences such as chemistry and biology. Although this creates an added burden on students, it ensures they all arrive in medical school with minimally sufficient BSS knowledge along with chemistry or biology. Subsequent BSS instruction, even in the first year, can begin at more sophisticated levels and progress more rapidly to complex translation of BSS into clinical experiences. Second, medical school applicants will receive an unequivocal message that BSS constructs (e.g. diversity, disparities, equity, empathy, communication, prevention, behavior) are not only central to medical education but equally valued as part of the foundational preparation for a career in medicine. Although the effects of the new MCAT are yet to be determined, preliminary data from a cross-sectional validity study show that new MCAT scores correlate with older MCAT scores in expected ways and correlate with undergraduate grade point averages, and BSS scores are higher in students who have studied those disciplines.[139]

Interprofessional education (IPE) and the acceptance of more team-based care provide a second example of a transformative landscape change on both national and international levels.[140] Although the notion of collaboratively and concurrently training health professionals from multiple disciplines is not new, recent systematic reviews and studies of IPE suggest it should be more seriously considered as a way to address the triple aim of: (1) improving the patient experience of care (including quality and satisfaction), (2) improving the health of populations, and (3) reducing the per capita cost of health care.[140-142] In general, IPE is well received by learners, can effectively teach teamwork and communication skills, and will likely improve service efficiencies, reduce clinical errors, and improve patient satisfaction.[143,144] Genuine acceptance and scaling up of IPE could have profound implications. In a not too distant future, applicants may be able to apply to a broad health professional program and then choose to specialize in a discipline during training instead of applying to nursing school or medical school and remaining locked into an initial training track. More likely is the addition and/or revision of existing curricula to include team work and communication skills, as well as a better understanding of the roles and abilities of diverse team members. From a patient and public health perspective, health care needs could soon be met by any one of a team of multidisciplinary professionals who are well coordinated and best suited to the task at hand. Rather than seeing IPE as a threat to discipline dominance, medical schools should recognize IPE's promise in meeting the needs of the public and potentially lightening the burden currently placed on physicians.

Implications for Practice

Each of our three central questions – who gets trained, what are they taught, and what teaching methods are used – have indirect yet important implications for clinical practice. Who gets trained (i.e., admitted to medical schools) will directly affect the composition of the physician workforce. Composition shifts may alter what field graduates choose to go into and what populations they are willing to serve.[46,47] Admissions committees have already begun the shift toward "holistic" admissions[38] with some de-emphasizing the MCAT and others incorporating new and more equitable interview procedures to assess "non-cognitive" factors (e.g. multiple mini-interview[145]). Both faculty and student composition should accurately reflect the clinical populations being served and will require a substantial investment in recruitment and retention efforts. Curricular content and teaching methods could be revised to better represent the leading causes of morbidity and mortality while employing strategies learned from educational and cognitive sciences to improve classroom efficiency and learner engagement.[109,111,113] Ultimately, the challenge (and potential reward) rests on a graduate's ability to "translate" classroom skills into clinical practice. Including more required workplace learning opportunities may help with this translational leap.

From a policymaker and Federal funding perspective, it will be essential to align financial incentives (and student financial support) with the goals of training a physician workforce capable of meeting the needs of the public. More primary care residency slots should be funded, but medical centers should also be incentivized to include other types of primary care providers such as nurse practitioners and physician assistants. Similarly, interprofessional training and practice could be promoted by creating a performance rubric that assesses and rewards integrative and coordinated interprofessional care. National accrediting bodies for medical schools and medical specialties (e.g., the Liaison Committee on Medical Education [LCME] and the American Board of Internal

Medicine [ABIM]) should be lobbied to include competencies or milestones that reflect the best of behavioral and social sciences and are seen as required competencies necessary for accreditation or credentialing. While global "report cards" remain controversial, the scope of medical school performance categories could be enriched to include more than NIH funding received or percent of applicants admitted. Schools scoring high on "social mission" (i.e., those meeting the needs of the communities they serve) should be recognized and rewarded.

Implications for Research

As with Implications for Practice, each of our three central questions has implications for both clinical and educational research. Who gets trained or admitted to medical school and the career choices these medical school graduates make in terms of their career paths need intense study using longitudinal approaches. The change in the MCAT exam with questions on social and behavioral sciences represents an important investment in understanding knowledge of entering medical students. Building upon this foundation by studying how to construct effective instructional designs that foster both clinical skills development and effective tracking of health behaviors in patients and population will be vital. Even more important will be linking more effective clinical practices in social and behavioral science to trends in social and behavioral population-based health indices. Research in neighborhood design —such as distance to healthier food choices and the installation of parks and areas that promote more exercise and physical activity—are all important topics that merit further development in terms of population-based research and the role physicians can play in conducting this kind of research.

The cognitive neurosciences will certainly advance in identifying more effective approaches to address detrimental health behaviors, such as smoking and alcohol or recreational drug use. With advances in clinical care, understanding how best to get best practices into clinical care will be a top priority. Implementation science will inform both clinicians and health systems about how best to integrate effective strategies into practice, which should be a top research priority. Additionally, getting these new techniques included in both medical school and residency training curricula will be important, as will understanding the best educational approaches that ensure use of these techniques in patient care. In summary, research in both effective educational techniques and in ways to advance clinical practice at both the point of patient care and at the population level will all be vital areas of future research.

An important limitation in advancing best educational practices in the health professions is limitations in funding. Though the Health Resources and Services Administration (HRSA) has supported efforts to understand the contributions that medical education makes, more rigor is needed to fully investigate how changes in medical school admissions influence both effective clinical practice in social and behavioral sciences and career path choices regarding primary care. Unfortunately, the mechanisms for funding educational research are limited. While the NIH typically funds clinical and population research, including studies in the social and behavioral sciences, and the National Science Foundation funds programs to enhance science, technology, engineering, and mathematics (STEM) education in undergraduate education, ongoing funding for research in health professions education is extremely limited. This limitation is, unfortunately, slowing the pace of identifying effective educational strategies for social and behavioral sciences

in medicine and other health professions. Addressing this problem should be a top priority for educational leaders and policymakers.

Conclusions

It is clear that the enterprise of medical education worldwide is a rich, robust, and impressive incubator for stunning innovations that could both improve educational processes and address the determinants of health for individuals and communities. The three core questions remain open to debate: (1) Are we training the right people? (2) Are we teaching the right content and competencies? (3) Are we teaching using the best techniques? However, impressive innovations largely derived from the behavioral and social sciences highlight potential solutions to address applicant recruitment, specialty selection, commitment to the underserved, and essential new content and competencies required of the 21st century physician who may be collaboratively treating older, more diverse patients with multiple chronic diseases. Larger landscape changes such as the Affordable Care Act, the new MCAT, and IPE offer examples of how dramatically different medical education may be in just 10 years from now. Our stewardship of this process is vitally important if real change is to occur. This stewardship must include medical educators, scientists, learners at all levels, stakeholders in local communities, health policymakers, health systems administrators, and patients themselves. Medical education can and should be realigned to better meet the Nation's health priorities. Training innovations drawn from the behavioral and social sciences offer one essential pathway to success.

Acknowledgments

Funding for this work was provided by NIH/NCCAM/OBSSR grant 5R25AT006573 and NIH/NCI/OBSSR grants 1R25CA158571-01A1 and 5R25CA158571-02. The opinions presented herein are those of the authors and may not represent the position of the Agency for Healthcare Research and Quality, the National Institutes of Health, or the U.S. Department of Health and Human Services.

Authors' Affiliations

Jason M. Satterfield, PhD, Professor of Medicine, University of California, San Francisco, San Francisco, CA. Patricia A. Carney, PhD, Professor of Family Medicine and of Public Health & Preventive Medicine, Oregon Health & Science University, Portland, OR.

Address correspondence to: Jason Satterfield, PhD, University of California, 1701 Divisadero Street, Suite 500, San Francisco, CA 94115; email: jsatter@medicine.ucsf.edu

References

1. Table 27: Total graduates by U.S. medical school and sex, 2010-2014. Washington, DC: Association of American Medical Colleges; 2013. Available at https://www.aamc.org/download/321532/data/factstable27-2.pdf. Accessed October 20, 2014.
2. Cooke M, Irby DM, Sullivan W, et al. American medical education 100 years after the Flexner report. N Engl J Med 2006 Sep 28;355(13):1339-44.
3. Irby DM, Cooke M, O'Brien BC. Calls for reform of medical education by the Carnegie Foundation for the Advancement of Teaching: 1910 and 2010. Acad Med 2010 Feb;85(2):220-7.
4. McLaughlin JE, Roth MT, Glatt DM, et al. The flipped classroom: a course redesign to foster learning and engagement in a health professions school. Acad Med 2014 Feb;89(2):236-43.
5. Cook DA, Hatala R, Brydges R, et al. Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. JAMA 2011;306(9):978-88.
6. Ruiz JG, Mintzer MJ, Leipzig RM. The impact of e-learning in medical education. Acad Med 2006 Mar;81(3):207-12.
7. Lucey CR. Medical education: part of the problem and part of the solution. JAMA Intern Med 2013 Sep 23; 173(17):1639-43.
8. Cooke M, Irby DM, O'Brien BC, et al. Educating physicians: a call for reform of medical school and residency. San Francisco, CA: Jossey-Bass; 2010.
9. Mullan F, Chen C, Petterson S, et al. The social mission of medical education: ranking the schools. Ann Intern Med 2010 Jun 15;152(12):804-11.
10. Satterfield JM, Adler SR, Chen HC, et al. Creating an ideal social and behavioural sciences curriculum for medical students. Med Educ 2010 Dec;44(12):1194-1202.
11. Health at a glance 2007. Danvers, MA: OECD Publishing. Available at http://www.oecd-ilibrary.org/social-issues-migration-health/health-at-a-glance-2007_health_glance-2007-en. Accessed October 20, 2014.
12. U.S. Burden of Disease Collaborators. The state of U.S. health, 1990-2010: burden of diseases, injuries, and risk factors. JAMA 2013 Aug 14;310(6):591-608.
13. Woolf SH, Aron LY, National Academies (U.S.), et al. U.S. health in international perspective: shorter lives, poorer health. Washington, DC: National Academies Press; 2013.
14. Frenk J, Chen L, Bhutta ZA, et al. Health professionals for a new century: transforming education to strengthen health systems in an interdependent world. Lancet 2010 Dec 4;376(9756):1923-58.
15. Boelen C, Heck JE, World Health Organization. Defining and measuring the social accountability of medical schools. Geneva: World Health Organization; 1995.
16. Mitka M. Report: growth of medical schools brings opportunity to redefine their mission. JAMA 2009 Mar 18; 301(11):1114-5.
17. McCurdy L, Goode LD, Inui TS, et al. Fulfilling the social contract between medical schools and the public. Acad Med 1997 Dec;72(12):1063-70.
18. Schroeder SA, Zones JS, Showstack JA. Academic medicine as a public trust. JAMA 1989 Aug 11;262(6):803-12.
19. Freeman J, Ferrer RL, Greiner KA. Viewpoint: developing a physician workforce for America's disadvantaged. Acad Med 2007 Feb;82(2):133-8.
20. Woollard RF. Caring for a common future: medical schools' social accountability. Med Educ 2006 Apr;40(4):301-13.
21. Total physicians by race/ethnicity. Chicago, IL: American Medical Association; 2006.
22. Annual estimates of the resident population by sex, race, and Hispanic origin for the United States: April 1, 2000 to July 1, 2008. Washington, DC: U.S. Census Bureau; 2008. Available at http://www.census.gov/popest/data/historical/2000s/vintage_2008/index.html. Accessed May 5, 2010.
23. Greico E, Acosta Y, de la Cruz P, et al. The foreign-born population in the United States: 2010. Washington, DC: U.S. Census Bureau; 2012.
24. Shin H, Bruno R. Language use and English-speaking ability: 2000. Washington, DC: U.S. Census Bureau; October 2003.
25. Regenstein M, Andres E, Wynia MK. Appropriate use of non-English-language skills in clinical care. JAMA 2013 Jan 9;309(2):145-6.
26. Diamond LC, Schenker Y, Curry L, et al. Getting by: underuse of interpreters by resident physicians. J Gen Intern Med 2009 Feb; 24(2):256-62.
27. Lee KC, Winickoff JP, Kim MK, et al. Resident physicians' use of professional and nonprofessional interpreters: a national survey. JAMA 2006 Sep 6; 296(9):1050-3.
28. Cooper LA, Roter DL, Johnson RL, et al. Patient-centered communication, ratings of care, and concordance of patient and physician race. Ann Intern Med 2003 Dec 2;139(11):907-15.

