



Comprehensive interpretation of in vitro micronucleus test results for 292 chemicals: from hazard identification to risk assessment application

Byron Kuo¹ · Marc A. Beal² · John W. Wills^{1,3} · Paul A. White¹ · Francesco Marchetti¹ · Andy Nong¹ · Tara S. Barton-Maclaren⁴ · Keith Houck⁵ · Carole L. Yauk^{1,6} 

Received: 3 November 2021 / Accepted: 23 February 2022 / Published online: 21 April 2022
© Crown 2022

Abstract

Risk assessments are increasingly reliant on information from in vitro assays. The in vitro micronucleus test (MNvit) is a genotoxicity test that detects chromosomal abnormalities, including chromosome breakage (clastogenicity) and/or whole chromosome loss (aneugenicity). In this study, MNvit datasets for 292 chemicals, generated by the US EPA's ToxCast program, were evaluated using a decision tree-based pipeline for hazard identification. Chemicals were tested with 19 concentrations ($n = 1$) up to 200 μM , in the presence and absence of Aroclor 1254-induced rat liver S9. To identify clastogenic chemicals, %MN values at each concentration were compared to a distribution of batch-specific solvent controls; this was followed by cytotoxicity assessment and benchmark concentration (BMC) analyses. The approach classified 157 substances as positives, 25 as negatives, and 110 as inconclusive. Using the approach described in Bryce et al. (Environ Mol Mutagen 52:280–286, 2011), we identified 15 (5%) aneugens. IVIVE (in vitro to in vivo extrapolation) was employed to convert BMCs into administered equivalent doses (AEDs). Where possible, AEDs were compared to points of departure (PODs) for traditional genotoxicity endpoints; AEDs were generally lower than PODs based on in vivo endpoints. To facilitate interpretation of in vitro MN assay concentration–response data for risk assessment, exposure estimates were utilized to calculate bioactivity exposure ratio (BER) values. BERs for 50 clastogens and two aneugens had AEDs that approached exposure estimates (i.e., $\text{BER} < 100$); these chemicals might be considered priorities for additional testing. This work provides a framework for the use of high-throughput in vitro genotoxicity testing for priority setting and chemical risk assessment.

Keywords Micronucleus · Genotoxicity · Benchmark dose · Toxicokinetics · IVIVE · Point of departure

Introduction

In many countries, legislation is in place requiring genotoxicity testing before new chemicals can be manufactured or imported above a certain volume or based on intended

use. This poses a challenge as the rate of new chemicals production for commercial and industrial use continues to increase. Traditional in vitro genotoxicity testing tools such as the Ames assay, Comet assay and Micronucleus test, along with in vivo testing on animals, lack the necessary throughput to maintain pace with the emergence of new chemicals. An additional challenge is the lack of such data for the large numbers of chemicals that already exist on the market (Brendt et al. 2021; Shibai-Ogata et al. 2011). Moreover, the practice of risk assessment is moving towards reducing the use of animals (Krewski et al. 2020; Liu et al. 2019; Scholz et al. 2013; Wallace Hayes et al. 2020). Thus, genotoxicity testing and assessment would benefit from new in vitro approaches that can more efficiently test, classify, and prioritize chemicals for evaluation and regulation.

The ToxCast project of the Center for Computational Toxicology and Exposure (CCTE) at the U.S. Environmental

Byron Kuo and Marc A. Beal contributed equally.

Disclaimer: This article has been reviewed by the U.S. Environmental Protection Agency, Office of Research and Development and approved for publication. Approval does not signify that the contents reflect the views of the Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

✉ Carole L. Yauk
carole.yauk@uottawa.ca

Extended author information available on the last page of the article

Protection Agency (US EPA) aims to prioritize and assess hazard for environmental chemicals by applying screening technologies that are more cost-effective and higher-throughput than conventional tests (Knight et al. 2009). The datasets generated from these high-throughput screening assays are analyzed and modeled by novel computational approaches to predict potential chemical toxicities to human physiologies. The ToxCast framework has been implemented in phases (Dix et al. 2007; Kavlock et al. 2012; Kligerman et al. 2015). Phases I and II tested chemicals that had extensive toxicity data (primarily pesticides) and potential endocrine-related activities, respectively. Phase III added new assays and endpoints, as well as new chemicals, including endocrine disruptors, flame retardants, and mixtures. However, the number of assays in ToxCast that assess genotoxicity are minimal (<1%) and these tests are known to lack sensitivity (approximately 13%) for this endpoint (Hsieh et al. 2019). Thus, alternative and more sensitive approaches are needed to support the identification of genotoxic hazards as part of high-throughput screens.

The *in vitro* micronucleus (MNvit) test is one of the most sensitive and widely used tests for genotoxicity (Decordier and Kirsch-Volders 2006; Kirsch-Volders et al. 2011). The test detects DNA fragments or extra chromosomes that are caused by clastogenic and aneugenic mechanisms, respectively. These chromosomes or chromosome fragments manifest as observable micronucleus events in the cytoplasm of the daughter cells in interphase (Decordier and Kirsch-Volders 2006; Fenech 2000; OECD 2016). In a MNvit test, cells are treated with a chemical, and micronucleus frequencies are measured against solvent control treated cells (i.e., controls). Traditionally, micronucleus frequencies have been measured by manually counting micronuclei in interphase spreads using microscopy, which can be very time-consuming. In the past decade or more, automated flow cytometric methods have become increasingly popular for micronucleus analysis. The Organisation for Economic Cooperation and Development (OECD) test guideline indicates a positive response for MNvit when: (a) at least one of the test concentrations (of a minimum of three recommended plus controls) exhibits a statistically significant increase compared with the concurrent negative control; (b) a trend-test demonstrates that the increase is concentration-related in at least one experimental condition; and (c) the positive results are outside the distribution of the historical negative control data (OECD 2016).

The MNvit test is typically conducted in six-well plates and consists of at least three concentrations plus controls in duplicate (OECD 2016). A low or non-cytotoxic chemical is usually tested with concentrations that are at two- to three-fold intervals. Chemicals are generally tested at non-cytotoxic, moderately cytotoxic and cytotoxic concentrations, where the maximum concentration seeks to achieve

$55 \pm 5\%$ cytotoxicity. In the absence of cytotoxicity, a top concentration of 10 mM (or 2 mg/mL) or the highest soluble concentration is recommended. As presently conducted according to the OECD-compliant protocol with a limited concentration range, the MNvit test is not easily integrated into the current ToxCast framework where testing is conducted using 96- or 384-well plate formats across a broad concentration range. Here, MNvit testing was conducted using automated flow cytometry of 19 chemical concentrations to a maximum of 200 μ M ($n = 1$ per concentration) alongside solvent controls, making it feasible to integrate the MNvit assay with the standard battery of tests conducted within the ToxCast framework. Moreover, this study design is more suited to application of the benchmark concentration (BMC) approach for assessment of chemical potency. Specifically, BMC precision is augmented by distribution of the experimental replicates to as many concentrations as possible (Slob et al. 2005; Woutersen et al. 2001).

