Commentary: Medical Statistics, Austin Bradford Hill, and a Celebration of 40 years of *Statistics in Medicine*

Vern Farewell[1] and Tony Johnson[2]

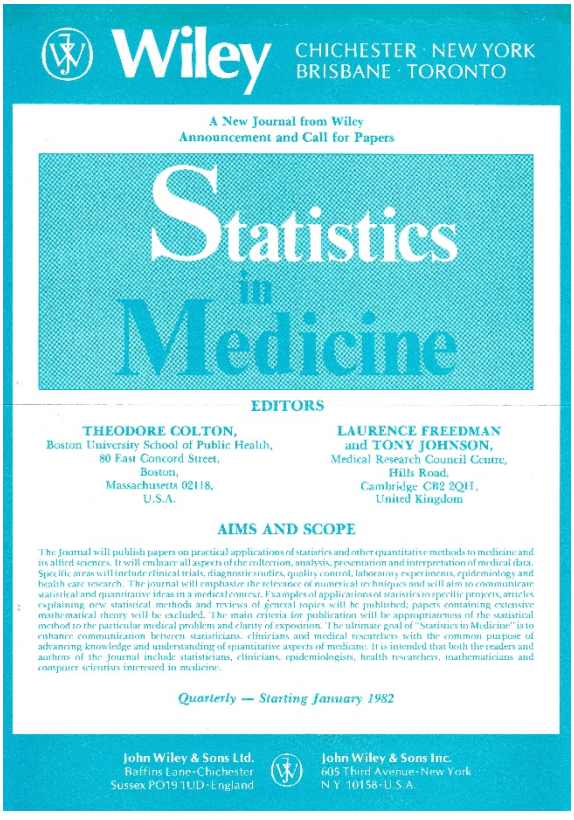
[1] *MRC Biostatistics Unit, University of Cambridge, UK*

[2] *MRC Clinical Trials Unit at UCL, London, UK*

Correspondence: V. Farewell, MRC Biostatistics Unit, Forvie Site, Robinson Way, Cambridge CB2 0SR. *e-mail:* vern.farewell@mrc-bsu.cam.ac.uk

1. INTRODUCTION

The first flyer (Figure 1) announcing a new journal, *Statistics in Medicine*, was distributed by the publisher John Wiley & Co., in May 1981, although the first issue did not arrive until July 1982, seven months after the scheduled launch date. The first volume was completed in February 1983 with 378 pages, with the last of the four issues being a celebration of the life and work of one of the great pioneers of medical statistics, Sir Austin Bradford Hill, on the occasion of his 85th birthday.


Wiley
CHICHESTER · NEW YORK
BRISBANE · TORONTO

A New Journal from Wiley
Announcement and Call for Papers

Statistics in Medicine

EDITORS

THEODORE COLTON,
Boston University School of Public Health,
80 East Concord Street,
Boston,
Massachusetts 02118,
U.S.A.

LAURENCE FREEDMAN
and TONY JOHNSON,
Medical Research Council Centre,
Hills Road,
Cambridge CB2 2QH,
United Kingdom

AIMS AND SCOPE

The Journal will publish papers on practical applications of statistics and other quantitative methods to medicine and its allied sciences. It will embrace all aspects of the collection, analysis, presentation and interpretation of medical data. Specific areas will include clinical trials, diagnostic studies, quality control, laboratory experiments, epidemiology and health care research. The journal will emphasize the relevance of numerical techniques and will aim to communicate statistical and quantitative ideas in a medical context. Examples of applications of statistics to specific projects, articles explaining new statistical methods and reviews of general topics will be published; papers containing extensive mathematical theory will be excluded. The main criteria for publication will be appropriateness of the statistical method to the particular medical problem and clarity of exposition. The ultimate goal of "Statistics in Medicine" is to enhance communication between statisticians, clinicians and medical researchers with the common purpose of advancing knowledge and understanding of quantitative aspects of medicine. It is intended that both the readers and authors of the Journal include statisticians, clinicians, epidemiologists, health researchers, mathematicians and computer scientists interested in medicine.

Quarterly — Starting January 1982

John Wiley & Sons Ltd.
Baffins Lane·Chichester
Sussex PO19 1UD·England

John Wiley & Sons Inc.
605 Third Avenue·New York
N.Y. 10158·U.S.A.

Figure 1: The first flyer announcing *Statistics in Medicine* (John Wiley & Sons Ltd, May 1981)

Since then the journal has grown at a prodigious rate to 30 issues per year with nearly 5,700 pages published in volume 38 (2018). Remarkably, the *Aims and Scope* of the journal remain almost unchanged from the original, and the ultimate aim to enhance communication between statisticians, clinicians and medical researchers has, we believe, been realised.

Such success could be achieved only with the combined aspirations, cooperation, and vision of many, many publishing staff, editorial staff, members of the Editorial Board, referees, and above all authors contributing to a truly international, global venture.

Now in 2021, by kind invitation of the Editors, Simon Day (UK), Els Goetghebeur (Belgium), Joel Greenhouse (USA) and Robert Platt (Canada) we celebrate forty years of this remarkable journal by commenting on the Alfred Watson Memorial Lecture delivered by Austin Bradford Hill, in 1962, after first reflecting on some issues raised in the first editorial of the journal and providing some details on Hill's career.