29. Fernandez A, Schillinger D, Grumbach K, et al. Physician language ability and cultural competence. An exploratory study of communication with Spanish-speaking patients. J Gen Intern Med 2004 Feb;19(2):167-74.
30. Saha S, Guiton G, Wimmers PF, et al. Student body racial and ethnic composition and diversity-related outcomes in U.S. medical schools. JAMA 2008 Sep 10;300(10):1135-45.
31. UCLA prep program. Los Angeles, CA: UCLA Office of Diversity, Inclusion, and Outreach. Available at http://www.medstudent.ucla.edu/offices/aeo/prep.cfm. Accessed May 10, 2014.
32. U.S. Census Bureau. California State and County QuickFacts. 2012. Available at http://quickfacts.census.gov/qfd/states/06000 html. Accessed May 3, 2014.
33. Beacham T, Askew RW, William PR. Strategies to increase racial/ethnic student participation in the nursing profession. ABNF J 2009 Sum;20(3):69-72.
34. Brunson WD, Jackson DL, Sinkford JC, et al. Components of effective outreach and recruitment programs for underrepresented minority and low-income dental students. J Dent Educ 2010 Oct;74(10 Suppl):S74-86.
35. Rumala BB, Cason FD Jr. Recruitment of underrepresented minority students to medical school: minority medical student organizations, an untapped resource. J Natl Med Assoc 2007 Sep;99(9):1000-4, 1008-9.
36. Guevara JP, Adanga E, Avakame E, et al. Minority faculty development programs and underrepresented minority faculty representation at U.S. medical schools. JAMA 2013 Dec 4;310(21):2297-304.
37. Rodriguez JE, Campbell KM, Fogarty JP, et al. Underrepresented minority faculty in academic medicine: a systematic review of URM faculty development. Fam Med 2014 Feb;46(2):100-4.
38. Addams A, Bletzinger R, Sondheimer H, et al. Roadmap to diversity: integrating holistic review practices into medical school admission processes. Washington, DC: Association of American Medical Colleges; 2010.
39. Physician distribution and health care challenges in rural and inner-city areas. Rockville, MD: Council on Graduate Medical Education; 1998.
40. Jones C, Parker T, Ahearn M, et al. Health status and health care access of farm and rural populations. Washington, DC: U.S. Department of Agriculture, Economic Research Service; 2009.
41. Ricketts TC, Randolph R. The diffusion of physicians. Health Aff 2008 Sep-Oct;27(5):1409-15.
42. Access transformed: building a primary care workforce for the 21st century. Bethesda, MD: National Association of Community Health Centers; August 2008.
43. Rosenblatt RA, Andrilla CH, Curtin T, et al. Shortages of medical personnel at community health centers: implications for planned expansion. JAMA 2006 Mar 1;295(9):1042-9.
44. Colwill JM, Cultice JM, Kruse RL. Will generalist physician supply meet demands of an increasing and aging population? Health Aff 2008 May-Jun;27(3):232-41.
45. Ogunyemi D, Edelstein R. Career intentions of U.S. medical graduates and international medical graduates. J Natl Med Assoc 2007 Oct; 99(10):1132–7.
46. Minorities in medicine: an ethnic and cultural challenge for physician training. Rockville, MD: Council on Graduate Medical Education; 2005.
47. Komaromy M, Grumbach K, Drake M, et al. The role of black and Hispanic physicians in providing health care for underserved populations. N Engl J Med 1996 May 16;334(20):1305-10.
48. Statement on the physician work-force. Washington, DC: Association of American Medical Colleges; 2006. Available at www.aamc.org/workforce. Accessed March 20, 2007
49. Ebell MH. Future salary and U.S. residency fill rate revisited. JAMA 2008 Sep 10;300(10):1131-2.
50. Hauer KE, Durning SJ, Kernan WN, et al. Factors associated with medical students' career choices regarding internal medicine. JAMA 2008 Sep 10;300(10):1154-64.
51. Lakhan SE, Laird C. Addressing the primary care physician shortage in an evolving medical workforce. Int Arch Med 2009;2(1):14.
52. Deutschlander S, Suter E, Grymonpre R. Interprofessional practice education: is the 'interprofessional' component relevant to recruiting new graduates to underserved areas? Rural Remote Health 2013 Oct-Dec;13(4):2489.
53. Fairman JA, Rowe JW, Hassmiller S, et al. Broadening the scope of nursing practice. New Engl J Med 2011 Jan 20;364(3):193-6.
54. Grumbach K, Hart LG, Mertz E, et al. Who is caring for the underserved? A comparison of primary care physicians and nonphysician clinicians in California and Washington. Ann Fam Med 2003 Jul-Aug;1(2):97-104.
55. Institute of Medicine. The future of nursing: leading change, advancing health. Washington, DC: National Academies Press; 2011.
56. Donelan K, DesRoches CM, Dittus RS, et al. Perspectives of physicians and nurse practitioners on primary care practice. N Engl J Med 2013 May 6;368(20):1898-1906.
57. Patchin RJ. AMA responds to IOM report on future of nursing [press release]. Chicago, IL: American Medical Association; October 5, 2010.
58. Norris TE, Schaad DC, DeWitt D, et al. Longitudinal integrated clerkships for medical students: an

innovation adopted by medical schools in Australia, Canada, South Africa, and the United States. Acad Med 2009 Jul;84(7):902-7.
59. Bell SK, Krupat E, Fazio SB, et al. Longitudinal pedagogy: a successful response to the fragmentation of the third-year medical student clerkship experience. Acad Med 2008 May;83(5):467-75.
60. Accreditation Council for Graduate Medical Education. ACGME Outcome Project 2000. Available at http://www.acgme.org. Accessed November 10, 2004.
61. Swing SR. The ACGME outcome project: retrospective and prospective. Med Teach 2007 Sep;29(7):648-54.
62. Carraccio C, Wolfsthal SD, Englander R, et al. Shifting paradigms: from Flexner to competencies. Acad Med 2002 May;77(5):361-7.
63. Frank JR, Snell LS, Cate OT, et al. Competency-based medical education: theory to practice. Med Teach 2010;32(8):638-45.
64. Lurie SJ, Mooney CJ, Lyness JM. Measurement of the general competencies of the accreditation council for graduate medical education: a systematic review. Acad Med 2009 Mar;84(3):301-9.
65. Behavioral and social science foundations for future physicians. Washington, DC: Association of American Medical Colleges; 2011.
66. Association of American Medical Colleges and the Howard Hughes Medical Institute. Report of Scientific Foundations for Future Physicians Committee. Washington, DC: Association of American Medical Colleges; 2009. Available at https://members.aamc.org/eweb/upload/Scientific%20Foundations%20for%20Future%20Physicians%20%20Report2%202009.pdf. Accessed October 20, 2014.
67. Satterfield JM, Becerra C. Developmental challenges, stressors and coping strategies in medical residents: a qualitative analysis of support groups. Med Educ 2010 Sep;44(9):908-16.
68. McGinnis JM, Foege WH. Actual causes of death in the United States. JAMA 1993 Nov 10;270(18):2207-12.
69. Mokdad AH, Marks JS, Stroup DF, et al. Actual causes of death in the United States, 2000. JAMA 2004 Mar 10;291(10):1238-45.
70. Schroeder SA. Shattuck Lecture. We can do better—improving the health of the American people. N Engl J Med 2007 Sep 20;357(12):1221-8.
71. Benbassat J, Baumal R, Borkan JM, et al. Overcoming barriers to teaching the behavioral and social sciences to medical students. Acad Med 2003 Apr;78(4):372-80.
72. Cuff PA, Vanselow NA, Institute of Medicine (U.S.). Committee on Behavioral and Social Sciences in Medical School Curricula. Improving medical education: enhancing the behavioral and social science content of medical school curricula. Washington, DC: National Academies Press; 2004.
73. Hollar D, Satterfield J, Carney P, et al. The National Institutes of Health Social and Behavioral Science Consortium: an introduction and progress report on UME curricular innovations. Ann Behav Sci Med Educ 2007;13(2), 60-8.
74. Enhancing behavioral and social sciences in undergraduate medical education. Bethesda, MD: National Institutes of Health, Office of Behavioral and Social Sciences Research; updated August 2014. Available at http://obssr.od nih.gov/training_and_education/bss_medical_school/. Accessed October 20, 2014.
75. Clough RW, Shea SL, Hamilton WR, et al. Weaving basic and social sciences into a case-based, clinically oriented medical curriculum: one school's approach. Acad Med 2004 Nov;79(11):1073-83.
76. Post DM, Stone LC, Knutson DJ, et al. Enhancing behavioral science education at the Ohio State University College of Medicine. Acad Med 2008 Jan;83(1):28-36.
77. Greenwald AG, McGhee DE, Schwartz JL. Measuring individual differences in implicit cognition: the implicit association test. J Pers Soc Psychol 1998 Jun;74(6):1464-80.
78. Teal CR, Shada RE, Gill AC, et al. When best intentions aren't enough: helping medical students develop strategies for managing bias about patients. J Gen Intern Med 2010 May;25(Suppl 2):S115-8.
79. Teal CR, Gill AC, Green AR, et al. Helping medical learners recognise and manage unconscious bias toward certain patient groups. Med Educ 2012 Jan;46(1):80-8.
80. Astin JA, Soeken K, Sierpina VS, et al. Barriers to the integration of psychosocial factors in medicine: results of a national survey of physicians. J Am Board Fam Med 2006 Nov-Dec;19(6):557-65.
81. Astin JA, Sierpina VS, Forys K, et al. Integration of the biopsychosocial model: perspectives of medical students and residents. Acad Med 2008 Jan;83(1):20-7.
82. Achievements in public health, 1900-1999: decline in deaths from heart disease and stroke—United States, 1900-1999. Atlanta, GA: Centers for Disease Control and Prevention; August 6, 1999.
83. Ray O. How the mind hurts and heals the body. Am Psychol 2004 Jan;59(1):29-40.
84. Adler N, Steward J, Cohen S, et al. Reaching for a healthier life: facts on socioeconomic status and health in the United States. San Francisco, CA: The John D. and Catherine T. MacArthur Foundation Research Network on Socioeconomic Status and Health; 2008.