Applying BMC analyses to MNvit datasets for potency comparisons and prioritization is aligned with initiatives to move away from using genotoxicity testing exclusively for dichotomous hazard classification (positive/negative) (White et al. 2020). Indeed, genotoxicity endpoints are increasingly being recognized as *bona fide* toxicological endpoints for risk assessment and regulatory decision making (MacGregor et al. 2015; White and Johnson 2016; White et al. 2020). To use *in vitro* data in risk assessment applications, it is necessary to apply *in vitro* to *in vivo* extrapolation (IVIVE) models to derive administered equivalent doses (AEDs). AEDs, based on IVIVE and toxicokinetics models, provide human dose context to *in vitro* (geno)toxicity data by determining the correlation between intake dose and chemical disposition to various body compartments (Honda et al. 2019). Specifically, the AED is the estimated dose required to reach an internal concentration in the plasma that is equivalent to the concentration seen to be genotoxic using *in vitro* assays, and AEDs can be viewed as surrogate points of departure in the absence of *in vivo* data. Therefore, AEDs provide a more relevant potency ranking of chemical hazard than what is achieved using *in vitro* toxicity concentrations alone (Rotroff et al. 2010; Wetmore et al. 2012).

AEDs can be compared to human exposure values to derive bioactivity exposure ratios (BERs). BERs are analogous to the traditional margin of exposure used in risk assessment in that chemicals with a lower BER possess a higher potential for risk. For example, recent work has demonstrated the utility of *in vitro* BER data and IVIVE in establishing points of departure (PODs) that are protective (i.e., conservative) relative to traditional and apical PODs (Beal et al. 2022; Gannon et al. 2019; Paul Friedman et al. 2020). Furthermore, this work also demonstrates how BERs may be used in a tiered risk assessment framework. As a follow-up, Health Canada prepared a case study evaluating

the BERs of 41 compounds with previous risk assessments conducted under the Chemicals Management Plan, and this work determined that a BER of less than 100, or a \log_{10} BER of less than 2, identified all six of the non-genotoxic compounds classified as toxic under the Canadian Environmental Protection Act (Health Canada 2021). Thus, BER data can be a useful tool for prioritizing chemicals for further work based on non-genotoxic endpoints, including risk assessment as relevant. However, the application of HTK models for IVIVE of in vitro genotoxicity data to derive BERs has not yet been explored.

In this study, we analyzed the MNvit datasets of 292 chemicals from ToxCast and applied our extensive experience with this assay to produce a pipeline to evaluate the potential genotoxicity of the agents tested. Our proposed decision-making scheme classifies a chemical as positive, negative, or inconclusive by extensively evaluating the relative survival, non-parametric concentration–response trends, and BMC values for MN induction. This large dataset provides a unique opportunity to study the application of toxicokinetics to in vitro genetic toxicology data to determine the relationship between our high-throughput MNvit AEDs and PODs from traditional in vivo genetic toxicology and cancer studies. We also derived BERs to support quantitative evaluation of the tested compounds. BERs serve to identify the compounds with the highest potential for concern (i.e., AEDs approaching human exposure levels), making them priorities for further examination in scoping and risk assessment activities. We envision that this modernized approach can be used to complement test batteries applied in tiered testing strategies (Thomas et al. 2013, 2019), increasing the robustness of genotoxicity assessment, as well as accelerating the rate of chemical risk characterization.

Methods

Chemical selection

Chemicals were selected for testing by cross-matching samples available in sufficient quantity from existing ToxCast chemical libraries and inclusion in available databases or literature providing genotoxicity information (Richard et al. 2016). The genotoxicity data sources included the Chemical Carcinogenesis Research Information System (<https://www.nlm.nih.gov/databases/download/ccris.html>), the Priority-based Assessment of Food Additives database (Benz and Irausquin 1991), the Carcinogenic Potency Database (<https://www.nlm.nih.gov/databases/download/cpdb.html>), The National Toxicology Program's genotoxicity data in the Chemical Effects in Biological Systems database (ftp://anonftp.niehs.nih.gov/ntp-cebs/datatype/GENETOX_Genetotoxicity%20Toxicology%20-%20Micronucleus/), genotoxicity

summary of food additives in Japan (Yamada and Honma 2018), and the Benigni/Bossa rulebase for mutagenicity and carcinogenicity data used in Toxtree (Benigni et al. 2008). There were 292 chemicals selected that had sufficient quantity for testing, structural features (i.e., chemotypes) indicating genotoxic potential, and/or, previous investigations of genotoxicity. Additional candidates beyond the 292 chemicals were identified but testing of these chemicals was beyond the time and funds allocated to this project.

MNvit experimentation and datasets

Experiments with 292 chemicals, and their concurrent DMSO solvent controls, were performed by BioReliance, and MNvit datasets were prepared by the US EPA. The MN assay dataset contains 79 files in Microsoft Excel format. Each file contains data generated from two 96-well plates, with or without the addition of S9. A total of 292 chemicals were tested in the MN assay, and some of the chemicals were tested in more than one plate, for a total of 614 tests. The assays were performed on 18 different days, and this was used to generate batch-specific control 95th percentile reference values.

All agents were tested in Chinese hamster ovary (CHO) cells obtained from the ATCC (American Type Culture Collection) (Manassas, VA). CHO cells were cultured in complete medium: Ham's F-12 medium, supplemented with 10% fetal bovine serum, 100 units penicillin/mL and 100 μ g streptomycin/mL. Cells were prepared in T75 flasks. The culture medium was aspirated and cells with 60–90% confluence were washed once with calcium and magnesium-free phosphate buffered saline. The cells were trypsinized with 2 mL of 0.1% trypsin for approximately 4 min at 37 °C. After the cells were detached, the flasks harboring cells were diluted with medium to achieve a final density suitable to 96-well format: 5×10^4 cells/mL. A volume of 0.2 mL of the cell suspension was added to each well of the 96-well plates for chemical exposure with and without S9. Cultures were incubated at 37 °C for 16–24 h before treatment. Cells were tested with 19 concentrations (4.50, 5.63, 7.04, 8.80, 11.0, 13.7, 17.2, 21.5, 26.8, 33.6, 41.9, 52.4, 65.5, 81.9, 102, 128, 160, and 200 μ M) for 24 h in the absence of rat liver S9 at 37 °C. In the S9-activated study, cells were treated for 3–4 h in the presence of Aroclor 1254-induced rat liver S9 at 37 °C, washed, and then incubated in test article-free complete medium at 37 °C for 24 h. Replicate vehicle controls (1% DMSO) were included on each plate. Three concentrations of a standard genotoxicant, vinblastine (0.025, 0.0125, and 0.00625 μ g/ml) or cyclophosphamide (5.0, 2.5 and 1.25 μ g/ml), were tested in parallel as positive controls without and with activation systems, respectively. In total, eight wells of negative controls (DMSO),

six wells of positive controls (vinblastine – S9 or cyclophosphamide + S9) and four chemicals, each having 19 concentrations plus a control, were applied to each plate. Following the treatments, the treatment medium was aspirated and the plates were placed on ice for 20 min.