2. THE FIRST *EDITORIAL* (1982)

In that first Editorial[1]the Founding Editors indicated the dualistic nature of medical statistics, and pointed out that those who practice it sought to straddle two quite different branches of knowledge, in their quest to deploy the ideas, principles and methods of one, statistics, to stimulate deeper understanding of the other, namely medicine. This theme was also addressed by Hill in his Alfred Watson Memorial Lecture[2] some twenty years previously (see subsequent discussion) and, indeed the Editorial quoted his view of the statistician's role in a medical project, that "If he is a layman he must endeavour to learn and understand its medical as well as its statistical aspects and should, I believe, be as intensely absorbed and interested in the one as the other."[3]

In attempting to set out the compass of the Journal, the Founding Editors attempted to define the boundaries of medical statistics, commenting that "it constitutes more than the simple Boolean intersection of each of the two disciplines," but pointed out that definition was a particularly onerous task when the practitioners cut across several traditional disciplines having long heritages and stereotyped professional roles.[1] In the end they came to the conclusion that there were no clear boundaries and the only advice they could offer was to avoid the obvious extremes of statistical theory with little medical pertinence and substantive medical research with little statistical content. We believe that Hill would have endorsed this viewpoint. In the end the Founding Editors adopted a pragmatic approach and simply asked authors to send in their papers and let the Editors assess suitability for publication. Judging by the journal's success in terms of its publication history, this simple trust appears to have worked. However, the Founding Editors had a broader vision than this, and viewed the

venture as a success if, in some future evaluation of advances in our knowledge of medical statistics, at least some of the noteworthy contributions cited will have appeared among the pages of the Journal.[1] We believe that while this objective has been realised, the definition of the boundaries remains unresolved. In the celebration of 25 years, the Founding Editors pointed out that contemporary references still opted simply for "the application of statistics to medical problems,"[4] that is, a unidirectional relationship between the two disciplines, not a fusion of the two. Forty years on this does not appear to have changed substantially.

3. SIR AUSTIN BRADFORD HILL CBE FRS PhD DSc (1897 – 1991)

The first Editorial[1] portrayed Hill as a renowned pioneer in the discipline of medical statistics and the last issue of the first volume was dedicated to him with the Founding Editors' assertion that "we can think of no person with any knowledge of the field of medical statistics who when asked to associate one single name with the field, would respond with other than the name of Sir Austin Bradford Hill."[5] This collection of invited papers provided the reader with a perspective on the niche Sir Austin occupied as a founding father of the field of medical statistics, the scope and nature of his distinct contributions to the profession in particular and to medicine and science in general, the influence he exerted on contemporary research and some insight on the dimensions of his personality. Appropriately, this became the first *Special Issue* of the Journal, and the publishers, John Wiley, kindly produced a hardbound copy with the frontispiece signed by all the authors. From the correspondence that followed we know that Sir Austin was thrilled with it and offered his profuse thanks on several occasions.

The issue ended with a bibliography of his published works but, as was discovered later, it was incomplete. His son, I David Hill (1926-2015), who was himself a medical statistician and worked for the Medical Research Council Unit that his father had directed, pointed out that we had missed book reviews published under the authorship of ABH. This was rectified when the journal republished the list in 1993, along with the obituary from the *British Medical Journal*, written by his colleague, Sir Richard Doll.[6]

Hill wanted to follow his father, Sir Leonard Erskine Hill (1866-1952), into the medical profession.[7] Leonard Hill was a renowned physiologist who was Head of the Physiology Department at the London Hospital and then Head of the Applied Physiology Section of the

fledging Medical Research Committee (later Council) from 1914. His ambition was never realised as, on the outbreak of WW1, he joined the Royal Naval Air Service and was sent to the Greek Island of Imbros to fight against the Turks. While there, he contracted tuberculosis that was misdiagnosed and mistreated, and eventually he returned to UK, given little chance of survival, and subsequently awarded a 100% disability pension for life. Faced with a long convalescence, his ambition to take a medical degree could not be fulfilled, and instead he studied by correspondence for an external degree from the University of London, in economics, the only subject on the verge of science that could be done without any practical work.[7] It would be interesting to know if the Institute of Actuaries was aware of his background in economics; he did not mention it in his lecture.

Having gained his degree, Hill gravitated back towards medicine, and was offered a one-year post with the Industrial Fatigue Research Board (subsequently the Industrial Health Research Board (IHRB)), a semi-independent body under the auspices of the Medical Research Council (MRC).[7] This move must have been enabled by Major Greenwood (1880-1949) who at the time was Chair of the MRC Statistical Committee.[8] The Greenwood and Hill families lived close by one another in Essex, and in 1906 Leonard Hill had rescued Greenwood from an unwanted career in medicine, by giving him a post in his laboratory at the London Hospital.[9] In a sense then, Greenwood was repaying a family debt, but being a very good judge of character and ability, the appointment would not have been made on the basis of nepotism. Hill's achievements would justify Greenwood's judgment. To learn statistics Hill attended lectures at University College London given by Karl Pearson who had years earlier advised Greenwood about undertaking a career in medical statistics. On the opening of the London School of Hygiene and Tropical Medicine (LSHTM), in 1927, Greenwood was appointed Professor of Epidemiology and Vital Statistics as well as Director of the MRC Statistics Unit and he took Hill with him, still as an MRC employee. Hill would later succeed Greenwood in both posts.[8]