85. Hofrichter R. Health and social justice: a reader on the politics, ideology, and inequity in the distribution of disease. San Francisco, CA: Jossey-Bass; 2003.
86. Kawachi I, Kennedy B, Wilkinson R. The society and population health reader: income inequality and health. New York, NY: The New Press; 1999.
87. Siegrist J, Marmot MG. Social inequalities in health: new evidence and policy implications. New York: Oxford University Press; 2006.
88. Hauer KE, Carney PA, Chang A, et al. Behavior change counseling curricula for medical trainees: a systematic review. Acad Med 2012 Jul;87(7):956-68.
89. Larson N, Story M. A review of environmental influences on food choices. Ann Behav Med 2009 Dec;38(Suppl 1):S56-73.
90. Pust S, Mohnen SM, Schneider S. Individual and social environment influences on smoking in children and adolescents. Pub Health 2008 Dec;122(12):1324-30.
91. Woolf SH, Dekker MM, Byrne FR, et al. Citizen-centered health promotion: building collaborations to facilitate healthy living. Am J Prev Med 2011 Jan;40(Suppl 1):S38-47.
92. Buckner AV, Ndjakani YD, Banks B, et al. Using service-learning to teach community health: the Morehouse School of Medicine Community Health Course. Acad Med 2010 Oct;85(10):1645-51.
93. McNeal MS, Blumenthal DS. Innovative ways of integrating public health into the medical school curriculum. Am J Prev Med 2011 Oct;41(4 Suppl 3):S309-11.
94. Kohn LT, Corrigan JM, Donaldson MS (Eds). To err is human: building a safer health system. Washington, DC: National Academy Press, Institute of Medicine; 1999.
95. Institute of Medicine, Crossing the quality chasm: a new health system for the twenty-first century. Washington, DC: National Academies Press; 2001.
96. Kroch E, Duan M, Silow-Carroll S, et al. Hospital performance improvement: trends in quality and efficiency—a quantitative analysis of performance improvement in U.S. hospitals. New York, NY: The Commonwealth Fund; 2007.
97. Berwick DM, Finkelstein JA. Preparing medical students for the continual improvement of health and health care: Abraham Flexner and the new "public interest." Acad Med 2010 Sep;85(9 Suppl):S56-65.
98. Headrick L, Baron R, Pingleton S, et al. Teaching for quality: integrating quality improvement and patient safety across the continuum of medical education. Washington, DC: Association of American Medical Colleges; 2013.
99. Davis NL, Davis DA, Johnson NM, et al. Aligning academic continuing medical education with quality improvement: a model for the 21st century. Acad Med 2013 Oct;88(10):1437-41.
100. Headrick LA, Barton AJ, Ogrinc G, et al. Results of an effort to integrate quality and safety into medical and nursing school curricula and foster joint learning. Health Aff 2012 Dec;31(12):2669-80.
101. Flexner A, Carnegie Foundation for the Advancement of Teaching, Pritchett HS. Medical education in the United States and Canada; a report to the Carnegie Foundation for the Advancement of Teaching. New York, NY: Carnegie Foundation; 1910.
102. Haist SA, Katsufrakis PJ, Dillon GF. The evolution of the United States Medical Licensing Examination (USMLE): enhancing assessment of practice-related competencies. JAMA 2013 Dec 4;310(21):2245-6.
103. Kaplan RM. Health outcomes and communication research. Health Comm 1997;9(1):75-82.
104. Skochelak SE. A decade of reports calling for change in medical education: what do they say? Acad Med 2010 Sep;85(9 Suppl):S26-33.
105. Norman G. Research in medical education: three decades of progress. Br Med J 2002 Jun 29;324(7353):1560-2.
106. Custers EJ. Long-term retention of basic science knowledge: a review study. Adv Health Sci Educ Theory Pract 2010;15(1):109-28.
107. Finnerty EP, Chauvin S, Bonaminio G, et al. Flexner revisited: the role and value of the basic sciences in medical education. Acad Med 2010 Feb;85(2):349-55.
108. Stead WW, Searle JR, Fessler HE, et al. Biomedical informatics: changing what physicians need to know and how they learn. Acad Med 2011 Apr;86(4):429-34.
109. Bordage G. Elaborated knowledge: a key to successful diagnostic thinking. Acad Med 1994 Nov;69(11):883-5.
110. Mann KV. Reflection: understanding its influence on practice. Med Educ 2008 May;42(5):449-51.
111. Mitchell R, Regan-Smith M, Fisher MA, et al. A new measure of the cognitive, metacognitive, and experiential aspects of residents' learning. Acad Med 2009 Jul;84(7):918-26.
112. Battinelli D, Smith L. Correspondence: "Alliance groups respond to proposed changes to USMLE." Acad Int Med Insight 2008;6(4):5.
113. Ludmerer KM. Learner-centered medical education. N Engl J Med 2004 Sep 16;351(12):1163-4.

114. Carney PA, Eiff MP, Saultz JW, et al. Assessing the impact of innovative training of family physicians for the patient centered medical home. J Grad Med Educ 2012 Mar:16-22.
115. Burgess AW, McGregor DM, Mellis CM. Applying established guidelines to team-based learning programs in medical schools: a systematic review. Acad Med 2014 Apr;89(4):678-88.
116. Cook DA, Brydges R, Zendejas B, et al. Technology-enhanced simulation to assess health professionals: a systematic review of validity evidence, research methods, and reporting quality. Acad Med 2013 Jun;88(6):872-83.
117. Gaffney M, Cottingham A. Four medical ethics team based learning modules. MedEdPORTAL 2012;ID=9268.
118. Gonzalez CM, Kim MY, Marantz PR. Implicit bias and its relation to health disparities: a teaching program and survey of medical students. Teach Learn Med 2014;26(1):64-71.
119. Chen HC, Sheu L, O'Sullivan P, et al. Legitimate workplace roles and activities for early learners. Med Educ 2014 Feb;48(2):136-45.
120. Bellini LM, Baime M, Shea JA. Variation of mood and empathy during internship. JAMA 2002 Jun 19;287(23):3143-6.
121. Bellini LM, Shea JA. Mood change and empathy decline persist during three years of internal medicine training. Acad Med 2005 Feb;80(2):164-7.
122. Hojat M, Mangione S, Nasca TJ, et al. An empirical study of decline in empathy in medical school. Med Educ 2004 Sep;38(9):934-41.
123. Ratanawongsa N, Wright SM, Carrese JA. Well-being in residency: effects on relationships with patients, interactions with colleagues, performance, and motivation. Patient Educ Couns 2008 Aug;72(2):194-200.
124. Shanafelt TD, Bradley KA, Wipf JE, et al. Burnout and self-reported patient care in an internal medicine residency program. Ann Intern Med 2002 Mar 5;136(5):358-67.
125. Epstein RM. Mindful practice. JAMA 1999 Sep 1;282(9):833-9.
126. Arora S, Ashrafian H, Davis R, et al. Emotional intelligence in medicine: a systematic review through the context of the ACGME competencies. Med Educ 2010 Aug;44(8):749-64.
127. Firth-Cozens, J. Interventions to improve physicians' well-being and patient care. Soc Sci Med 2001;52:215–22.
128. Shapiro SL, Shapiro DE, Schwartz GE. Stress management in medical education: a review of the literature. Acad Med 2000 Jul;75(7):748-59.
129. Amiel J, Armstrong-Cohen A, Bernitz MJ, et al. Narrative medicine in education and practice. In Feldman, Christensen, Satterfield (Eds.), Behavioral medicine, 4th Ed. New York, NY: McGraw-Hill; in press.
130. Balmer DF, Richards BF. Faculty development as transformation: lessons learned from a process-oriented program. Teach Learn Med 2012;24(3):242-7.
131. Charon R. Narrative medicine: honoring the stories of illness. New York, NY: Oxford University Press; 2006.
132. Engel J, Zarconi J, Pethtel L, et al. Narrative in health care: healing patients, practitioners, profession, and community. Oxford, UK: Radcliffe Publishing Ltd; 2008.
133. Koles PG, Stolfi A, Borges NJ, et al. The impact of team-based learning on medical students' academic performance. Acad Med 2010 Nov;85(11):1739-45.
134. Wald HS, Davis SW, Reis SP, et al. Reflecting on reflections: enhancement of medical education curriculum with structured field notes and guided feedback. Acad Med 2009 Jul;84(7):830-7.
135. Wass V, Jones R, Van der Vleuten C. Standardized or real patients to test clinical competence? The long case revisited. Med Educ 2001 Apr;35(4):321-5.
136. Green ML. Identifying, appraising, and implementing medical education curricula: a guide for medical educators. Ann Intern Med 2001 Nov 20;135(10):889-96.
137. Preview guide for the MCAT 2015 exam (2nd Ed). Washington, DC: Association of American Medical Colleges; 2012.
138. Kaplan RM, Satterfield JM, Kington RS. Building a better physician—the case for the new MCAT. N Engl J Med 2012 Apr 5;366(14):1265-8.
139. Bereknyei S, Satterfield J, Sein A, et al. A collaborative research model across eleven medical schools and the Association of American Medical Colleges: investigating the predictive validity of the psychological, social, and biological foundations of behavior (PSBB) section of the new MCAT exam. Poster presented at Western Group on Educational Affairs, 2014; Honolulu, HI.
140. Framework for action on interprofessional education and collaborative practice. Geneva: World Health Organization; 2010.
141. Reeves S, Goldman J, Burton A, et al. Synthesis of systematic review evidence of interprofessional education. J Allied Health 2010 Fall;39(Suppl 1):198-203.

142. Institute for Healthcare Improvement. IHI Triple Aim Initiative. Available at http://www.ihi.org/Engage/Initiatives/TripleAim/Pages/default.aspx. Accessed October 22, 2014.
143. Institute of Medicine. Interprofessional education for collaboration: learning how to improve health from interprofessional models across the continuum of education to practice: workshop summary. Washington, DC: National Academies Press; 2013.
144. Reeves S, Zwarenstein M, Goldman J, et al. Interprofessional education: effects on professional practice and health care outcomes. Cochrane Database Syst Rev 2008 Jan 23;1.
145. Eva KW, Rosenfeld J, Reiter HI, et al. An admissions OSCE: the multiple mini-interview. Med Educ 2004;38:314–26.

Jason M. Satterfield, PhD, is Professor of Clinical Medicine in the Division of General Internal Medicine, University of California San Francisco. His current interests include the integration of behavioral science in medical education, dissemination and implementation of evidence-based behavioral practices, and emerging mobile health technologies for behavioral health. He is a member of the NIH Social and Behavioral Sciences Curriculum Consortium and the NIH Evidence-Based Practice Training Council. Dr. Satterfield evenly divides his time between ongoing patient care, teaching, and health services research.

Patricia A. Carney, PhD, is Professor of Family Medicine and of Public Health and Preventive Medicine at Oregon Health & Science University. She has contributed to the development of research grants funded by the National Institutes of Health, the Agency for Healthcare Research and Quality, and the Health Resources and Services Administration (HRSA), as well as the Macy Foundation and the American Medical Association. Dr. Carney has developed educational research programs in family medicine, internal medicine, pediatrics, surgery, and obstetrics and gynecology. In 2009, she received the North American Primary Care Research Group's President's Award, and in 2011, she became the first non-physician to obtain a Bishop Fellowship, which funds 1 year of protected time to undergo leadership training in educational research, conducted at her host institution, The Wilson Center at the University of Toronto.