The flow cytometry-based MNvit and cytotoxicity assay was performed using the MicroFlow[®] kit (Litron Laboratories, Rochester, New York, USA) with more details related to protocol and compositions of buffers/solutions described previously (Avlasevich et al. 2006). A brief description of the MicroFlow methods is provided here. In each well of the plate, 50 µL of Nucleic Acid Dye A solution was added, and the plates were placed under a white fluorescent light (approximately 15 cm away). Samples were exposed to visible light for 30 min while on ice. Approximately 150 µL of 1X Buffer solution with 2% Fetal Bovine Serum was added to each well and the Dye/Buffer was then aspirated. Afterwards, 100 µL of Complete Lysis Solution I containing RNase and Nucleic Acid B (SYTOX Green) was added to each well, and the plates were incubated at 37 °C for 30 min, followed by 30 min at room temperature with shaking at 200–300 rpm. Complete Lysis Solution II, containing counting beads and Nucleic Acid B (SYTOX Green), was added to each well (100 µL), and plates were incubated for 30 min at room temperature. The plates were stored at 2–8 °C for up to 3 days prior to flow cytometry.

Micronucleus frequency, nuclei to beads ratio (i.e., relative survival) and hypodiploid events in treated cells were assessed using a BD FACSCanto II flow cytometer with 488 nm excitation laser. Samples were equilibrated at room temperature for approximately 30 min, and then the solvent control cell suspension was transferred to a 5 mL flow cytometry tube and analyzed on the cytometer. Unwanted events, including debris and apoptotic or necrotic events, were removed by gating out the regions containing these events (see Fig. 1 of Avlasevich et al. 2006 as an example). Micronucleus values were expressed as a percent frequency by dividing the number of events within the micronucleus-gated region by the number of events within the nucleated region and multiplying by 100. Hypodiploid nuclei were gated during flow cytometric analysis with the percentage frequency used to identify aneugenicity. Relative survival rate was obtained by dividing the nuclei to bead ratio in the sample by the ratio of the vehicle control and multiply by 100 (Avlasevich et al. 2006).

Counts of micronucleus and hypodiploidy events were reported as percentages relative to the control concentration's nucleated events. Fold-changes are the ratio of each concentration's micronucleus or hypodiploidy percentage over that of the DMSO control.

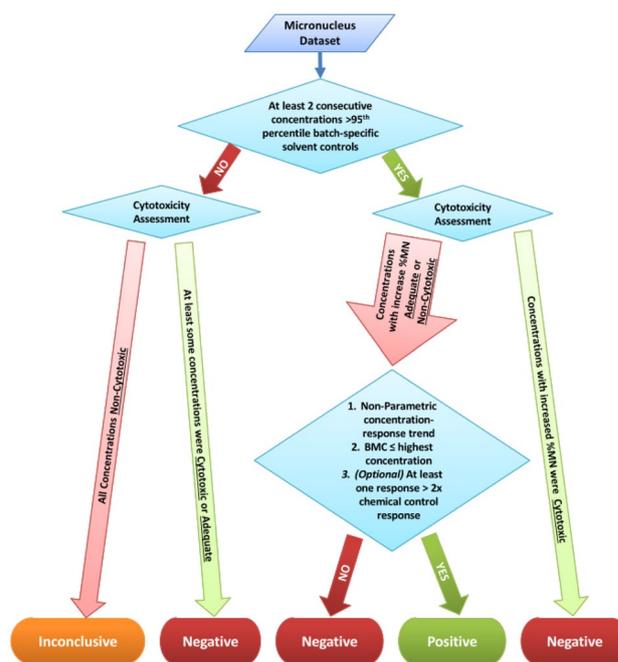


Fig. 1 Decision-making scheme for evaluating MN assay results (color figure online)

Data preparation and analysis

In vitro datasets were downloaded in Microsoft Excel formats, with each plate as an individual file. An open-source software, XLS to CSV Converter, (<https://cwestblog.com/2013/04/12/batch-excel-to-csv-converter-application/>) was used to convert each MS Excel sheet and file into a tab-delimited text file. Perl scripts were created to separately pool control datasets and treatment datasets from each plate for further analyses.

Data analysis was performed using the R Project for Statistical Computing (<https://www.r-project.org/>) under R Studio (<https://www.rstudio.com/>). Non-parametric trend tests to determine concentration–response were applied using the Jonckheere test function (<https://www.rdocumentation.org/packages/clinfun/versions/1.0.14/topics/jonckheere.test>) of the *Clinfun* package in the R statistics environment (Seshan 2018).

Data quality assessment

Solvent control (DMSO) and positive controls (cyclophosphamide and vinblastine) were separately parsed and pooled to individual files. Data were imported into R for analysis and assessment. For positive control datasets, plots of MN frequency against relative survival rates were generated for each treatment concentration, 1.25, 2.5 and 5 µm for cyclophosphamide (+ S9) and 0.00625, 0.0125 and 0.025 µm for

vinblastine (–S9), to identify distributions and outliers. For solvent controls, four histograms were plotted for MN and hypodiploidy frequencies, in the presence or absence of S9, to identify outliers. 95th percentile values for MN and hypodiploidy frequencies were computed for each batch date. Using this approach, one plate was identified as an outlier and removed, i.e., 20131009 Plate 2.

Benchmark concentration (BMC) analysis

BMC modeling was conducted using the PROAST package (version 61.5) for the R statistical environment. PROAST was selected for BMC analysis because it is capable of modeling BMC without replicates. The benchmark response (BMR) for %MN and %hypodiploid was based on first calculating one standard deviation (SD) of the mean control values, excluding the outlier, noted above. One SD was an approximate 30% and 60% change in %MN and %hypodiploid, relative to control, respectively. Thus, BMR values of 30% and 60% increases were used for %MN and %hypodiploid. For each dataset (i.e., chemical ± S9), concentration–response modeling was performed on concentrations categorized as ‘non-cytotoxic’ and ‘adequate’ (described in more detail below) using the model 5 (Hill and Exponential) model families, based on the European Food Safety Authority (EFSA) Science Committee’s recommendations for continuous data (EFSA 2009, 2017). Model selection was informed using the Akaike Information Criterion (AIC). For each compound, we also retained information describing the two-sided 90% confidence interval (e.g., bounded by the BMC lower limit (BMCL) and BMC upper limit (BMCU)).

Micronucleus assessment

A decision-tree approach was developed to assess MN based on %MN against batch-specific solvent controls, cytotoxicity assessments, and concentration–response tests (Fig. 1). Within this decision tree, cytotoxic refers to concentrations at which relative survival was <40%; adequate refers to levels of cytotoxicity between 40–60% (i.e., a requirement in the OECD MNvit test guideline, unless the top concentration tested is 10 mM, 2 mg/mL or 2 µL/mL); and non-cytotoxic refers to test concentrations with relative survival >60%.