Hill was a gifted teacher and communicator both in his lectures at LSHTM and in his textbook *Principles of Medical Statistics* that emerged from a series of articles invited by the Editor of *The Lancet* in 1937. The publication history of this precious volume has been described elsewhere.[11] His lectures would inspire others, among them Richard

Doll, like Greenwood a physician more interested in medical research than in hospital medicine. In 1946 Doll was at the Central Middlesex Hospital engaged in research on peptic ulcers, supported and sponsored by the IHRB. He realised that his work would require statistical analysis, and along with his wife, Joan, he attended Hill's three months course at LSHTM.[12]

Figure 2
Three great pioneers of epidemiology and medical statistics in UK Professor Major Greenwood (1880-1949) , Sir Austin Bradford Hill (1897-1991), and Sir Richard Doll (1912-2005)

Hill recognised Doll's abilities and in 1948 offered him a post in the MRC Statistical Research Unit at LSHTM. Thus the great partnership of these two pioneers was commenced, Keating[12] proclaiming that "at the time, epidemiology was not a recognised medical discipline; only a few enlightened policy makers realised the importance it would have in the fight for the nation's health. Collectively Doll and Hill laid the basis for the explosive development of epidemiology by showing how the old science could be refurbished as a tool to discover the causes of non-infectious diseases. Neither man knew it yet, but their subsequent collaborations ushered in a new era of research." Of the three Directors shown in Figure 2, Greenwood and Doll were medically qualified, Hill was not, but all three had some statistical background and understanding, all three were epidemiologists, and two, Hill and

Doll were clinical trialists; Greenwood was not though he was an experimentalist.[9,13] Epidemiology and clinical trials constitute two of the main areas of activity within medical statistics as pointed out by Hill in his memorial lecture. Keating[12] actually goes further and even suggests an apostolic succession – perhaps the only one in British medical research – that started with Karl Pearson, who established the discipline of mathematical statistics at UCL, then passed to Major Greenwood, the first professor of epidemiology and vital statistics at LSHTM, to Hill, and subsequently to Doll".

We suggest that Karl Pearson was not the anchor and certainly he was not a medical statistician; he needed Greenwood's help in this field.[8] The succession more properly started with John Brownlee (1868-1927)[14], then passed successively to the other three with all of them, at some point, becoming Directors of the MRC Statistics Unit, that Brownlee first directed. In this context credit must be given to the MRC that presciently included statistics as one of its four foundation pillars in 1913; without it, it is unlikely that the succession could have taken place. The succession created by the MRC in UK now resides partly in the Universities with their own departments of medical statistics, epidemiology, and clinical trials, and partly in the MRC with their Units centred on epidemiology, biostatistics and clinical trials.

We should remember above all in celebrating the global achievements of *Statistics in Medicine* that the four figures portrayed in our succession above were all internationalists and that the foundations of statistics, epidemiology, medical statistics, and all the allied disciplines, were international collaborations.

We turn now to Alfred William Watson and the Memorial Lectures created in his memory.

4. SIR ALFRED WILLIAM WATSON KCB FIA (1870 – 1936)

4.1 Career

Figure 3
Sir Alfred William Watson
(1870 – 1936)

Sir Alfred William Watson[15] (Figure 3) was a British actuary and civil servant. In 1896, aged 26, he was appointed to the Treasury-controlled Rothschild Committee on Old Age Pensions, which had been set up to investigate the establishment of a state pension in the UK. Although the committee concluded in 1898 that none of the systems presented to them were viable, Watson continued to give advice about the writing of the legislation that became the National Insurance Act of 1911, the same Act that created the Medical Research Committee (later Council). He was then appointed chief actuary to the National Health Insurance Joint Committee and advised many government departments during World War I.

In 1917 he became Britain's first government actuary and was very influential in setting up the funding, through National Insurance, for the newly introduced state pension. He worked inside HM Treasury, aligned with the department's desire for more control over social policy and, indeed, no social policy could become law without his approval. Watson served on Royal Commissions on decimalisation in 1918 (not introduced in UK until 1971), and national health insurance in 1926. He served as President of the Institute of Actuaries from 1920 to 1922 and served on the Council of the *Royal Statistical Society* and the committee of the Reform Club.

4.2 The Memorial Lectures

The Alfred Watson Memorial Lectures were arranged as the result of a gift of £500 made anonymously to the Institute of Actuaries (now the Institute and Faculty of Actuaries) by a senior member as a tribute to the memory of Sir Alfred Watson, and with the intention that it should be used to advance the interests of the younger members of the profession. It is most likely that the lectures took place at Staple Inn Hall, a part-Tudor building in High Holborn in the City of London that was used for meetings of the Institute and Faculty of Actuaries, and is the last surviving Inn of Chancery (Figure 4). It was proposed that the Lectures be delivered by leaders of various important industries or by other persons of eminence.

Figure 4: Staple Inn, London

The first Lecture was given on 10 April 1946 by Sir William V. Wood, President of the Executive of the London Midland and Scottish Railway Company. Subsequent lectures were indeed delivered by leaders of industry but the seventh (1956) was given by a statistician MG Kendall (Professor of Statistics, University of London) on *Modern statistics in business and commerce,* and the eighth (1958), by Alex Comfort (Nuffield Research Fellow in Gerontology, University College, London) on *Mortality and the nature of age processes.*

Hill delivered the Tenth Alfred Watson Memorial Lecture which the Journal has republished in this issue and on which we now comment more directly.