Conclusion

Determinants of Health and Longevity

Nancy E. Adler and Aric A. Prather

Abstract

The chapters in this volume document the increasing evidence over the past few decades regarding the contribution of social and behavioral sciences to understanding the etiology and progression of disease and the patterning of disparities in health across groups. During this same period, scientific advances in understanding the genetic substrates of human biology and disease processes have increased the focus of scientific research on genetic determinants and the development of individually targeted treatments. The biological sciences operate largely in isolation from the social and behavioral sciences. In this chapter we consider current approaches to health and argue for greater integration of approaches to discovery and treatment in order to achieve our national goals of improving health and reducing disparities.

Poor Population Health

Despite spending far more on health care than any other nation, the United States ranks near the bottom on key health indicators. This paradox has been attributed to underinvestment in addressing social and behavioral determinants of health. A recent Institute of Medicine (IOM) report[1] linked the shorter overall life expectancy in the United States to problems that are either caused by behavioral risks (e.g., injuries and homicides, adolescent pregnancy and sexually transmitted infections (STIs), HIV/AIDS, drug-related deaths, lung diseases, obesity, and diabetes) or affected by social conditions (e.g., birth outcomes, heart disease, and disability).

While spending more than other countries per capita on health care services, the United States, spends less on average than do other nations on social services impacting social and behavioral determinants of health. Bradley et al., found that Organization for Economic Co-operation and Development (OECD) nations with a higher ratio of spending on social services relative to health care services have better health and longer life expectancies than do those like the United States that have a lower ratio.[2]

Health Determinants

A series of analyses have examined the factors accounting for overall health and longevity. In a landmark 1993 paper, McGinnis and Foege observed that although deaths are attributed to a specific disease (e.g., heart disease or cancer), the actual causes of death reside in the factors that determine whether and when an individual develops and succumbs to disease. These factors include tobacco, diet, activity patterns, alcohol use, microbial and toxic agents, firearms, risky sexual behaviors, motor vehicle accidents, and illicit drug use.[3]

Genetic vulnerabilities, health care, and exposures in the physical environment also contribute to health and mortality, but Centers for Disease Control and Prevention (CDC) analyses suggested a relatively greater impact of social and behavioral factors: 10 percent of premature mortality was attributed to inadequacies of health care, 20-30 percent to genetics, 5 percent to the physical environment, 15 percent to the social environment, and 40-50 percent to health-damaging behaviors.[4-6] Taken together, over half of all deaths in the United States can be "attributed to a limited number of largely preventable behaviors and exposures" (p. 1242).[6]

While "actual causes" referred to social and behavioral determinants of individual health, Link and Phelan[7] drew attention to the "fundamental causes" of these more proximal determinants. Fundamental causes reflect upstream social and economic policies that drive the behavioral and biological risk factors. Galea et al. estimated that social determinants (e.g., poverty, income inequalities, racial segregation) accounted for more than 800,000 deaths in 2000, which is "comparable to the number attributed to pathophysiological and behavioral causes (p. 1456)."[8]

The above estimates relied on available data. Although imprecise and failing to account for overlapping effects, they provide a rough order of magnitude of the health impact of various levels of determinants. The greatest uncertainty involved the contribution of genetics, since direct epidemiological evidence is lacking, and its estimated effect was not based on direct calculation. Rather, after calculating deaths attributable to other factors, remaining deaths were attributed to genetic causes. This likely overestimates the effect, since the residual variation that was attributed fully to genetics includes the interaction of genes with social and physical environments, as well as error variance across all determinants.

Social and behavioral determinants affect a wide range of health problems, which substantially increase the burden of disease (depression, diabetes, cardiovascular disease) and shorten life expectancy. For example, smoking increases risk for specific cancers, respiratory diseases (asthma, chronic obstructive pulmonary disease [COPD]), diabetes, heart disease, and stroke.[9] Eating behaviors (e.g., consuming excess fat and insufficient fruits and vegetables) and inadequate levels of physical activity contribute to coronary heart disease, type 2 diabetes, and some neurodegenerative diseases.

Genetic Determinants

The largest health research initiative in recent years was the mapping of the human genome. It has generated a great deal of research and some important advances in diagnosis and treatment. However, the impact on overall population health has been modest. Most of the advances have occurred in relation to cancer, where it has enabled development of chemotherapeutic drugs based on the genetic composition of an individual's tumor. In contrast, there have been relatively few findings on genetic determinants of disorders such as diabetes that create the largest burden of disease and greatest percent of overall mortality.

Obesity is a major determinant of diabetes and represents a threat to population health. Genetic factors appear to play a minor role; only 7 percent of severe obesity in young children can be traced to monogenetic causes, and affected children represent less than .01 percent of the population.[10] The impact of genetic variation on degree of obesity is also limited. On average, adults homozygous for a risk allele for obesity (found in 16 percent of individuals) weigh only 3 kilograms more than those lacking the risk allele; such adults have only modestly elevated odds of being classified as obese and

even smaller elevations in their odds of diabetes.[10,11] A genome-wide association (GWA) study of 2.8 million single-nucleotide polymorphisms (SNPs) identified 32 loci for body mass index (BMI), which together accounted for only 1.45 percent of variance in BMI. Despite the attention given to the *FTO* allele as the "fat gene," the variation directly accounted for by *FTO* SNP amounted to only 0.34 percent.[12]

Importantly, genetic factors could not have changed rapidly enough in a few decades to account for the marked increase in the prevalence of obesity and diabetes in this time period. Rather, changes in behaviors linked to diet and exercise, fostered by environmental changes such as "super-sizing" of food portions, an increasing proportion of meals eaten outside the home, the availability of "fast food," and agricultural policies that have increased the availability of cheap, low-nutrient food are more likely explanations.[13,14]

Social and Behavioral Determinants

A vast literature documents direct associations between sociodemographic characteristics (e.g., income, education, race/ethnicity), personality characteristics and affective states (e.g., hostility, anxiety, hopelessness, optimism, conscientiousness), and health behaviors (e.g. smoking, exercise, diet) with bio-markers that reflect dysregulation of cardiovascular, metabolic, and immune function (e.g., high blood pressure, excess inflammation, insulin resistance).[15,16,17] Perturbations in these biological processes contribute to the premature onset and progression of diseases, such as coronary heart disease, diabetes, and acute infectious illness, and to accelerated biological aging and mortality risk.

Many of the social and behavioral variables linked to disease risk are involved in the stress response. Stress experiences – especially those that are severe and chronic – produce changes in brain and body that promote disease onset and progression.[18] Acute fluctuations in cortisol, sympathetic nervous system activity, and metabolic hormones are adaptive for meeting short-term demands. However, they lay the foundation for chronic illness if persistent activation occurs. Neuroscience research is establishing the neural mechanisms through which stress shapes threat perception, the brain areas where emotion regulation resides, and how external stimuli are transduced from brain to body.[19]

Social conditions determine the acuity and frequency of stress exposures including major life events (e.g., loss of a loved one) and conflicts of daily living and their impact on mental and physical health. Physiological stress responses occur when individuals encounter situations in which threats or demands exceed their capacity to overcome or ameliorate them.[20] The chronic stress experienced by socially disadvantaged individuals who are subjected to more adverse situations and have fewer resources with which to address them help account for the pervasive health disparities associated with socioeconomic status and race/ethnicity.[17,21]

Getting into the Body

Research on the ways in which social conditions "get under the skin" to produce the social patterning of morbidity and mortality has identified several important pathways and mechanisms by which this may occur.

Cellular Aging

Accelerated cellular senescence is indexed by telomere length in immune cells. Telomeres are DNA-protein complexes that cap the ends of the chromosome, conferring chromosomal stability. In mitotic human cells, telomeres shorten with each successive cell division. When critically short, they can send cells into replicative senescence, causing cell cycle arrest and malfunction. Short telomere length in immune cells may serve as a marker, and possibly a mechanism, of earlier onset of diseases of aging.

In a break-through discovery in 2004, shorter telomere length was found to be associated with greater exposure to stress and psychological distress.[22] Subsequently, the stress-telomere link has been replicated in numerous samples with varied ages and demographic distributions. Shorter telomere length has also been observed among individuals characterized by elevated levels of hostility, depression, and/or low social support.[23-25] Accelerated telomere attrition has been associated with social determinants, including various indicators of socioeconomic status (e.g., income, education, employment status) and neighborhood social environment.[26-28]

Gene-Environment Interaction

Social conditions and the environment interact with one's genetic endowment to affect health. For instance, a history of childhood maltreatment predicts development of antisocial characteristics among individuals with the monoamine-oxidase risk allele[29] but not in those lacking it. Reliance on a candidate gene approach has been criticized since many failed to replicate in independent samples,[30] and the polymorphisms under investigation are rarely identified in GWA studies. The *FTO* gene is an exception; its effect on obesity has been shown to be modified by activity. A meta-analysis covering over 218,000 adults and nearly 20,000 children, showed that the risk of obesity among those with the *FTO* risk allele was reduced by more than 25 percent in individuals who were physically active but not among the sedentary.[31]

Epigenetics represents a biological mechanism through which social determinants can modulate the genome. Epigenetic processes, such as DNA methylation and chromatin modification, regulate developmental programming and cellular identity and serve as conduits through which the social environment can interact with the genome. It provides a mechanism though which the environment can regulate the transcriptional control of a gene that can persist for prolonged periods, even across generations. Seminal animal research demonstrated effects of early rearing on an animal's lifelong response to stress via epigenetic modifications of their hypothalamic-pituitary-adrenal (HPA) axis.[32] In humans, early evidence is showing variation in methylation patterns across levels of psychological stress, early life socioeconomic status, and sociodemographic factors.[33] Demethylation of the DNA near the glucocorticoid response element in *FKBP5* gene may reflect an epigenetic process that underpins a previously demonstrated gene by environmental interaction between a polymorphism in the *FKBP5* gene and prior exposure to childhood trauma in predicting rates of posttraumatic stress disorder (PTSD).[34,35]

Efforts to identify methylation status across the entire genome are underway, as are investigations to determine whether epigenetic modifications vary by tissue type. Although still a nascent area

of work, it has great potential for linking socially-patterned exposures with the "omics" inside the body. It underlines the need for a more systematic mapping of the "exposome," including social and behavioral exposures along with physical and chemical exposures.

Overcoming Obstacles to Integration

Despite the promising examples given above, obstacles remain to creating a full, integrated model of health determinants to inform care for individuals and achieve optimal health of populations. Spanning the full range of levels requires harmonizing different methods, languages, and values. While true in any cross-disciplinary collaboration, specific issues arise in collaborations between researchers studying social and behavioral determinants and those working on biological processes.