A positive assessment for MN required: (a) that two consecutive concentrations with a %MN greater than their respective ± S9 95th percentile of batch-specific solvent controls; (b) that the observed increase in %MN occurred at non-cytotoxic or adequate cytotoxicity levels; (c) a concentration–response trend, which was assessed using a non-parametric test; and (d) a BMC (BMR30) value that was lower than or equal to the highest concentration used in BMC modeling. We also assessed the outcome if an additional filter requiring at least a two-fold increase above controls for at

least one concentration was also applied, i.e., as a more conservative approach. If conditions (a) and (b) were met, but there was no concentration–response by the non-parametric test (i.e., condition (c)), or the chemical had a BMC > highest test concentration [i.e., condition (d) whereby the model suggested that the chemical did not induce the BMR within the given test concentrations], the assessment was classified as negative. When a %MN increase was observed, but only within overtly cytotoxic concentrations (i.e., relative survival rate <40%), the chemical was also classified as negative.

Negative calls were made if the chemical did not have at least two consecutive %MN above the 95th percentile of controls, and the top concentration tested achieved the minimal required level of cytotoxicity (i.e., either ‘cytotoxic’ or ‘adequate’). In cases where the test did not achieve the minimum specified top threshold for cytotoxicity, the negative result was considered inconclusive.

Hypodiploidy (aneugenicity) assessment

Two approaches were used to identify agents that induce hypodiploidy. In the first approach, the decision scheme outlined in Fig. 1 was applied, but the %MN was replaced with %hypodiploid and a BMR60 was used. The second approach applied the method suggested by Bryce et al. (2011). In this approach, a positive hypodiploidy call was made if fold change relative to concurrent controls in %MN was greater than three, and fold change for %hypodiploid cells was greater than ten. In this approach, the fold-change values were directly obtained from the datasets obtained from EPA and the call is either positive or negative.

Derivation of administered equivalent doses (AED)

IVIVE modeling of BMC in µM to determine AEDs in mg/kg bw/day was performed using the HHTK package v1.10 (Pearce et al. 2017) in R following methods used previously (Beal et al. 2022; Paul Friedman et al. 2020). Specifically, the three compartment steady-state model (“3compartmentss”), modified from Wetmore et al. (2012, 2015); Wetmore (2015), was used to calculate the steady-state concentration in the plasma (C_{ss}) at a constant infusion dose rate of 1 mg/kg bw/day. Assuming a linear relationship, IVIVE can be used to model the dose rate (i.e., AED) required to reach an internal C_{ss} equal to the BMC shown to be genotoxic in the CHO cells. The AED was derived using the formula:

$$\text{AED} = \text{benchmark concentration } (\mu\text{M}) \times \frac{1 \frac{\text{mg}}{\text{kg}}}{C_{ss}(\mu\text{M})}$$

The `calc_mc_css()` function in HHTK was used to model the C_{ss} using parameters `well.stirred.correction = TRUE` and `output.units = “µM”`. To run the 3compartmentss model and

return units in mg/kg bw/day, the in vitro parameters consisting of intrinsic clearance (Cl_{int}) and fraction unbound in the plasma protein (F_{up}) were needed. In addition, the physico-chemical properties consisting of molecular weight and the octanol/water partition coefficient ($\log P$) were also required. These data were unavailable for many of the compounds tested in this study; therefore, in silico predictions were used to obtain predictions for these parameters (Supplemental File 1). Previous work has demonstrated that ADMET Predictor can provide reliable estimates of in vitro parameters for HTTK modeling to derive stable estimates of C_{ss} (Pradeep et al. 2020). Thus, ADMET Predictor version 9.5 was used to provide F_{up} percentage (hum_fup%) and human liver microsomal clearance (CYP_HLM_CLint) by CASRN (CAS number). CYP_HLM_CLint ($\mu\text{L}/\text{min}/\text{mg}$) were adjusted to Cl_{int} HTTK units ($\mu\text{L}/\text{min}/10^6$ cells) by dimensional analysis using scaling factors (Barter et al. 2007) that have been previously applied (Sipes et al. 2017):

$$Cl_{int} = \text{CYP_HLM_CLint} \times \frac{32 \text{ mg of microsomal protein}}{\text{g of liver}} \times \frac{1 \text{ g of liver}}{99 \times 10^6 \text{ cells}}$$

The ChemmineOB R package (Horan and Girke 2020), was used to interface with the OpenBabel C++ project (O'Boyle et al. 2011) to provide molecular weight and $\log P$ estimates for each chemical. The required parameters for IVIVE were imported into HTTK using add_chemtable() function. The overwrite parameter was set to FALSE to prioritize any existing experimental data in the HTTK database over the supplemented in silico parameters.

Comparison of administered equivalent doses to traditional in vivo PoDs

Traditional in vivo genotoxicity data were compared against AEDs to build confidence in the developed workflow and the utility of AEDs in the context of risk assessment. Specifically, traditional in vivo PODs in mg/kg bw/day were extracted from the ToxValDB, which is a highly-structured database containing cancer and genotoxicity data in addition to many other endpoints (Williams et al. 2017). The AEDs derived in this study for all of the compounds classified as positive clastogens or aneugens in MN assay were compared against the traditional PODs where the study type is listed as “cancer” or “genetox.”

Determination of bioactivity exposure ratios (BERs)

BERs were determined using the most recent ExpoCast exposure estimates (Cohen Hubal et al. 2010; Wambaugh

et al. 2013) downloaded from the CompTox Chemicals Dashboard (Williams et al. 2017) on October 22, 2020. Both the median and 95th percentile exposure estimates were used to derive BERs. BERs were reported as the \log_{10} BER on the logarithmic scale ($\log_{10}\text{AED} - \log_{10}\text{Exposure}$), and as BER on the arithmetic scale (AED/Exposure).

Results

Data quality assessment using positive controls

Positive controls, cyclophosphamide with S9 and vinblastine without S9, were pooled according to their concentrations, and scatter plots were generated of %MN against their respective relative survival rates (Fig. 2). In general, positive controls behaved as expected demonstrating declines in relative survival and increases in %MN with increased variance in %MN as concentration increased. Of the 158 data points in each concentration, one outlier (0.6%) was identified for cyclophosphamide for each of 2.5 μM and 5 μM . For vinblastine, eight (5%) outliers were identified for all three concentrations. These all belong to the same date batch (20130903 plates 1, 2, 3, and 4), indicating a potential batch effect only for vinblastine. Overall, no clear plate-effects were observed.

Data Quality Assessment Using Solvent Controls

Solvent controls (DMSO) were pooled from each plate and grouped into DMSO – S9 and DMSO + S9 to investigate distributions and outliers (Fig. 3). In the – S9 %hypodiploid analysis, four outliers were observed at 3.2%, 15.5%, 28.3%, 61.8%, and 132.5%, the latter four were removed from the figure; the respective plates were 20130911 Plate 5, 20130923 Plate 5a, 20130911 Plate 6, 20130923 Plate 5a, and 20130923 Plate 5a. A plate effect was observed in the + S9 distributions for both %MN and %hypodiploid cells (colored in red). Further investigations revealed that 20131009 Plate 2 contributed to these outliers. These outliers were subsequently removed from further analysis.