5. Hill AND "THE MEDICAL STATISTICIAN"

The aim of the Alfred Watson Memorial Lectures was, to use Hill's figurative language, to find out what goes on "behind the shutters of the house next door". Thus, in the lecture, Hill was taking on the task of revealing "a little of the life of a statistician in medicine". And, one of his key messages was that a medical statistician's life must itself bring two houses together, medicine and statistics. As we remarked earlier, this is a view highlighted often by Hill and it is no surprise to see its emphasis in this lecture.

In considering Hill's remarks on this topic however, it is worth reflecting that Hill was brought into medical statistics by Major Greenwood, as has been outlined earlier, and that

Greenwood was medically qualified. As evidenced in a draft memorandum written by Greenwood in 1927,[16] the distinction between the medically qualified worker in medical statistics and others was important to Greenwood. For example, he wrote that "75% of the organisers [*of medical statistics in the UK]* have not been medical men" and "of new workers introduced during these ten years *[the last ten]*, the only demonstrably first rate additions, Miss Newbold and A.B. Hill, are non-medicals." He also wrote that "The present position is that the group of statisticians working for the *[Medical Research]* Council is much stronger on the purely statistical side than on the medical side and that there is no immediate prospect of strengthening it on the medical side" and that "I should prefer some young medical genius who would make great discoveries in applied statistics. But there isn't such a person and I can't wait."

By 1962, the route to a career in medical statistics rarely began with formal training in medicine but, given his training by Greenwood, the influence of his father, and his own convalescence, it is not surprising that Hill was very conscious of the value of medical knowledge to a medical statistician and made this a major aspect of his lecture. He writes, for example, "Unless himself medically qualified, and thus armed with the weapons of both disciplines, he must learn a great deal of medicine and even, I believe, absorb the doctor's way of thinking." This, of course, is difficult, and Hill does not expand on it. The difficulty lies in medicine having two essential components, one non-clinical that can be studied from textbooks and courses, and the other, clinical, that comes only from the training acquired by talking to, examining, and assessing patients, to reach a diagnosis. While Hill provides a fairly wide ranging description of medicine, distinguishing clinical medicine and public health and the different sources of knowledge, animal experimentation, experiments in man including clinical trials and observational studies, we will only consider some of his remarks on clinical trials and observational studies.

6. HILL AND "CLINICAL TRIALS"

6.1 SOME OBSERVATIONS

As expected, Hill has a number of important observations about clinical trials, and the statistician's role in trials, in his Alfred Watson Memorial Lecture. Some comments on the need for the medical statistician to "learn a great deal of medicine" are followed, with respect to trials in particular, with criticism of a narrow view of the statistician's role in trials that

assumes they have nothing to contribute to the "medical" aspects of a trial. Unfortunately, such attitudes can still be found today. The situation has, however, certainly dramatically improved since 1962 and it is now widely accepted that a statistician has to understand the details of a trial and can synergistically help to frame the questions being asked.

Hill also comments that "the statistician who is concerned with clinical trials must no less endeavour to acquire that similar code of honour that is second nature to the physician and surgeon." This is perhaps even more relevant today than when Hill gave his lecture as statisticians are now essential members of trial Steering Committees, Management Groups, and Data Monitoring Committees that have become central to trial organisation

There is also a useful reminder in Hill's lecture that there are always going to be questions about generalizability when designing and interpreting trials. Hill characterizes this both as "living dangerously" and necessary. He also raises the issue of assessing "intent to treat" and the historical background to this is of some interest.

6.2 INTENT TO TREAT

In his section on *The Statistician in Clinical Trials* Hill warns that randomization may be disturbed by selective withdrawal of patients who cease to take a treatment or are sensitive to it so that they have to be withdrawn. Of course, this is a problem well recognised in research papers published prior to 1962, and Hill stresses that "the medical statistician is aware of it." However, Hill continues, "The experiment is necessarily weakened - indeed we may on occasions have to access the value of an *intent* to treat rather than a treatment." The italics are Hill's, not ours, and clearly, he wants to draw attention to the underlying concept though he says nothing further about it.

Was an intent-to-treat analysis one of Hill's original ideas, or did he get it from someone else? We believe it was the latter. The first textbook to distinguish between explanatory and pragmatic clinical trials was published in French in 1970,[17] eight years after Hill's lecture, and written by Schwartz, Flamant, and Lellouch.[18] These authors admit in the Preface to their book that the distinction between the two approaches took at least five years to develop, perhaps longer.

In March 1959 Hill organised a five-day meeting on clinical trials in Vienna to which one hundred participants were invited.[19] It was intended that the introductory papers would be published in French only and not in English. However, demand was far higher than expected and eventually 23 papers (two of them summaries of conclusions) were published in both languages. Hill prepared the English version, published under the title, *Controlled Clinical Trials*,[19] while Schwartz, Flamant, Lellouch and Rouquette prepared the French version under the title, *Les essais thérapeutiques cliniques: méthode scientifique d'appreciation d'un traitement*.[20] Although there is no paper by the French authors in the proceedings, could the original ideas of intent-to-treat analysis have been presented and discussed with Hill at this meeting? The prospect is intriguing.

Despite Hill's simple statement of the concept it is worth remembering that it gave rise to many discussions, some heated, between statisticians and clinicians as the concept of pragmatic analysis entered successive specialties of medicine starting with cancer. Hill lived to witness much of this debate.