Different Perspectives, Valued Differently

An iconic cover of the New Yorker magazine depicts the New Yorker's view of the United States. It has fine detail of streets east of the Hudson River, with a largely undifferentiated flat plane west of the Hudson to the Pacific. A subsequent variant depicts the reciprocal Californian's view. Figure 1 shows the scientific equivalent of those two maps. On the top is the central figure from a New England Journal of Medicine (NEJM) article linking biomedical advances and deaths from cardiovascular disease.[36] Below it, from a letter to the editor, the same graph depicts public health advances.[37] Yet another letter, not shown here, shows the similarity of the graph to changing rates of cigarette consumption in the United States over the same time period.[38]

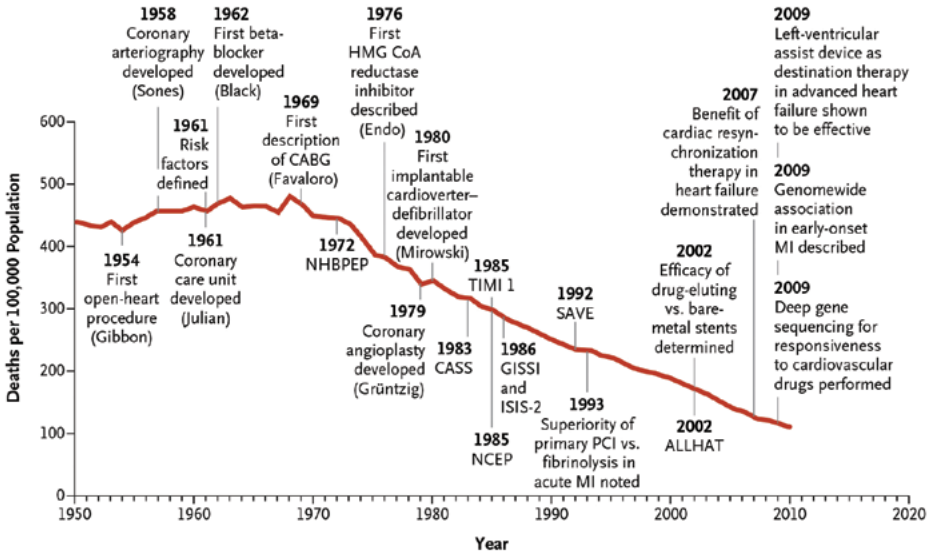


Figure 1a. Original depiction of biomedical contributions to drops in cardiovascular mortality

Source: Nabel EG, Braunwald E. A tale of coronary artery disease and myocardial infarction. New Engl J Med 2012;366:54-63. Used with permission.

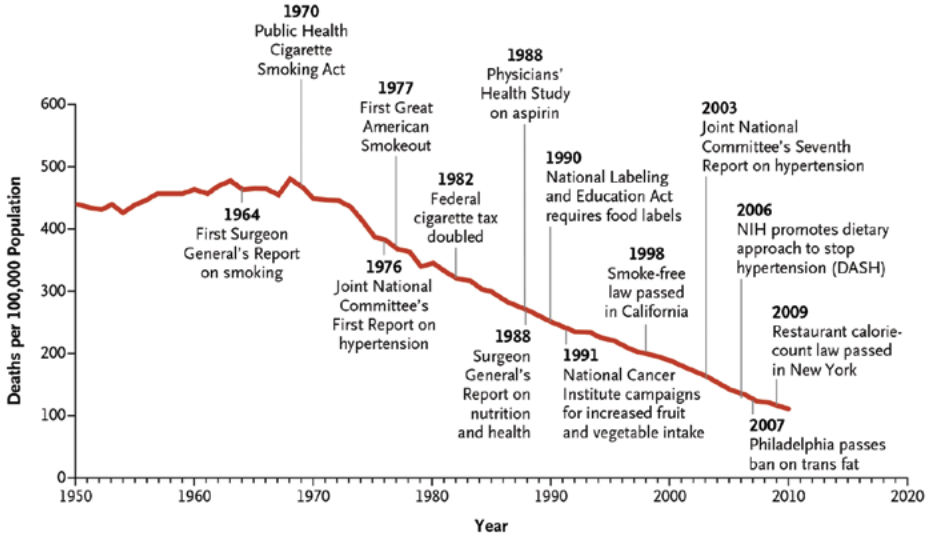


Figure 1b. The public health perspective

Source: Laing BY, Katz MH. Coronary arteries, myocardial infarction, and history. N Engl J Med 2012;366:1258-9. Used with permission.

The above examples show that the same data may be construed and understood in different ways. Under the right conditions, this divergence can spark new ideas and formulations. Discoveries often result from encounters across the bounds of disciplines where new ways of understanding findings can foster paradigm shifts. Ongoing interaction among researchers from different fields fosters such insights. The value of "water cooler" interactions was demonstrated in a study that found greater proximity among research collaborators, even within the same building, to be associated with a higher impact of their research.[39] Unfortunately, the organization of most universities and academic health centers does not facilitate interactions between social and behavioral scientists and biomedical researchers. "Desktop" research is generally done in different places than is "wet lab" research.

Proximity alone will not guarantee meaningful interaction. Unless there is mutual respect and valuing of the perspective of other disciplines, interactions are less likely to result in meaningful engagement. Although rarely explicit, the view that social and behavioral sciences are less valuable for understanding disease than are biological sciences may inhibit productive encounters. The omission of behavioral and public health landmarks associated with drops in cardiovascular disease noted above is, regrettably, not an isolated example of social and behavioral data going unnoticed. Such data may be ignored because they are assumed to be less rigorous or valuable than evidence emanating from the bench. The scientific community implicitly construes a hierarchy in the value and prestige of various types of science. As Jared Diamond noted, even the terms describing hard versus soft science reflect this valuation; in the extreme, the former are the only ones qualifying as real sciences.[40]

Social and behavioral sciences may be viewed as less valuable because of the relatively greater challenge in operationalizing and controlling the variables they study. The gold standard for making

causal inferences is the randomized experiment. However, the challenges of gaining sufficient control over social and behavioral factors to enable random assignment have fostered development of alternative methods to allow rigorous tests of predictions about causal associations. If the defining characteristic of good science is rigor in testing theoretical predictions against empirical findings, Diamond argued it would be more appropriate to view the sciences whose phenomena lend themselves easily to manipulation and control as the "easy" sciences, in contrast to those where the difficulty of doing so makes them "hard" sciences.[40] Thus, rather than our current classification of "hard" versus "soft" sciences, biological sciences might be characterized as the "hard" sciences, with social and behavioral sciences characterized as the "harder" sciences.

The consequences of placing a lower valuation on social and behavioral sciences are difficult to quantify since they mostly represent lost opportunities. For example, social scientists were not eligible to become members of the National Academy of Sciences (NAS) until the mid-1970s. As a result, the perspectives and methods of these disciplines were lacking, or at the very least, underrepresented, in the deliberations, reports, and culture of the major institution providing science advice to the Nation. Within the Institute of Medicine of the NAS, the membership sections appropriate for social and behavioral scientists are substantially smaller than are most other sections. Funding for social and behavioral research constitutes a small fraction of the National Institutes of Health (NIH) budget, and the vast majority of intramural research is in the biological sciences. Taken together, these result in a lack of visibility of social and behavioral scientists in positions that would allow them to shape the culture of research and encourage inclusion of social and behavioral measures and analyses into interdisciplinary investigations.

Changing Views

Recent developments within and outside of the research world may enhance the perceived value of social and behavioral data and encourage greater collaboration among the range of disciplines. The Clinical & Translational Science Awards (CTSAs) established by NIH have helped initiate interdisciplinary programs in over 60 institutions that aim to advance the translation of research findings from "bench" to "bedside" to "community." Social and behavioral issues are inherent aspects of the translation of findings at the bench into better care and better health. Acceptability, adherence, adoption of innovation, and diffusion of knowledge all involve cognitive, affective, and social factors that need to be understood and addressed. Insofar as Clinical and Translational Science Institutes (CTSIs) will be evaluated for renewal not only on the basis of their bench science discoveries, but also by their ability to move these discoveries into practice and improve individual and population health, the CTSIs should be motivated to include social and behavioral scientists in their work.

A similar "pull" for social and behavioral data is coming from the health care system as a result of the co-occurrence of high health care cost and relatively poor population health outcomes. Just as NIH leaders enacted policies establishing CTSIs to increase the yield on NIH investments in basic research, Congress—following a recommendation from its policy advisory group, the Medicare Payment Advisory Commission (MEDPAC)—mandated the establishment of "Accountable Care Organizations" (ACOs), which hold health systems and providers financially at risk for poor health

outcomes of the patients they serve. ACOs provide incentives for health providers and systems to address modifiable determinants of health. Given the powerful contribution of social and behavioral factors to health, health systems are motivated to address these factors in order to reduce utilization and cost. Solutions will not only require social and behavioral knowledge regarding effective translation of findings from the bench to the population, but also basic research on social and behavioral phenomena that are linked to onset or progression of disease.

Explicit examples of how social and behavioral understanding can improve diagnosis and treatment should accelerate this demand. Currently, for example, misclassifications regarding coronary heart disease (CHD) based on the Framingham Risk Score (FRS) results in both under- and over-treatment. The latter is especially likely in low socioeconomic status (SES) populations. Adding SES information to a patient's FRS has been shown to result in a better match of predicted cases of CHD with observed cases. The improved prediction exceeds that of adding information on genetic factors.[41,42]

The Diabetes Prevention Program provides another compelling example of the value of addressing behavioral determinants directly. This randomized trial assigned pre-diabetic patients to one of four arms, three of which involved drug interventions with standard lifestyle recommendations (Metformin in one and Rezulin in another, and placebo), and an intensive lifestyle intervention targeting exercise and diet. The Rezulin arm was ended early based on evidence of liver damage. Significantly fewer patients in the lifestyle intervention (4.8 percent) subsequently developed diabetes than did those assigned to Metformin (7.8 percent) or to placebo (11 percent).[43]

Finally, there is increasing interest in mining "big data" and using new analytic methods to test associations. By linking genomic information to biobanks, electronic health records, biosensor data, and environmental data, researchers are hoping to make new discoveries about the determinants of disease and the effectiveness of various treatments. One of the largest projects to date, the Research Program on Genes, Environment and Health (RPGEH) links data from electronic health records to biological specimens (i.e., saliva and blood samples) to enable genetic screening and to survey and geocode information to assess environmental and behavioral factors. Analyses may reveal direct effects of social and behavioral factors. However, even more important may be the interactions of these factors with environmental exposures and genetic vulnerabilities in determining disease risk. Such findings would fuel interest in gathering data on social and behavioral determinants and including these variables in analyses.

Increasing Visibility of Social and Behavioral Determinants

Social and behavioral determinants of disease may be underestimated in part because they involve factors that are more complex and less easily observed than are medical interventions. In addition, addressing these determinants often takes more personal effort; they are rarely modified once and for all by a given action. Understandably, people prefer simple, high-leverage solutions that involve a single action with long-lasting effects over those requiring ongoing action. Apart from considerations like cost, difficulty and side-effects, a one-time vaccination is likely to be preferred to a drug that has to be repeatedly taken. In turn, taking a drug is likely to be preferred to having to modify a behavior.

"Miracle cures" like the polio vaccine and the new drug, sofosbuvir (Sovaldi), for hepatitis C are examples of high-leverage discoveries. These fuel expectations that bench discoveries will provide solutions to health problems. Such solutions do not force us to change habitual patterns or gratifying behaviors. The quest for the "fat gene" to enable a pharmacologic solution to obesity is appealing; it promises a pill that would keep one's weight in check despite a poor diet or lack of exercise.

The occasional discovery of high-leverage cures can foster unrealistic expectations that similar cures will be found for a wide range of ailments. Research on the "availability heuristic" shows that people tend to overestimate the probability that a given event will occur if it is dramatic and concrete and can easily be brought to mind.[44] The vividness and appeal of a new vaccine or drug encourages people to anticipate other such break-throughs.