One of the criteria to assess MN is to compare the %MN of a chemical at every concentration to reference values generated from batch-specific solvent controls. To compute these reference values, we pooled the control datasets within dates. Each date was considered a batch, and their control %MN and %hypodiploidy frequencies were used to compute the 95th percentile values (i.e., 95% confidence interval of the population) (Table 1). This revealed that %MN for – S9

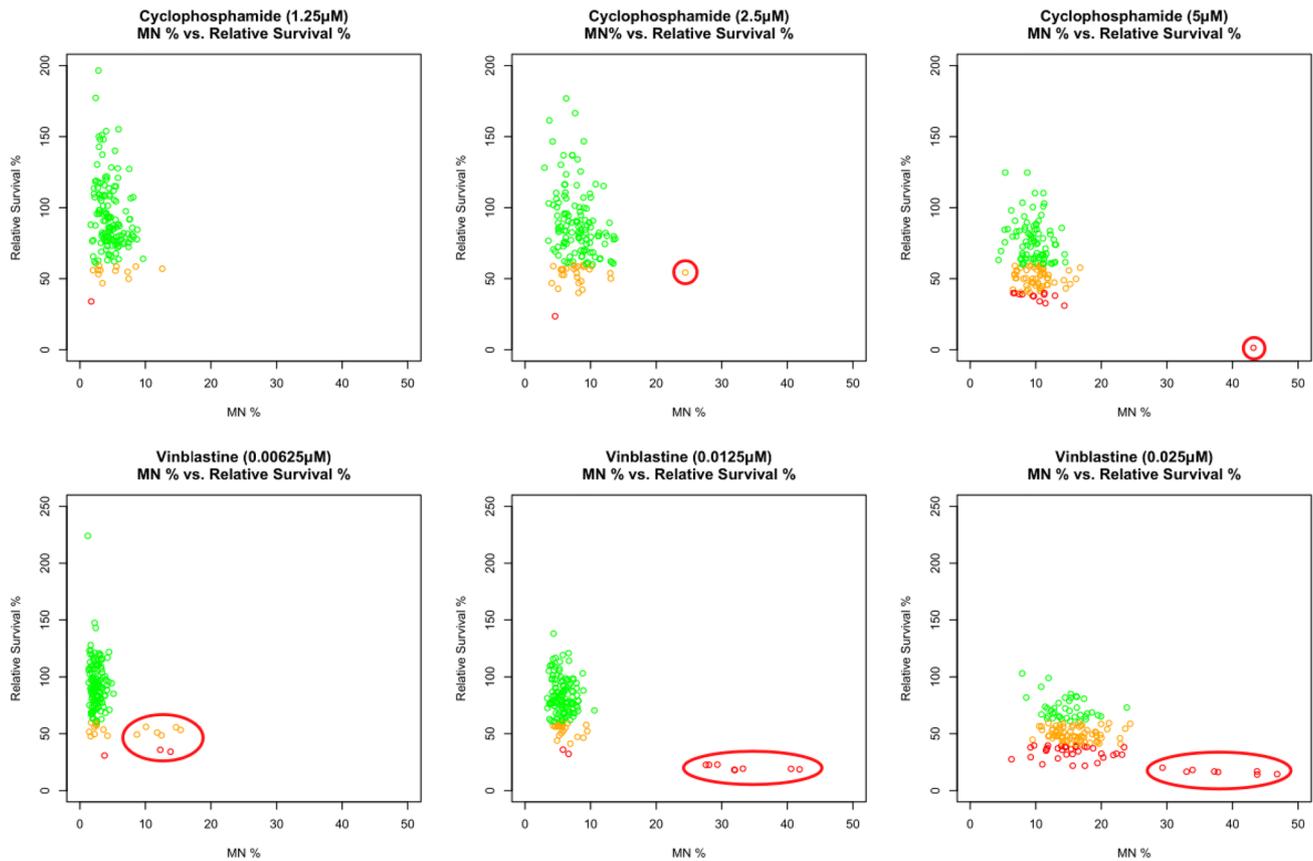


Fig. 2 Scatter plots of positive controls, vinblastine (–S9) and cyclophosphamide (+S9), versus relative survival rate. Each plot represents increasing concentrations of test agents (1.25, 2.5, and 5 μM for cyclophosphamide, and 0.00625, 0.0125, and 0.025 μM for vinblastine). Wells with relative survival rates below 40% are colored in red,

between 40 and 60% in orange, and above 60% in green. One outlier was identified for cyclophosphamide for each of 2.5 μM and 5 μM and eight outliers were identified for all three vinblastine concentrations (circled in red). Two outliers were not plotted for cyclophosphamide at 1.25 μM (color figure online)

and +S9 ranged from 1.65% to 2.96%, and from 1.22% to 2.31%, respectively.

Relative survival rate and cytotoxicity assessment

Cytotoxicity for each chemical, concentration, and S9 treatment, was evaluated using its relative survival data (Supplemental File 2). Relative survival rates below 40%, between 40 and 60%, and above 60% were classified as cytotoxic, adequate, and non-cytotoxic, respectively. Of the 11,570 unique data points, 571, 274, and 10,725 were classified as cytotoxic, adequate, and non-cytotoxic, respectively. Among all chemical treatments, 56%, 28.4%, 13%, and 2.6% had a relative survival rate < 100%, 100–120%, 121–150%, and > 150%, respectively. In the positive controls, 80%, 12.6%, 5.4% and 2% for cyclophosphamide +S9, and 85%, 11.8%, 3% and 0.2%, for vinblastine –S9, had relative survival rates of < 100%, 100–120%, 120–150%, and > 150%, respectively.

BMC analysis

BMC30 and BMC60 values were generated for each dataset (\pm S9) for MN and hypodiploid frequencies, respectively. Of the 614 datasets, 284 did not produce a BMC30 value for MN (i.e., had no concentration–response or had a BMC > highest dose) and 311 did not enable calculation of a BMC60 value for hypodiploidy. None of the BMC30 values for MN exceeded the highest concentrations included during BMC modeling, whereas five BMC60 values for hypodiploidy exceeded the highest concentrations used.

Micronucleus assessments

We performed MN assessment according to the decision-making scheme outlined in Fig. 1. Using our criteria, 180 data sets yielded positive results, either in the presence or absence of S9 (Table 2; further details in Supplemental Files 3 and 4). This amounts to 157 of the 292 chemicals