7. HILL AND "CAUSATION"

7.1 CRITERIA FOR CAUSATION

Because Austin Bradford Hill is now associated with a set of criteria to infer causation, as outlined in his 1965 paper, "The Environment and Disease: Association or Causation?",[21] it is interesting to see his remarks on this made in 1962. In particular, it is, retrospectively, somewhat puzzling to see that Hill makes reference to attempts "in the U.S.A. to lay down some decisive rules of evidence … which will determine our judgment of cause and effect in chronic diseases" and to indicate that he has "nothing of that kind to offer you for I am not convinced that we can simply reduce a wide variety of complex situations to a simple common denominator …".

The paper which prompted these remarks, by Yerushalmy and Palmer,[22] attempted to develop parallels to Koch's postulates for biological causation of infective diseases and had generated considerable comment, e.g. Lilienfeld[23] and Sartwell.[24] The thinking reflected in these publications likely influenced the "Criteria for the Epidemiologic Method" used for evaluating causal factors section of the 1964 Surgeon General's report on Smoking and Health.[25] Indicating that ``Statistical methods cannot establish proof of a causal relationship in an association" and that "The causal significance of an association is a matter of judgment", the criteria for this judgment were summarized as"

1) Consistency of the association
2) Strength of the association
3) Specificity of the association.
4) Temporal relationship of the association.
5) Coherence of the association.

These are very similar to Hill's 1965 Criteria to be assessed when non-experimental data show an association. These were given in his Presidential address to the newly formed Section of Occupational Medicine of the Royal Society of Medicine,[21] and are:

1) Strength of the association.
2) Consistency of the association.
3) Specificity of the association.
4) Temporal relationship of the association.
5) Biological gradient (dose response relationship)
6) Plausibility
7) Coherence
8) Experimental or semi-experimental evidence.
9) Analogy. (e.g. drugs in pregnancy given the experience with thalidomide).

The Surgeon General's fifth criterion could be viewed broadly to encompass the last 5 criteria of Hill, although it seems particularly useful to separate out the dose-response criterion. It is noteworthy that temporal relationship of the association refers to the *ordering* of events, specifically, that the *cause must precede the outcome,* and does not refer to *proximity* in time, although we suggest that causality is easier to establish when events are closer together.

The temporal sequence of the development of criteria could generate the speculation that Hill's negative view of Yerushalmy and Palmer[22] may have been influenced by their use of the link between smoking and cancer as an example in their development and that they presented, as illustration, some of the arguments that might be made against this causal link. They also, interestingly, included a footnote in their published paper indicating that the comments of some readers of their paper in manuscript suggested that the "exploratory

nature" of their discussion might have been overlooked and that they wished only to "illustrate the methodology" and "not to express a judgment".

Hill, and his colleague Richard Doll, had by the late 1950s published a number of epidemiologic papers with evidence for a link between smoking and lung cancer (*e.g.* Doll and Hill[26]) and the arguments for this link were increasingly infiltrating public consciousness. For Hill though, this came with personal cost. Notably, his relationship with R.A. Fisher which had been warm and friendly began to deteriorate towards the end of 1958[27] due to Fisher's antagonistic response to this research. Bodmer attributes this response to Fisher's strong belief that correlation does not prove causation, perhaps not being fully aware of the strength of the correlation and the unlikelihood of it being linked to a correlated cause, and also his libertarian view that people should be left to come to their own views during a time when there was growing publicity about the dangers of smoking. Naively perhaps[28] Fisher also believed he was free to give any advice on this matter even though he accepted payment for his advice to tobacco manufacturers and did not consider the view that others would take of this link. One can easily see how this might influence Hill's attitude to arguments similar to those made by Fisher even though Bodmer[27] claims that Hill was more bemused by Fisher's attitude to him than antagonistic. And, it must be recognized that others also made these arguments[29], for example, the distinguished psychologist, Hans Eysenck, who built on Fisher's arguments and suggested that personality might be a causal influence.[30]

By 1965 however, it seems that Hill had moved somewhat from the position he took in his Alfred Watson Memorial Lecture, although he did quote from the lecture in his 1965 Presidential address in his discussion of the plausibility criterion. There seems considerable overlap of approach between Hill in his Memorial Lecture and Yerushalmy and Palmer and just three years later Hill goes even further in identifying specific criteria to be considered. Nevertheless, even in 1965, Hill emphasizes that "What I do not believe … is that we can usefully lay down some hard-and-fast rules of evidence that *must* be obeyed before we accept cause and effect.

7.2 THE NEED TO CONSIDER CAUSATION

Of course, of much more importance than attempts to formalize the consideration of causation is Hill's emphasis in his lecture on the need to recognize that causation must be considered based on observational studies. Little could Hill have recognized that his

argument that some information must come from observational studies because "the medical statistician cannot convince thousands of his countrymen… to eat no animal fats, smoke no cigarettes, take regular exercise and drive their cars more slowly" would have to be nuanced in later years as large scale preventive health trials, such as the Women's Health Initiative[31] , were successfully launched and completed. Nevertheless, his primary message is as relevant today as it was in 1962: observational data is the only source of information on many epidemiological, and some clinical questions and we cannot simply say "association is not causation" or that "evidence is only statistical" in these situations. As the epidemiologist Weiss[32] argues, wholly consistently with Hill's arguments, "it is necessary to attempt to draw inferences of cause and effect, even from inevitably incomplete data, for the alternative is to make no inference at all, which would preclude taking preventive or therapeutic action".