The limitations of "break-through" treatments are harder to grasp and get less attention. For example, crizotinib (Xalkori) is a drug for lung cancer patients whose tumors have the right genetic match. One media account described its use as having "commuted" the "death sentence" of a diagnosis for lung cancer in a 64-year-old woman.[45] This dramatic account failed to mention that only about 4 percent of lung cancer patients have tumors with the appropriate genetic composition to benefit from the drug, or that "progression-free added survival" is limited among those whose tumors have the right genetic match (7.7 months versus 3.0 months for traditional chemotherapy). While gaining even a few additional months is highly meaningful to a person with cancer, on a population basis, an addition of a few months for 4 percent of patients does not translate into a major overall advance in longevity.

In addition to overestimating the probability and value of future pharmacologic "cures," people may underestimate the obstacles to their effective use. In addition to cost (e.g., $84,000 for a typical course of Solvadi), obstacles include side-effects, behavioral demands for adherence, and possible interactions with other drugs. Currently, a quarter of newly prescribed drugs are never obtained by the patient,[46] and half of those obtained are not taken as prescribed.[47] The more prescriptions one has, the greater the chances for non-adherence, and this overload is likely to get worse as more drugs are made available for a variety of ailments.

The biases noted above not only affect the lay public, but researchers and funders as well. The preference for dramatic, easily visualized results may skew resources towards work on approaches that have a chance, no matter how slim, of resulting in a high-leverage cure. Such approaches may be favored, even if the likelihood of success and the population-wide impact are small, over research on more distal factors that may not eliminate a given problem but reduce risk across a much larger number of people. Explicit public debate is needed about the appropriate mix of research that varies on the probability and value of the potential results. In the meantime, social and behavioral researchers may want to more vividly describe the problems their work addresses and convey more compellingly to their scientific colleagues and the public alike how their findings are improving outcomes.

Market Forces

The rewards and demands of the market may affect what research questions are asked. Research that yields marketable products is likely to be favored over equally impactful research that does not produce a profitable commodity. The success of some genetically targeted drugs, along with substantial drops in the cost of gene sequencing, is likely to generate even greater interest in the genetic underpinnings of disease. Such interest is not problematic, but greater focus on genetic risk could shift attention and funding away from social and behavioral research.

Market forces associated with health care financing create demands for specific types of research. The predominant fee-for-service system provides little financial incentive for discovering or implementing social and behavioral interventions, and it has been difficult to make the business case for taking social and behavioral discoveries to scale. However, as noted earlier, changes in health care financing associated with the Affordable Care Act (ACA) and Accountable Care Organizations (ACOs) are pressuring health providers to seek effective ways to address the major drivers of disease and disability. They will, in brief, need to incorporate social and behavioral determinants of health into more traditional health care.

Finally, two trends will enable new types of research on social and behavioral determinants and create demand for the findings. One is the rapidly increasing development of sensors, mobile monitors, and digital communication. Their use will expand the reach of research on health determinants, and findings can inform improved design. Finally, increasing inclusion of social and behavioral data in electronic health records, as recommended in a recent IOM report,[48] will not only support analyses linking these factors to disease risk and treatment efficacy, but also will enable providers to address these factors in their patients.

Conclusion

Advancing both individual and population health will require health systems to address the whole range of the determinants of health from genetic inheritance to the society in which we live. To accomplish this, knowledge generation needs to be supported across the entire spectrum. NIH represents the major force shaping the health research agenda of our Nation. Its mission is to "seek fundamental knowledge about the nature and behavior of living systems and the application of that knowledge to enhance health, lengthen life, and reduce illness and disability."[49] Each aspect of this mission requires an understanding of the role of social and behavioral factors.

While NIH supports this range of research, it has allocated far fewer resources to the social and behavioral determinants of health than to biological substrates of disease. The establishment of CTSAs should direct greater attention to the whole range of determinants and how they play out in all aspects of translation from bench to application in the real world. However, if these efforts are going to be successful, specific policies and plans for incorporating social and behavioral sciences as an integral part of the spectrum will be required. Achieving optimal health for the population requires not just the biological knowledge, but also its union with social and behavioral knowledge. Our biological bodies develop in a physical and social context. So, too, must our science.

Acknowledgments

We are grateful to Stephanie Chernitskiy for her able assistance in preparation of this manuscript. The opinions presented herein are those of the authors and do not necessarily represent the position of the Agency for Healthcare Research and Quality, the National Institutes of Health, or the U.S. Department of Health and Human Services.

Authors' Affiliations

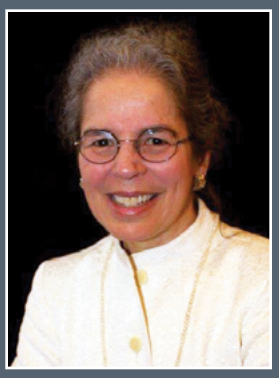
Nancy E. Adler, PhD, Center for Health and Community, Department of Psychiatry, University of California San Francisco. Aric A. Prather, PhD, Center for Health and Community, Department of Psychiatry, University of California San Francisco.

Address correspondence to: Nancy E. Adler, PhD, University of California, San Francisco, 3333 California St., Suite 465, San Francisco, CA 94118; email Nancy.Adler@ucsf.edu

References

1. Woolf SH, Aron L. U.S. health in international perspective: shorter lives, poorer health. Washington, DC: National Academies Press; 2013.
2. Bradley EH, Elkins BR, Herrin J, et al. Health and social services expenditures: associations with health outcomes. BMJ Qual Saf 2011;20(10):826-31.
3. McGinnis JM, Foege WH. Actual causes of death in the United States. JAMA 1993;270(18):2207-12.
4. Lee P, Paxman D. Reinventing public health. Annu Rev Public Health 1997; 8:1-35.
5. McGinnis JM, Williams-Russo P, Knickman JR. The case for more active policy attention to health promotion. Health Aff 2002;21(2):78-93.
6. Mokdad AH, Marks JS, Stroup DF, et al. Actual causes of death in the United States, 2000. JAMA 2004;291(10):1238-45.
7. Link BG, Phelan J. Social conditions as fundamental causes of disease. J Health Soc Behav 1995;Spec No: 80-94.
8. Galea S, Tracy M, Hoggatt KJ, et al. Estimated deaths attributable to social factors in the United States. Am J Public Health 2011;101(8):1456-65.
9. Smoking & tobacco use: health effects of cigarette smoking. Atlanta, GA: Centers for Disease Control and Prevention; 2014. Available at http://www.cdc.gov/tobacco/data_statistics/fact_sheets/health_effects/effects_cig_smoking/index.htm. Accessed November 4, 2014.
10. Frayling TM, Timpson NJ, Weedon MN, et al. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. Science 2007;316(5826):889-94.
11. Xi B, Takeuchi F, Meirhaeghe A, et al. Associations of genetic variants in/near body mass index-associated genes with type 2 diabetes: a systematic meta-analysis. Clin Endocrinol 2014;81(5):702-10.
12. Speliotes EK, Willer CJ, Berndt SI, et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. Nat Genet 2010;42(11):937-48.
13. Novak NL, Brownell KD. Role of policy and government in the obesity epidemic. Circulation 2012;126(19):2345-52.
14. Wallinga D. Agricultural policy and childhood obesity: a food systems and public health commentary. Health Aff 2010;29(3):405-10.
15. Everson-Rose SA, Lewis TT. Psychosocial factors and cardiovascular diseases. Annu Rev Public Health 2005;26:469-500.
16. Miller G, Chen E, Cole SW. Health psychology: developing biologically plausible models linking the social world and physical health. Annu Rev Psychol 2009;60:501-24.
17. Seeman T, Epel E, Gruenewald T, et al. Socio-economic differentials in peripheral biology: cumulative allostatic load. Ann NY Acad Sci 2010;1186:223-39.
18. McEwen BS. The brain on stress: how behavior and the social environment "get under the skin." In Kaplan R, Spittel M, David D (eds), Emerging behavioral and social science perspectives on population health. Rockville, MD: Agency for Healthcare Research and Quality; in press.

19. McEwen BS, Gianaros PJ. Central role of the brain in stress and adaptation: links to socioeconomic status, health, and disease. Ann NY Acad Sci 2010;1186:190-222.
20. Lazarus R, Folkman S. Stress, appraisal and coping. New York: Springer; 1984.
21. Adler NE, Stewart J. Health disparities across the lifespan: meaning, methods, and mechanisms. Ann NY Acad Sci 2010;1186:5-12.
22. Epel ES, Blackburn EH, Lin J, et al. Accelerated telomere shortening in response to life stress. Proc Natl Acad Sci U S A 2004;101(49):17312-5.
23. Verhoeven JE, Revesz D, Epel ES, et al. Major depressive disorder and accelerated cellular aging: results from a large psychiatric cohort study. Mol Psychiatry 2014;19(8):895-901.
24. Carroll JE, Diez Roux AV, Fitzpatrick AL, et al. Low social support is associated with shorter leukocyte telomere length in late life: multi-ethnic study of atherosclerosis. Psychosom Med 2013;75(2):171-7.
25. Brydon L, Lin J, Butcher L, et al. Hostility and cellular aging in men from the Whitehall II cohort. Biol Psychiatry 2012;71(9):767-73.
26. Carroll JE, Diez-Roux AV, Adler NE, et al. Socioeconomic factors and leukocyte telomere length in a multi-ethnic sample: findings from the multi-ethnic study of atherosclerosis (MESA). Brain Behav Immun 2013;28:108-14.
27. Needham BL, Carroll JE, Diez Roux AV, et al. Neighborhood characteristics and leukocyte telomere length: the Multi-Ethnic Study of Atherosclerosis. Health Place 2014;28:167-72.
28. Adler N, Pantell MS, O'Donovan A, et al. Educational attainment and late life telomere length in the Health, Aging and Body Composition Study. Brain Behav Immun 2013;27(1):15-21.
29. Caspi A, McClay J, Moffitt TE, et al. Role of genotype in the cycle of violence in maltreated children. Science 2002;297(5582):851-4.
30. Risch N, Herrell R, Lehner T, et al. Interaction between the serotonin transporter gene (5-HTTLPR), stressful life events, and risk of depression: a meta-analysis. JAMA 2009;301(23):2462-71.
31. Kilpelainen TO, Qi L, Brage S, et al. Physical activity attenuates the influence of FTO variants on obesity risk: a meta-analysis of 218,166 adults and 19,268 children. PLoS Med 2011;8(11):e1001116.
32. Meaney MJ, Ferguson-Smith AC. Epigenetic regulation of the neural transcriptome: the meaning of the marks. Nat Neurosci 2010;13(11):1313-8.
33. Lam LL, Emberly E, Fraser HB, et al. Factors underlying variable DNA methylation in a human community cohort. Proc Natl Acad Sci U S A 2012;109(Suppl 2):17253-60.
34. Klengel T, Mehta D, Anacker C, et al. Allele-specific FKBP5 DNA demethylation mediates gene-childhood trauma interactions. Nat Neurosci 2013;16(1):33-41.
35. Boyce WT. Epigenomics and the unheralded convergence of the biological and social sciences. In Kaplan R, Spittle M, David D (eds), Emerging behavioral and social science perspectives on population health. Rockville, MD: Agency for Healthcare Research and Quality; in press.
36. Nabel EG, Braunwald E. A tale of coronary artery disease and myocardial infarction. New Engl J Med 2012;366(1):54-63.
37. Laing BY, Katz MH. Coronary arteries, myocardial infarction, and history. Comment. N Engl J Med 2012;366(13):1258-9; author reply 1260.
38. Tarone RE, McLaughlin JK. Coronary arteries, myocardial infarction, and history. Comment. N Engl J Med 2012;366(13):1259-60; author reply 1260.
39. Lee K, Brownstein JS, Mills RG, Kohane IS. Does collocation inform the impact of collaboration? PLoS One 2010;5(12):e14279.
40. Diamond J. Soft sciences are often harder than hard sciences. Discover 1987:34-9.
41. Fiscella K, Tancredi D, Franks P. Adding socioeconomic status to Framingham scoring to reduce disparities in coronary risk assessment. Am Heart J 2009;157(6):988-94.
42. Thanassoulis G, Peloso GM, Pencina MJ, et al. A genetic risk score is associated with incident cardiovascular disease and coronary artery calcium: the Framingham Heart Study. Circ Cardiovasc Genet 2012;5(1):113-21.
43. Knowler WC, Barrett-Connor E, Fowler SE, et al. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. N Engl J Med 2002;346(6):393-403.
44. Tversky A, Kahneman D. Availability: a heuristic for judging frequency and probability. Cogn Psychol 1973;5:207-32.
45. Denver Post. CU system resets health care with $63M personalized medicine division (April 20, 2014). 2014. http://www.denverpost.com/news/ci_25600933/cu-system-resets-health-care-63m-personalized-medicine. Accessed November 5, 2014.
46. Fischer MA, Choudhry NK, Brill G, et al. Trouble getting started: predictors of primary medication nonadherence. Am J Med 2011;124(11):1081, 39-22.
47. Osterberg L, Blaschke T. Adherence to medication. N Engl J Med 2005;353(5):487-97.
48. Capturing social and behavioral domains in electronic health records: Phase 1. Washington, DC: National Academies Press; 2014.
49. U.S. Department of Health and Human Services: National Institutes of Health. About NIH - Mission. 2013. Available at http://www nih.gov/about/mission.htm. Accessed November 18, 2014.