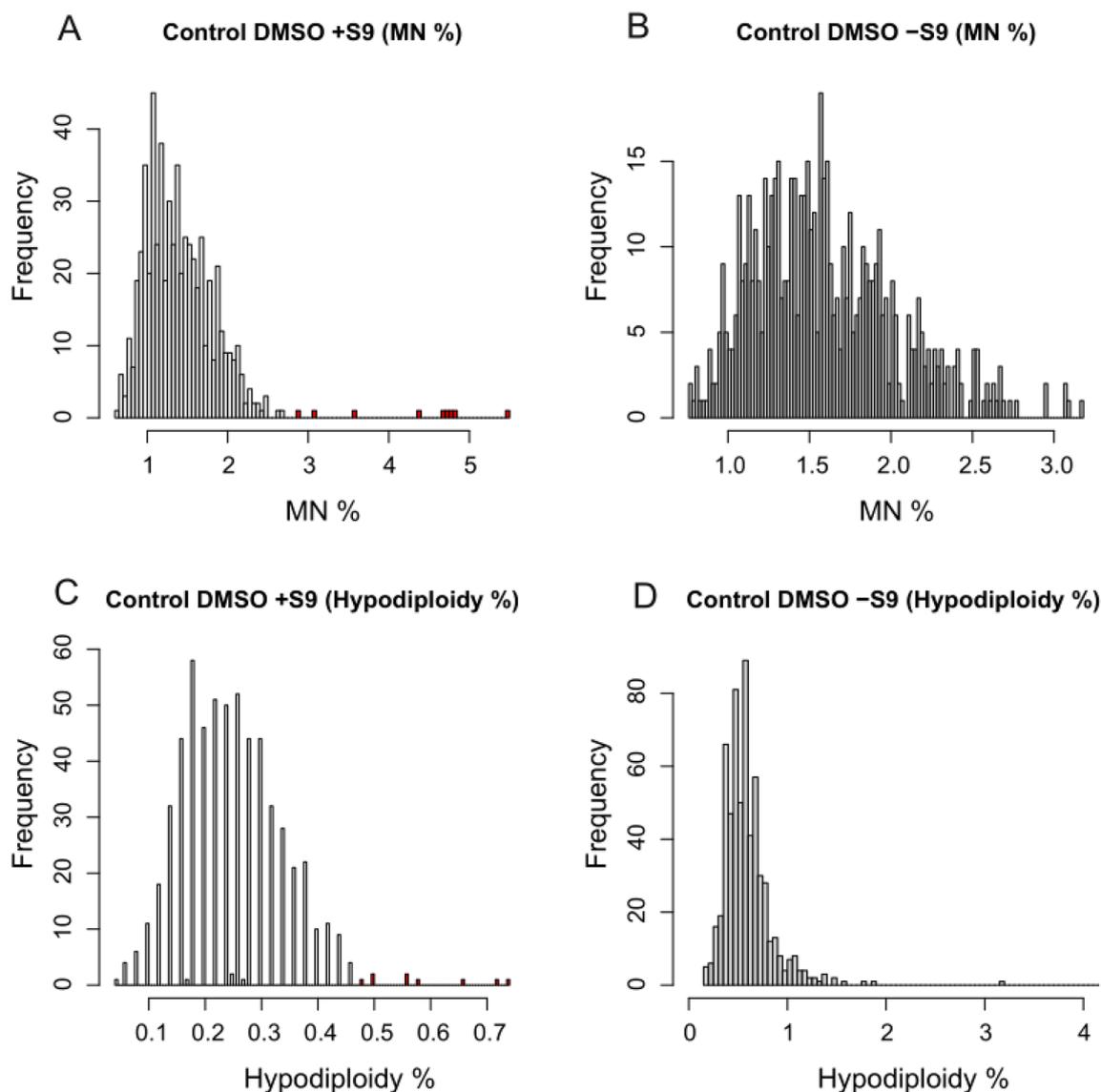


Fig. 3 Overall distributions for solvent control (DMSO) %MN and %hypodiploid, \pm S9. For the %MN and %hypodiploid+ \pm S9 distributions (A and C), nine outliers belonging to 20131009 Plate 2 are

colored in red. Four data points from solvent control DMSO – S9 (hypo %) (D) were also removed (hypodiploid ranges from 15 to 132%). No clear outliers were observed in (B) (color figure online)

with a positive MN response in at least one of two MN tests (\pm S9), and thus classified as clastogenic. In contrast, 25 chemicals were negative in both the presence and absence of S9. The remaining 110 chemicals were categorized as inconclusive; these chemicals either yielded inconclusive calls in both \pm S9 MN tests, or were inconclusive in one and negative in the other MN test. Note that each chemical is found only once in the ‘Total number of unique chemicals’ row (for a total of 292 chemicals), reflecting that each chemical was given a single positive, negative or inconclusive call. Of the 157 positive chemicals, 38 were MN positive only in the presence of S9, indicating that these chemicals require metabolic generation of a clastogenic metabolite.

We also applied an additional filter of at least one response greater than two-fold above the DMSO solvent control (i.e., concentration of zero μ L/mL) prior to modeling to derive a BMC, as in vitro chromosome damage assays are often considered overly sensitive (i.e., low specificity). With this filter, the numbers of chemicals classified as positive, negative and inconclusive were 115, 41 and 136, respectively.

To identify chemicals causing hypodiploidy (i.e., aneugens), we first applied an identical decision-making scheme. This analysis yielded 209 chemicals with hypodiploidy positive results in either the presence or absence of S9, and 27 and 56 chemicals with negative and inconclusive calls,

Table 1 Upper 95th percentiles of the distribution of solvent control (DMSO) %MN and %hypodiploid (%hypo) frequencies calculated based on the distributions of DMSO datasets per batch (date—year, month, day)

Batch date	%MN 95th percentile		%Hypodiploidy 95th percentile	
	–S9	+S9	–S9	+S9
20130827	1.86	1.22	0.84	0.30
20130829	1.80	1.75	0.64	0.38
20130903	2.31	1.81	0.56	0.32
20130905	1.92	2.00	0.61	0.42
20130906	2.31	2.29	0.81	0.37
20130909	2.47	2.24	0.66	0.39
20130911	2.96	2.10	1.46	0.36
20130916	2.65	2.06	0.76	0.37
20130918	2.47	2.24	0.82	0.44
20130923	2.28	1.99	0.80	0.41
20130925	2.21	2.31	0.77	0.38
20130930	2.36	2.12	1.36	0.49
20131007	2.19	2.12	1.03	0.44
20131009	2.46	2.00 ^a	0.75	0.32 ^a
20131015	1.65	1.61	0.93	0.54
20131017	1.84	1.36	0.74	0.32
20131024	2.28	2.00	0.70	0.41

^aData collected on “20131009 Plate 2” were removed from the calculations

Table 2 Hazard classifications based on %MN assessments following the decision tree (Fig. 1)

MN assessment	Positive	Negative	Inconclusive
# of Datasets			
–S9	115 (77) ^a	114 (152)	79 (79)
+S9	65 (56)	83 (92)	158 (158)
Total	180 (133)	197 (244)	237 (237)
Total # of unique chemicals	157 (115) ^P	25 (41) ^N	110 (136) ^I

^aNumbers in brackets are chemicals assessed with an additional filter of at least one response greater than two-fold above the solvent control

^P‘Positive’ chemicals had positive calls in either \pm S9 tests

^N‘Negative’ chemicals had negative calls in both \pm S9 tests

^I‘Inconclusive’ chemicals had either (a) inconclusive calls in both \pm S9 tests; (b) inconclusive calls in one of the \pm S9 tests and negative in the other

respectively (Table 3; with further details in Supplemental File 3). As a comparison, we also applied a more stringent approach to the identification of aneugens (Bryce et al. 2011). Using this approach, 15 chemicals were hypodiploidy positive and 277 were negative, either in the presence or absence of S9 (Table 4). Of the 15 chemicals identified as hypodiploid positive, there was literature to support

Table 3 Hypodiploidy (aneugenicity) hazard classification using the same decision tree as shown in Fig. 1, but replacing 95th percentile for %MN with the 95th percentile for %hypodiploid cells