From the statistician's perspective, it is also essential to take up the challenge of Aalen *et al*,[33] who wrote (page 348):

"One major danger of avoiding the subject of causality in statistical education and statistical literature, is that one never gets any insight into this fascinating concept, which has such an old history in philosophy and science. The fact is that statistics plays a major role in looking for causal connections in many fields and statisticians who know next to nothing about causality as a larger concept will be far less useful than they could have been."

These two quotes were also used by O'Keeffe *et al*[34] to motivate their application of Hill's criteria for causation to the question of whether joint tenderness and inflammation is causally related to subsequent permanent joint damage in patients with psoriatic arthritis, informed by Schweder's concept[35] of local (in)dependence between components of a composable finite Markov process. Hill's criteria were as informative for this complex clinical question in 2011 as they had been for studying the link between smoking and lung cancer 50 years previously.

Currently, causal inference represents a topic of considerable interest in the statistical literature. Graphical models (*e.g.* Didelez[36,37]) are playing a major role in establishing when causal effects are identifiable and can contribute much to discussions of causality. The drawing of these graphs may correspond somewhat to the "process of arranging, sorting and classifying a miscellaneous heap of objects" that Hill quotes Lord MacMillan as saying "affords a certain pleasure." However, there are perhaps two warnings that might be taken from Hill's Alfred Watson Lecture concerning this. The first is to ensure that the medical

statistician does not lose sight of the fact that they live in two houses. They need to be as concerned with the medical aspects of the causal questions as the statistical methods being employed. In this regard, the recent technical work might be seen as serving to help the formalization of the problem and aiding in the interpretation of findings. This should not, however, rule out consideration of more pragmatic and biological and medical orientated criteria, such as those associated with Hill. Secondly, while Hill was correctly adamant that observational data must be used for causal inference, his lecture also illustrates clearly his belief that, when possible (and it is not possible in all areas of science), an experiment, e.g. a clinical trial or otherwise, is a much more direct, and sound, basis for causal inference if, as mentioned earlier, one is prepared to "live dangerously" in terms of generalizability. This remains true but current work on emulated trials[38,39], probably not feasible in 1962, suggests that a simple dichotomy may no longer be a totally appropriate characterization.

On a more positive note however, it is to be hoped that an increasing emphasis on the need to consider causal inference when examining observational data, and the more robust inferences that can be drawn, will lead to fewer apparent inconsistencies between observational and trial data where opposite views are held based largely on the source of data used. This type of situation was manifested to one of us (VF) when submitting the subsequently published paper by Guyatt *et al*[40] which, in a meta-analysis, compared randomized trials and observational studies in adolescent pregnancy prevention. A marked difference was found with significant effects for prevention programmes only found in the observational studies. One referee felt that this paper was "polemical advocacy of unhelpful orthodoxy".

A contemporary, albeit slightly more complex, illustration of such a situation has arisen with the publication of a randomised trial of pulmonary metastasectomy for patients with colo-rectal cancer.[41] Although the trial did not achieve its planned recruitment, the estimated hazard ratio comparing patients randomized to surgery with a control group was 0.93 with a wide 95% confidence interval of (0.56 to1.56). More significantly perhaps, the estimated 5 year survival rates were 36.4% with a 95% confidence interval of (21.3% to 53.0%) and 29.6% with an confidence interval of (15.3% to 45.7%) respectively. At the present time however, this trial has had no impact on NICE guidelines[42] or the Society of Thoracic Surgeons' Expert Consensus Document on Pulmonary Metastasectomy[43] which are largely based on observational series of patients undergoing metastasectomy and an assumption that

5 year survival in patients with metastatic disease not so treated is zero, along with "expert opinion". It seems reasonable to expect that the sensible application of causal arguments, either to form conclusions or to indicate that none can be drawn, might have led to more helpful advice for patients and direction for further studies.

8. LITERARY REFERENCES

Hill makes several literary references in his lecture conjuring quotations from *Alice through the Looking Glass* (Lewis Carroll (1832–1898)), *A Portrait of Aunt Ida* (James Thurber (1894-1961)),[44] and the biblical *Book of Proverbs,* as well as, De Morgan on antimony, and Lord MacMillan on law. The quotation from Thurber, regarding the effect of being right 1 in 20 times, is a gem, and we wonder how many statisticians are aware of it, and indeed, how Hill discovered it? However, Hill was invalided out of the RAF at the age of twenty, a young man but with a pneumothorax, and he then spent two years in hospital followed by two years of convalescence[7]. Much of that time he must have spent reading, book after book, and he may have acquired a taste for literature. Perhaps, *The New* Yorker was a staple of his later reading. In any event, in 1962 he remembered Aunt Ida from Thurber's 1934 article and obviously appreciated Thurber's witticisms. For example, how good Aunt Ida's predictions actually were may be judged from her perspicacity, Thurber says, in "foretelling the sex of unborn children, when she was right about half the time."

9. THE STORY OF THE NICKEL WORKERS IN SOUTH WALES

Hill describes an excess of lung and nasal cancer in a group of nickel workers at a refinery in South Wales despite no causal agent being identified, and without knowledge of the underlying genetics and general environment of the population.[2] His argument relies essentially on use of mortality statistics, "his statistical evidence", derived from both the local and national routinely collected mortality data. The investigations actually started with Hill's report to the Mond Nickel Company in 1939 but were then taken up by his colleague, Richard Doll in the paper to which Hill refers.[45] This paper is a classic in the compilation of "statistical evidence" in its use of epidemiological methods based on local employment records, contact tracing, and derivation of expected mortality from nationally collected "routine statistics."