Nancy E. Adler, PhD, is the Lisa and John Pritzker Professor of Psychology, in the Departments of Psychiatry and Pediatrics at the University of California, San Francisco, where she directs the Center for Health and Community. She directed the MacArthur Foundation Research Network on SES and Health, where she developed a widely used measure of subjective social status. Dr. Adler is a member of the American Academy of Arts and Sciences and the Institute of Medicine (IOM). She has received the James McKeen Cattell Award from the American Psychological Society, the Distinguished Scientific Award for the Application of Psychology from the American Psychological Association, the Marion Spencer Fay Award from the Institute for Women's Health and Leadership, the David Rall medal from the Institute of Medicine, and the Lloyd Holly Smith award and the Chancellor's Award for Advancement of Women from UCSF.

Aric A. Prather, PhD, is an Assistant Professor in the Department of Psychiatry, and Associate Director of the Center for Health and Community at the University of California, San Francisco. Prior to joining the faculty at UCSF, he completed a 2-year postdoctoral fellowship in the Robert Wood Johnson Foundation Health and Society Scholars program (RWJF-HSS). Dr. Prather is an affiliated faculty member at the UCSF Osher Center for Integrative Medicine, faculty member in the National Institute of Mental Health-sponsored postdoctoral program in Psychology and Medicine, program faculty in the RWJF-HSS program, and is on the executive board of the UCSF Center for Obesity Assessment, Study, and Treatment.

New Directions for Behavioral and Social Science Strategies to Improve Health

Daryn H. David, Michael L. Spittel, and Robert M. Kaplan

Abstract

This book offers an exciting collection of original chapters that highlight the varied contributions of the behavioral and social sciences to population health. The chapters also explore ways to increase the impact of these fields on innovations in health metrics and population health. In the following sections, we review some of the key points of this book and recommend some next steps for these continually evolving fields of research and practice.

Why the Behavioral and Social Sciences Matter

The vast majority of deaths in the United States and worldwide are due to non-communicable diseases (NCD) like cancer, cardiovascular disease (CVD), diabetes, and respiratory disease. The key determinants of these mortality rates are behavioral and social risk factors that include smoking, use of drugs and alcohol, poor diet, and lack of physical activity. According to the World Health Organization (WHO) World Health Assembly, eliminating these major risk factors, including unhealthy diet, tobacco use, and physical inactivity, could reduce the incidence of diabetes and CVD by 75 percent and the incidence of cancer by 40 percent. WHO estimates that these behavioral changes would reduce health inequalities by an estimated 50 percent.[1]

Further, a 2013 Institute of Medicine (IOM) report, U.S. Health in International Perspective: Shorter Lives, Poorer Health, showed that life expectancy for both men and women in the United States is below average in comparison with other high-income countries.[2,3] We could, however, achieve substantial gains in life expectancy in the United States by reversing the current trends responsible for declining health, including overweight/obesity, sedentary lifestyle, tobacco use, and excessive alcohol consumption, as well as the accidental deaths resulting from firearms, poisoning, and other risky behaviors.

One of the key questions tackled by this volume is exactly how the behavioral and social science communities could bring a range of tools to bear on developing interventions that have the potential to improve health and well-being. Another question concerns the optimal point at which to intervene, ranging from the national, to community, to individual levels. A final point focuses on the necessity of utilizing the findings of behavioral and social science research when building sustainable efforts that aim to advance population well-being.

Behavioral and social scientists study these problems using a variety of outcomes, levels of analysis, and empirical approaches. In their chapter, Williams and Purdie-Vaughns (pages 51-66) cite several efforts to enhance health through larger-scale reductions in income inequality. Zimmerman and colleagues (pages 347-384) highlight the importance of community-level education for well-being, while Kazdin (pages 305-326) explores the need for novel, large-scale modes of treatment delivery to adequately and efficiently meet the mental and behavioral health needs of larger segments of the population. Baldwin (pages 67-88) notes that youth, families, schools, and communities must all be engaged to successfully reduce risk factors for non-communicable diseases among youth. Holtgrave and colleagues (pages 199-216) highlight the contributions of the behavioral and social science community at the national level toward controlling the spread of HIV and providing necessary supports to those affected by the virus. Teutsch and his co-authors (pages 185-198) describe the use of a community-based health impact assessment in Los Angeles, CA, to address obesity and overweight; they provide evidence that interventions—such as requiring restaurants to identify how many calories are associated with each menu item—results in wiser food choices. The effectiveness of statewide policy interventions to reduce injury are described by Sleet and Gielen (pages 127-142), while Berkman (pages 39-49) explores the role that institutions, companies, and workplace policies can have on improving family health and well-being. Finally, Marteau and colleagues (pages 105-126) reveal how innovative population-level interventions that change the environment in which people live can affect the personal choices they make. In each of these chapters, the dynamic impact that the larger social, political, regulatory, treatment-delivery, and cultural milieus may have on individual decisionmaking, behaviors, and population health is brought to light.

The central importance of individual-level factors on behavior and well-being likewise should not be underestimated. The role of physical activity (Sallis and Carlson, pages 169-184) and smoking-related choices (Abrams and colleagues, pages 143-168) in maintaining, enhancing, or harming personal and population-level health are two examples. The interplay of biological and physiological factors with behavioral and social variables is likewise a key consideration when attempting to intervene to improve well-being. As the chapters by McEwen (pages 233-247) and Boyce (pages 219-232) suggest, the human body reacts to a range of pressures, and the social environment can have a major role in mitigating, attenuating, or accelerating what was once understood strictly as 'biological effects.' The traditional lines between behavioral and social science research and the human biological sciences are starting to blur. As Adler and Prather (pages 411-423) so aptly point out, future research and practice that aim to optimize population health will need to effectively draw on interdisciplinary perspectives and methods across the behavioral, social, and biological sciences to succeed.

Some behavioral and social interventions have the potential to increase life expectancy by years.[4] Unfortunately, however, the potential for enhancing health outcomes through social and behavioral interventions may be underappreciated. A few back-of-the envelope calculations concerning cigarette smoking help to illustrate this point. About 18 percent of adults in the United States smoke cigarettes[5] (see also Abrams et al., pages 143-168). Tobacco smoking contributes to an estimated 425,000 premature deaths per year. Each 1 percent reduction in cigarette use should result in about

21,500 premature deaths prevented. Even a modest 2 percent reduction in the smoking rate could have a large effect, perhaps rivaling that of completely eliminating breast cancer deaths. Further, meta-analysis suggests that minimal intervention in primary care would increase the marginal quit rate by about 0.94 percent.[6] Given the severe health consequences associated with smoking, this small percentage would translate into an impressive savings of about 373,000 quality-adjusted life years (QALYs). Brief counseling combined with nicotine replacement could increase the marginal quit rate by 8.4 percent and produce about 3,333,000 QALYs. On a population basis, these effects are profound in relation to many widely accepted health care interventions that have been evaluated using similar methods (see Russell, pages 251-270).

Unfortunately, these smoking interventions are used rarely, and at best they are offered to only about a quarter of smokers. If utilized more widely, these simple interventions may save many more lives than those lost to infections from antibiotic-resistant organisms (e.g., methicillin-resistant *Staphylococcus aureus* [MRSA]), measles, and a variety of other health problems that occupy national attention. The cost-effectiveness of these simple interventions is also much more favorable than almost all other primary care services that have been analyzed.[7]

The above example of staggering lost opportunity stands in sharp contrast to the success of another set of behavioral and social science-based interventions. The Centers for Disease Control and Prevention (CDC) has listed motor vehicle injury prevention among the 10 greatest public health achievements in the United States from 1900 to 2000.[8] This achievement was the result of multiple behavioral and cultural factors working in concert, including greater use of seat belts, better driver preparation, reduced frequency of driving while intoxicated, and improved road and automobile engineering. This example stands as a testament to the importance of integrating behavioral and social science considerations when working to improve population-level health.

Several chapters in this volume have explored the economic implications of investing in behavioral and social science interventions. Russell (pages 251-270) uses economic modeling methods to highlight which medical care alternatives may return the greatest health benefit for the investment made. She considers how many life years could be purchased for an investment of $1 million. In comparison to several medical, preventive, and environmental alternatives, smoking cessation consistently returns the most health benefits for the money invested. Further, Stewart and Cutler (pages 271-290) use economic analysis to demonstrate the importance of prioritizing behavioral and social interventions when addressing activities like smoking and dangerous driving. They found that changes in smoking between 1960 and 2010 improved quality-adjusted life expectancy in the United States by 1.42 years, and that cutting motor vehicle-related deaths resulted in another 0.43 year increase in life expectancy. Unfortunately, the costs saved by improvements in health behaviors such as smoking and reduced motor vehicle crashes were offset by increased costs associated with an uptick in obesity, poisonings, and firearm-related events. Finally, Frank and Glied (pages 291-304) discuss how health economics can yield improvements in mental health policy and practice. Taken together, these myriad chapters speak to the promise of behavioral interventions and point to areas for future research and investment.

The Path Forward

Perhaps it is best to conclude where we started. The mission of the National Institutes of Health (NIH) is "to seek fundamental knowledge about the nature and behavior of living systems and the application of that knowledge to enhance health, lengthen life, and reduce illness and disability." Through its programmatic development efforts and its efforts to build consensus across NIH Institutes and Centers, the NIH Office of Behavioral and Social Sciences Research (OBSSR) seeks to highlight and promote the crucial role that basic and applied research in the behavioral and social sciences plays in improving health and quality of life over the lifespan. Our other sponsor, the Agency for Healthcare Research and Quality (AHRQ), focuses on the translation of evidence-based research into the clinical practice of health care. AHRQ's mission is to produce evidence to make health care safer, higher quality, more accessible, equitable, and affordable and to work with other U.S. Department of Health and Human Services (HHS) agencies and other public- and private-sector partners to make sure that this evidence is understood and used. Several chapters in this book focus on AHRQ's efforts relevant to the establishment of new lines of evidence and to the efficiency and equity of health care.