Hypodiploidy assessment #1	Positive	Negative	Inconclusive
# of datasets			
–S9	193 (154) ^a	89 (128)	26 (26)
+S9	34 (31)	79 (82)	193 (193)
Total	227 (185)	168 (210)	219 (219)
Total # of unique chemicals	209(172)	27 (35)	56 (85)

^aNumbers in brackets are chemicals assessed with an additional filter of at least one response greater than two-fold above the solvent control

Table 4 Hypodiploidy (aneugenicity) hazard assessment using the classification method described by Bryce et al. (2011)

Hypodiploidy approach 2 ^a	Positive	Negative
# of datasets		
–S9	13	295
+S9	5	301
Total	18	596
Total # of unique chemicals	15	277

^aPositive if MN fold change > 3 and hypodiploidy fold-change > 10

that ten of them are aneugenic (Supplemental File 4). Two chemicals, Amiodarone hydrochloride (CAS# 19774-82-4) and Michlers ketone (CAS# 90-94-8) were positive only in the presence of S9, indicating the requirement of metabolic activation to generate an aneugenic metabolite. Nine of the 15 aneugenic positive chemicals according to Bryce et al. (2011) were also classified as clastogenic in our proposed decision-tree approach, regardless of the two-fold change filter.

Comparison of AEDs and PODs

BMC confidence interval plots for both %MN and %hypodiploidy for each chemical are rank ordered in Supplemental Files 5, 6 and 7. AEDs could be derived for 137 of the compounds exclusively classified as positive for clastogenicity using the decision-tree approach (i.e., without the optional two-fold filter), and 14 of the chemicals classified as aneugens using the Bryce et al. 2011 approach (Supplemental File 1) based on availability of the necessary kinetics data and models for IVIVE. For AED determination, the lower BMC \pm S9 was used. These in vitro AEDs are referred to as $POD_{\text{Clastogen}}$ and POD_{Aneugen} .

To compare traditional PODs and AEDs directly, genotoxicity and cancer study data were extracted from ToxValDB. The sources of the POD data were from the: (a) European Union COSMOS project for 111 PODs, (b)

European Chemicals Agency (ECHA) or ECHA International Uniform Chemical Information Database (IUCLID) for 45 PODs, and (c) Health Assessment Workplace Collaborative (HAWC) for 16 PODs. There were a total of 172 *in vivo* PODs for 33 chemicals classified as clastogenic using the decision-tree approach. Specifically, there were 109 PODs from cancer studies for 31 chemicals (referred to as $\text{POD}_{\text{Cancer}}$) and 63 PODs from *in vivo* genotoxicity studies for 12 chemicals (referred to as $\text{POD}_{\text{Genetox}}$). None of the chemicals positive for hypodiploidy had data in ToxValDB and therefore, comparisons could only be made between AEDs and traditional PODs for clastogens.

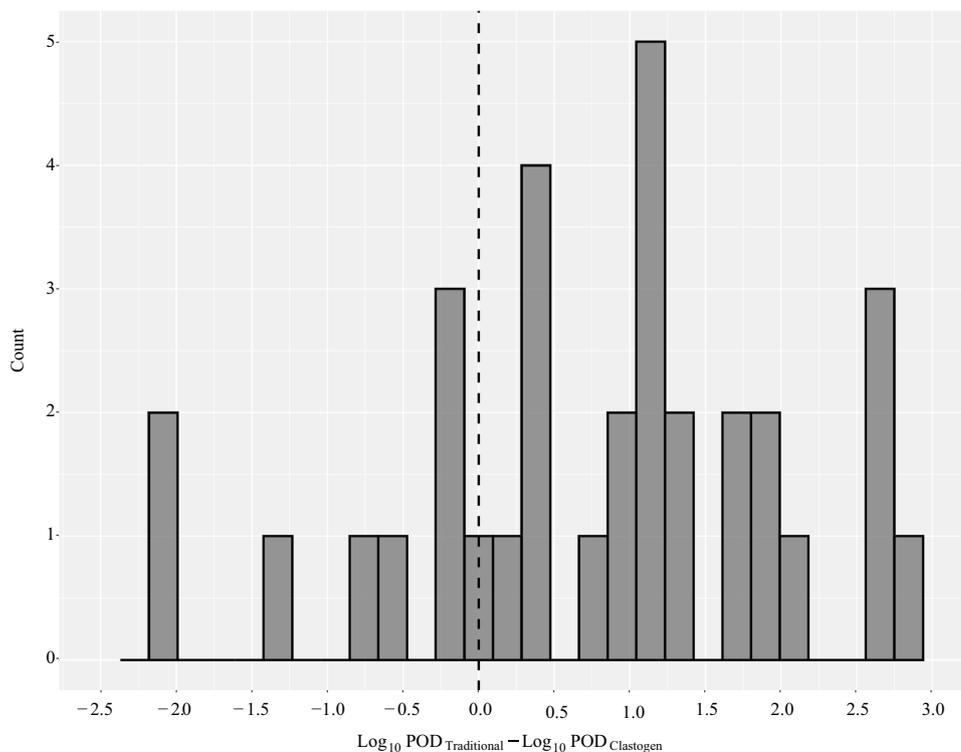
When using the lowest traditional POD from both genotoxicity and cancer studies, the AED was lower than the traditional POD for 24 out of 33 (72.7%) of possible comparisons (Fig. 4). The $\log_{10}\text{Traditional POD} - \log_{10}\text{AED}$ difference ranged from -2.10 to 2.82 (Supplementary File 8). The average difference was 0.77 (i.e., AED was 5.9-fold lower than traditional POD on arithmetic scale). There were only two compounds where the difference was below -2 (i.e., AED is 100-fold higher than traditional POD on an arithmetic scale). Folic acid and acrylamide both had a \log_{10} difference of -2.10 . The traditional POD for folic acid was based on *in vivo* micronucleus data with a highest no effect level of $0.0033 \text{ mg/kg bw/day}$, and the lowest POD for acrylamide was from a cancer study with a no-observed-adverse-effect level (NOAEL) of 0.1 mg/kg bw/day .

When the analysis was limited to the lowest POD from cancer studies, the AED was lower than the traditional POD for 26 out of 31 (83.9%) compounds with available cancer data. The difference ranged from -2.10 to 3.06 , with an average difference of 1.15 (i.e., AED is 14.2-fold lower on arithmetic scale). When the analysis was limited to the lowest POD from *in vivo* genotoxicity studies, the AED was lower than the traditional POD for 8 out of 12 (66.7%) of compounds with data. The difference ranged from -2.10 to 2.82 , with an average difference of 0.42 (i.e., AED is 2.6-fold lower on arithmetic scale). Thus, the AED is typically lower than the cancer POD by one order of magnitude and is on the same order of magnitude for *in vivo* genetic toxicity PODs.