This is not quite the end of the story and Hill teases us a little when he mentions "the apparent disappearance of the abnormality in workers who entered the changed environment after 1923" without stating exactly what the changes were and why the cancers disappeared.[2] Or is he really teasing us? Is he perhaps just being more cautious in his "disentangling of the chain of causation forged in Nature."?[2] The answer it appears lay in the refining processes of the nickel ore that are described in detail by Morgan,[46] a paper that Hill references (no. 14 in his list).

Morgan[42] states, regarding the period before 1924, that "it is not unreasonable to draw the inference that carcinogenetic material may possibly have been present in the heated calcined dusts – possibly in their *arsenic* content." (Our italics). Perhaps the underlying cause was mentioned if it was raised in a question at the end of the lecture, or maybe in the publication by Richard Doll (no. 19), referred to as "in press"? We have not been able to trace this publication, and as far as we aware, Doll's next paper on this issue appeared in 1970.[47] Here he states that the epidemiological evidence "strongly suggested that the risk was associated with the preliminary steps of the process preceding the formation of nickel carbonyl." He continues, "Before 1925, a partly refined ore was imported from Canada, which contained sulphides of nickel and copper and a variety of other elements including cobalt, selenium, tellurium, sulphur, and several of the precious metals. Preliminary treatment removed the copper by leaching with an impure form of sulphuric acid containing significant amounts of arsenic. The partially purified material was then subjected to an exceedingly dusty process of calcination. Between 1920 and 1925 considerable changes took place in the refinery; less dusty calcinators and some personal protection against dust were introduced and arsenic-free sulphuric acid began to be used, But the essential chemical nature of the process, based on the unique properties of nickel carbonyl, was not altered." Hill's example of the nickel workers and the complex nature of the refinery process involving many mineral ores and acidic compounds, demonstrates the difficulties faced by epidemiologists and medical statisticians in identifying the causes of carcinogenesis, from "natural experiments? Could it be that this study, as well as those on lung cancer and smoking, lead him to conclude that formalised guidance on causality was, after all, a necessity?

10. CONCLUSION

Compared to the present day, the scientific world, and society more generally, were very different in 1962. Nevertheless, Austin Bradford Hill's Alfred Watson lecture remains a

timely reminder of the challenges of working in medical statistics. Both medical and statistical understanding must still jointly contribute to the research efforts of the medical statistician and, hopefully, be evident in both medical and methodological publications. This creates a synergy which, for example, was evidenced in the past when the HIV/AIDS epidemic drove the need for the development of new statistical modelling techniques to solve important medical problems which at the same time was reflected in the expansion of methodological work reported in *Statistics in Medicine* and elsewhere.

In a similar way, Hill's lecture remains timely in considering topics of current interest related to the design of clinical trials and causal inference. The republication of the lecture serves both to provide a valuable historical perspective on these topics and to highlight the ongoing relevance of many of the issues raised by Hill. It reminds us again of why Austin Bradford Hill is regarded as one of the most influential pioneers in the field of medical statistics and why to him, and to other pioneers, we owe so much.

ACKNOWLEDGEMENTS

We thank the Editors of *Statistics in Medicine* for their invitation to contribute to this celebratory issue. It is more than a privilege to do so, and we have enjoyed stepping outside our remit to reminisce on some of the early history of the Journal, history indeed, to which we were both so fortunate to contribute.

In addition, we thank David Raymont, Librarian, Institute and Faculty of Actuaries, London, for information about the Alfred Watson Memorial Lectures. We acknowledge Wikipedia as the source of historical information and photographs.

REFERENCES

1. Colton T, Freedman L, Johnson T. Editorial. *Stat Med.* 1982 1 (1) 1-3.
2. Hill AB. The statistician in medicine. *JIA.* 1962 88(2):178-191.
3. Hill AB. "The clinical trial". *New Engl J Med.* 1952;247(4):113-119.
4. Colton T, Freedman L, Johnson T. A celebration of 25 years. *Stat Med* 2007;26 (3): 473-483.
5. Colton T, Freedman L, Johnson T. Dedication. *Stat Med* 1982;1(4):i.