OBSSR and AHRQ achieve programmatic success by working with a broad range of scientific constituencies to define mission-specific goals. The current plan to advance behavioral and social science research along with health services research includes several different 'pillars' or targets, several of which are described below.

Advance Research Methods and Approaches

In order to harness the new understanding of health determinants, we must encourage continued development of the changing paradigms of behavioral and social science research methodologies, approaches, and practices.

Findings in the behavioral and social sciences are often minimized or dismissed outright because they are correlational and not causal. The chapter by Pickett and Wilkinson (pages 15-38) takes up the challenge of estimating causation from observational data. The authors make a case that the relationship between social determinants (specifically, socioeconomic status) and health outcomes is, indeed, causal. This finding, and the rigorous methodology that the authors used to come to it, can be crucially important for shaping health and income policies. Preston (pages 89-102) also provides a detailed account of how social science research has contributed to better understanding of the social forces that have improved health through (1) accurate measurement, (2) attention to research design, and (3) a disciplinary focus on populations. These contributions include an evaluation of the impact of changes in living standards, medical care, and public health on improvements in morbidity and mortality.

The importance of utilizing robust methodologies has been further recognized by the OBSSR, which has been leading efforts in systems science for several years. Systems science methods can enable investigators to simultaneously examine the dynamic interrelationships of variables at multiple levels of analysis in complex systems (e.g., from cells to society), using modern computer modeling technologies. As one example, new methods based on systems science can help model the

complexity of the multiple interacting factors that contribute to health disparities, adolescent sexual behavior, and poor health outcomes (see Orr et al., pages 329-346).

In addition to systems science initiatives, behavioral and social science research requires a shift in focus from the individual to an emphasis on the community level and, concurrently, to the use of appropriate population modeling tools. "Big Data" and machine learning represent other important and useful areas of emergence. On the level of practice, fostering and supporting an informed workforce of physicians who can be on top of the most recent advances in the biomedical sciences while also effectively serving both patients and the larger community is also warranted (see Satterfield and Carney, pages 385-408).

Promote Interdisciplinarity, Team Science, and Collaboration

As emphasized by Adler and Prather (pages 411-423), there is also a critical need for an interdisciplinary research perspective that focuses on new ways to shift the boundaries and/or blur the lines between research disciplines. Such an approach should incorporate the methods of epidemiology and public health, sociology, psychology, environmental science, neurology, developmental biology, genetics, epigenetics, anthropology, political science, engineering, computer science, and other disciplines. One obstacle to this collaboration is the tendency to undervalue the behavioral and social sciences compared to biomedical science. Proactive communication among behavioral and social scientists and biologists is needed to spur collaboration. New scientific journals highlighting promising collaborations and professional societies recognizing the inter-connectedness between biology and behavior could play an important role, as could an expanded dialogue focused beyond the individual level to encompass communities and indeed whole populations.[3]

Not all investments in health are of equal value, but as many of the chapters in this volume have underscored, investments focused on health behaviors—including smoking, physical activity, and improved safety practices—represent a very good use of resources. Broadly speaking, behavioral and social science research should more fully embrace health economics to identify opportunity costs and return on investment for public health interventions that target behavior and social change. Health economic methods such as cost-effectiveness analysis (CEA), cost-benefit analysis (CBA) (see Russell, pages 251-270), and health impact assessment (HIA) (see Teutsch et al., pages 185-198) can prove increasingly helpful. The potential impact of these and other approaches is highlighted by the use of health economics research to spur improvements in the behavioral health care system.

This book is the product of an important collaboration between NIH and AHRQ. There are other important potential collaborators in population-level research, including the CDC and the environmental health community. Within NIH, the National Cancer Institute (NCI) is providing tools to advance interdisciplinary efforts. As one example, the NCI Team Science Toolkit is an online resource that serves to integrate and disseminate information and resources for engaging in tobacco science, as well as facilitating, supporting, evaluating, or studying team science.[a] Likewise, translational research opportunities may become available through NIH T4 awards and through

[a] More information about NCI's Team Science Toolkit is available at https://www.teamsciencetoolkit.cancer.gov/public/Home.aspx.

interaction with the National Center for Advancing Translational Sciences (NCATS) to identify opportunities for clinical translational science awards (CTSA).

Expand Visibility of Behavioral and Social Science Interventions

Behavioral and social scientists must increase their efforts to effectively communicate the importance of behavioral and social science interventions for public health. Researchers need to learn how much and what type of evidence is sufficient and of high enough quality to support public policy decisions; efforts are similarly required to better understand the methodologies and communication strategies that are needed to inform the public arena. The return on investment in behavioral and social science research and interventions, as noted in several of the chapters in this book, must be communicated to policymakers and the public alike. The value of behavioral and social sciences can be demonstrated by studies documenting how much additional heath benefit is added by interventions in areas such as diabetes prevention.[9]

Generally, investigators are not trained to make data accessible to nontechnical audiences, such as policymakers and the general public. Those interested in this undertaking may need to seek out opportunities to practice and hone these skills. Efforts are also needed to better communicate with regulatory agencies, such as the Food and Drug Administration (FDA). NCI is spearheading an effort in this area to develop an inter-agency initiative with FDA to inform regulatory decisions about tobacco products and their marketing. In keeping with AHRQ's efforts to ensure that the best scientific evidence is available for use in health care practice and that policymakers have access to high quality scientific evidence that can be used in their decisions about health care policy, the agency offers a variety of data resources for policymakers, including the Medical Expenditure Panel Survey (MEPS) and the Healthcare Cost and Utilization Project (HCUP).

Conclusions

Behavioral and social factors have profound effects on life expectancy and health-related quality of life. The magnitude of these effects has been under-recognized. The goal of extending the human life span and of improving health-related quality of life will require rigorous new research that establishes causal relationships, improves measurements, and systematically evaluates intervention strategies, as well as effective practices that emerge from this research.

This book deals less with specific diseases and more with the broader topical groupings such as the demographic, biological, behavioral, and policy determinants that affect health outcomes. We hope these contributions will stimulate efforts that can maximize the impact of behavioral and social science research on health outcomes that are meaningful to patients, families, and communities. By bringing greater visibility to behavioral and social science research, we hope to inspire a new generation of creative research and application that ultimately will produce better health for populations, both here in the United States and around the world.

Acknowledgments

The opinions presented herein are those of the authors and do not necessarily represent the official position of the Agency for Healthcare Research and Quality, the National Institutes of Health, or the U.S. Department of Health and Human Services.

Authors' Affiliations

Daryn H. David, PhD, and Michael L. Spittel, PhD, Office of Behavioral and Social Sciences Research, Office of the Director, National Institutes of Health. Robert M. Kaplan, PhD, Agency for Healthcare Research and Quality; formerly the Office of Behavioral and Social Sciences Research, Office of the Director, National Institutes of Health.

Address correspondence to: Robert M. Kaplan, PhD, Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850; email Robert.Kaplan@ahrq.hhs.gov.

References

1. Preventing chronic diseases: a vital investment. WHO Global Report. Geneva: World Health Organization; 2005.
2. Woolf SH, Aron L (eds). U.S. health in international perspective: shorter lives, poorer health. Panel on Understanding Cross-National Differences Among High-Income Countries. Washington, DC: National Academies Press; 2013.
3. Woolf SH, Aron LY. The U.S. health disadvantage relative to other high-income countries: findings from a National Research Council/Institute of Medicine report. JAMA 2013;309(8):771-2.
4. Kaplan RM. Behavior change and reducing health disparities. Prev Med 2014;68:5-10.
5. Smoking & tobacco use. Current cigarette smoking among adults in the United States (Web site). Atlanta, GA: Centers for Disease Control and Prevention; 2015. Available at http://www.cdc.gov/tobacco/data_statistics/fact_sheets/adult_data/cig_smoking/. Accessed July 9, 2015.
6. Kottke TE, Battista RN, DeFriese GH, et al. Attributes of successful smoking cessation interventions in medical practice. A meta-analysis of 39 controlled trials. JAMA 1988;259(19):2883-9.
7. Cromwell J, Bartosch WJ, Fiore MC, et al. Cost-effectiveness of the clinical practice recommendations in the AHCPR guideline for smoking cessation. JAMA 1997;278(21):1759-66.
8. Achievements in Public Health, 1900-1999 Motor Vehicle Safety: A 20th Century Public Health Achievement. MMWR 1999;48(18):369-74.
9. Dyson PA. Addressing noncommunicable disease at the population level: a focus on diabetes. Diabetes Manag 2014;4:153-63.

We thank the following individuals for their contributions to this book as peer reviewers:

Colin Baker
David Berrigan
Greg Bloss
Nancy Breen
Heather Cameron
Yen-Pin Chiang
Lee Eiden
Sheila Fleischhacker
Bob Freeman
Dionne Godette
Chris Hafner-Eaton
Martha Hare
Lynne Haverkos
Misty Heggeness
Carl Hill
Dionne Jones
Bill Lawrence
Amy Lossie
Patricia Mabry
Brett Miller
Carmen Moten
Peggy Murray
Susan Newcomer
Wendy Nilsen
Susan Persky
Barry Portnoy
Dan Raiten
Dianne Rausch
Dana Sampson
Michael Schoenbaum
Joel Sherrill
Bruce Simons-Morton
David Sommers
Erica Spotts
Shobha Srivasan
Steve Suomi
Derrick Tabor
Richard Troiano
Deborah Young-Hyman
Sheryl Zwerski

The editors gratefully acknowledge the exceptional efforts of four people who made this book possible. Mary Grady served as managing editor for the project. In this role, she set a very high bar for precision and accuracy and kept a close eye on the production schedule. Sabrina Liao served as the graphic designer for the project. With a long history of working with the Office of Behavioral and Social Sciences Research (OBSSR) at the National Institutes of Health (NIH), Sabrina produced the cover, formatted the copy, and coordinated the aesthetic aspects of the book. Joel Boches managed the print and Web production process for the book. Finally, our colleague Stephane Philogene from OBSSR worked with us on all aspects of the project, including selection of the authors, coordination between NIH and the Agency for Healthcare Research and Quality (AHRQ), and oversight of the production. Stephane's contributions were essential in guiding both the content and production of the book. We feel lucky to have such talented and committed partners in this effort.

In addition to NIH and AHRQ, the following NIH Institutes, Centers, and Offices participated in the development and peer review of the chapters in this book.

Center for Scientific Review
Office of the Director
Office of Disease Prevention
National Cancer Institute
National Human Genome Research Institute
National Heart, Lung, and Blood Institute
National Institute on Alcohol Abuse and Alcoholism
National Institute of Allergy and Infectious Diseases
Eunice Kennedy Shriver National Institute of Child Health and Human Development
National Institute on Drug Abuse
National Institute of Mental Health
National Institute on Minority Health and Health Disparities
National Institute of Neurological Disorders and Stroke
National Institute of Nursing Research

AHRQ Publication No. 15-0002
July 2015
ISBN 978-1-58763-444-4