6. Colton T, Freedman L, Johnson T. Sir Austin Bradford Hill, 1897-1991. *Stat Med* 1993;12(8):795-808.
7. Hill ID. Austin Bradford hill—ancestry and early life. *Stat Med.* 1982;1(4):297- 300.
8. Farewell V, Johnson T. Major Greenwood (1880-1949): a biographical and bibliographical study. *Stat Med.* 2016;35(5):645-670.
9. Farewell, V, Johnson T. Major Greenwood's early career and the first departments of medical statistics. *Stat Med.* 2014;33(13):2161-2173.
10. Farewell V, Johnson A. The origins of Austin Bradford Hill's classic textbook of medical statistics. *J Roy Soc Med.* 2012;105():483-489.
11. Farewell V, Johnson A. The origins of Austin Bradford Hill's classic textbook of medical statistics. *J Roy Soc Med.* 2012;105(11):483-489.
12. Keating C. *Smoking Kills: The Revolutionary Life of Richard Doll.* Oxford: Signal Books, 2009.
13. Farewell V, Johnson A. Major Greenwood and clinical trials. *J Roy Soc Med.* 2017; 110(11):452-457.
14. Farewell VT, Johnson TL. Commentary: Dr John Brownlee MA, MD, DSc, DPH (Cantab), FRFPS, FSS, FRMetS (1868–1927), public health officer, geneticist, epidemiologist and medical statistician. *Int J Epidemiol.* 2013;42(4):935–943
15. CRVC. Memoir of Sir Alfred W Watson KCB. *JIA.* 1936;68(1):2-7.
16. Farewell V, Johnson T, Armitage P. 'A memorandum on the present position and prospects of medical statistics and epidemiology' by Major Greenwood. *Stat Med.* 2006;25(13):2161-2177.
17. Johnson A. Textbooks and other publications on controlled clinical trials, 1948 to 1983. JLL Bulletin: Commentaries on the history of treatment evaluation.
18. Schwartz D, Flamant R, Lellouch J. *L'essai thérapeutique chez l'homme.* Paris: Editions Médicales Flammarion, 1970.
19. Hill AB. *Controlled Clinical Trials.* Oxford: Blackwell, 1960.
20. Schwartz D, Flamant R, Lellouch J, *et al. Les essais thérapeutiques cliniques: méthode scientifique d'appréciation d'un traitement.* Paris: Masson and Co., 1960.
21. Hill AB. The environment and disease: association or causation. *PRSM.* 1965,58(5),295-300.
22. Yerushalmy J, Palmer CE. On the methodology of investigations of etiologic factors in chronic diseases. *J Chronic Dis.* 1959;10(1):27-40.

23. Lilienfeld AM. On the methodology of investigations of etiologic factors in chronic diseases. Some comments. *J Chronic Dis.* 1959;10(1):41-46
24. Sartwell PE. On the methodology of investigations of etiologic factors in chronic diseases. Further comments. *J Chronic Dis.* 1960; 11(1):61-63.
25. United States Public Health Service. *Smoking and Health: Report of the Advisory Committee to the Surgeon General of the Public Health Service*, Washington DC, Department of Health, Education and Welfare, 1964,
26. Doll R, Hill AB. The mortality of British doctors in relation to their smoking habits; a preliminary report. *BMJ.* 1954;1(4877):1451–1455.
27. Bodmer W. RA Fisher, statistician and geneticist extraordinary: a personal view. *Int J Epidemiol.* 2003, 32(6):938-942.
28. Box JF. *R. A. Fisher, the life of a scientist.* New York: Wiley, 1978.
29. Vandenbroucke JP. Those who were wrong. *Am J Epidemiol.* 1989;130(1):3-5.
30. Eysenck HJ, Tarrant M, Woolf M, et al. Smoking and personality. *BMJ.* 1960:1(5184):1456-1460.
31. Chlebowski RT, Aragaki AK, Anderson GL, et al. Dietary modification and breast cancer mortality: Long-term follow-up of the Women's Health Initiative randomized trial. *J Clin Oncol.* 2020;38(13):1419-1428.
32. Weiss NS. *Clinical Epidemiology; the Study of the Outcome of Illness*, New York: Oxford University Press, 1986.
33. Aalen OO, Borgan O, Gjessing HK. *Survival and Event History Analysis: A Process Point of View*, New York Springer, 2008.
34. O'Keeffe AG, Tom, BDM, Farewell VT. A case-study in the clinical epidemiology of psoriatic arthritis: multistate models and causal arguments. *J Roy Stat Soc S C.* 2011;60(5):675-699.
35. Schweder T. Composable Markov processes. *J Appl Probab.* 1970;7(2):400-410.
36. Didelez V. Graphical models for composable finite Markov processes. *Scand J Stat.* 2007;34(1):169-185.
37. Didelez V. Graphical models for marked point processes based on local independence. *JR Stat Soc B* 2008;70(1):245-264.

38. Hernan MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am J Epi* 2016: 183, 758-764.

39. Didelez V. Commentary: Should the analysis of observational data always be preceded by specifying a target experimental trial? *Int J Epid* 2016: 45, 2049-2051.

40. Guyatt GH, DiCenso A, Farewell V, et al. Randomized trials versus observational studies in adolescent pregnancy prevention. *J Clin Epidemiol* 2000;53(2):167-174.

41. Milosevic M, Edwards J, Tsang D, et al. Pulmonary metastasectomy in colorectal cancer: updated analysis of 93 randomized patients – control survival is much better than previously assumed. *Colorect Dis* 2020, doi:10.1111/codi.15113

42. Baum M, Fallowfield L, Farewell V, Macbeth F, Treasure T. NICE guidelines: management of colorectal cancer metastases. (Letter). *British Journal of Surgery* 2020. DOI: 10.1002/bjs.11789

43. Treasure T, Milosevic M, Macbeth F. Pulmonary metastasectomy in colorectal cancer challenges STS Consensus belief. (Letter) *The Annals of Thoracic Surgery* 2020. To appear

44. Thurber J. *A Portrait of Aunt Ida.* The New Yorker. 1934;10Nov:p19.

45. Doll R. Cancer of the lung and nose in nickel workers. *Brit J Industr Med.* 1958;15():217-223.

46. Morgan JG. Some observations on the incidence of respiratory cancer in nickel workers. *Brit J Industr Med. 1958;15:224-234.*

47. Doll R, Morgan LG, Speizer FE, Cancers of the lung and nasal sinuses in nickel workers. *Brit J Cancer.* 1970:24(4):623-632.