Ranking of potential chemical risk based on bioactivity exposure ratios (BERs)

Available exposure estimates allowed for the derivation of BERs for 122 clastogens (Fig. 5) and 8 aneugens (Fig. 6; details in Supplemental File 1). Using the ExpoCast median exposure estimate, there was one clastogenic compound with a $\log_{10}\text{BER} < 2$ (dinoseb, $\log_{10}\text{BER} = 1.98$), two compounds with a $\log_{10}\text{BER}$ between 2 and 3 (azobenzene, $\log_{10}\text{BER} = 2.92$; 1,2-diphenylhydrazine, $\log_{10}\text{BER} = 2.96$), and the remaining 119 compounds had a $\log_{10}\text{BER} > 3$ (range 3.33–11.87). Using the ExpoCast 95th percentile exposure estimates, there were seven compounds with

Fig. 4 POD ratio distribution between lowest traditional POD (i.e., across *in vivo* genotoxicity and cancer studies) and AEDs determined herein using *in vitro* MN data. The dashed line indicates where the traditional POD and AED are equivalent. Chemicals to the left of the dashed line have a higher AED than the traditional POD, and chemicals to the right have a lower AED than the traditional POD (i.e., are more conservative)



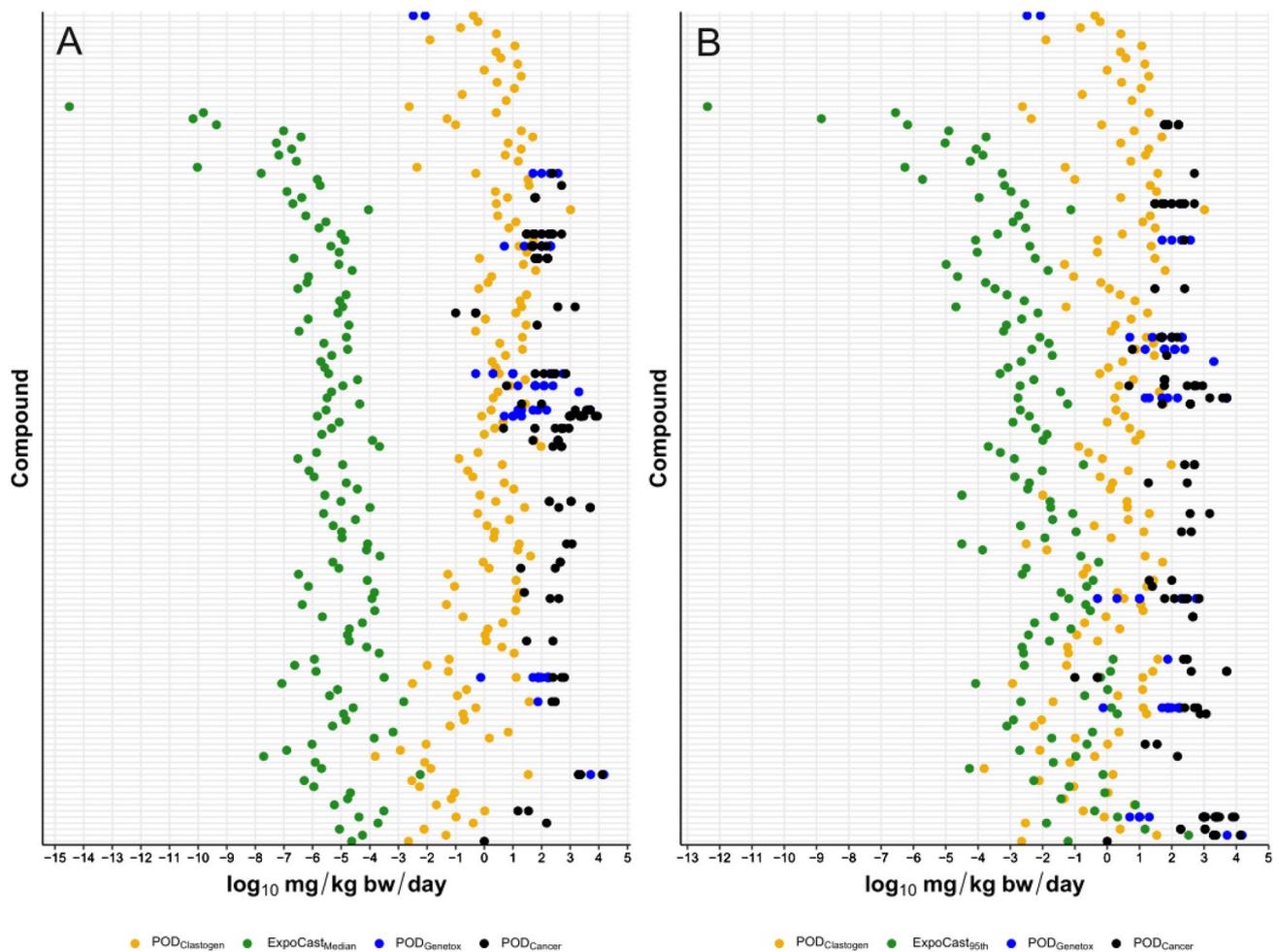


Fig. 5 Bioactivity Exposure Ratios (BERs) for compounds classified as clastogens based on the decision tree (yellow). **A** ExpoCast median exposure estimates (green), and **B** ExpoCast 95th percentile exposure estimates (green). Traditional genetic toxicology POD (blue) and cancer

POD (black) from ToxValDB are plotted for comparison. The full list of chemicals and their respective POD values are in Supplemental File 8 (color figure online)

negative \log_{10} BERs, where exposure estimates were above AEDs based on MN data. Specifically, dinoseb, 1,2-propylene glycol, caprolactam, 4-nitroaniline, FD&C yellow 5, 1,2-dimethyl-3-nitrobenzene, and di(2-ethylhexyl) adipate had \log_{10} BERs ranging from -1.44 to -0.04 . There were 43 chemicals where the \log_{10} BER was between 0 and 2 (range 0.09–1.99), 23 chemicals with \log_{10} BER between 2 and 3 (range 2.05–2.97), and 49 chemicals with \log_{10} BER > 3 (range 3.00–9.75).

When the ExpoCast median exposure estimates were used to evaluate chemicals classified as aneugenic, there was one compound with a negative \log_{10} BER. Specifically, 4-hexylresorcinol had a \log_{10} BER of -2.30 . This was driven by the extremely low BMC (i.e., 1×10^{-6} μM without S9) rather than by the C_{50} or exposure estimate. The remaining 7 compounds all had a \log_{10} BER > 3 (range 3.36–7.56). When the derivation of BERs for aneugenic compounds was based on

95th percentile ExpoCast exposure estimate, there were two compounds with negative \log_{10} BERs. These were 4-hexylresorcinol (-5.11) and benomyl (-0.19). There were no compounds with a \log_{10} BER between 0 and 2. There were two compounds with a \log_{10} BER between 2 and 3: 2-chloroacetophenone (2.32) and 1,5-naphthalenediamine (2.60). The remaining four compounds had a \log_{10} BER > 3 (triglycidyl isocyanurate, 3.11; 4,4-methylenebis(2-methylaniline), 3.16; carbendazim, 3.37; thiophanate-methyl, 5.55).

Discussion

Chemical genotoxicity assessment typically requires a combination of rigorous *in vitro* and *in vivo* studies, which can be costly and time-consuming. In this study, we paired *in vitro* and *in silico* methodologies to classify and assess

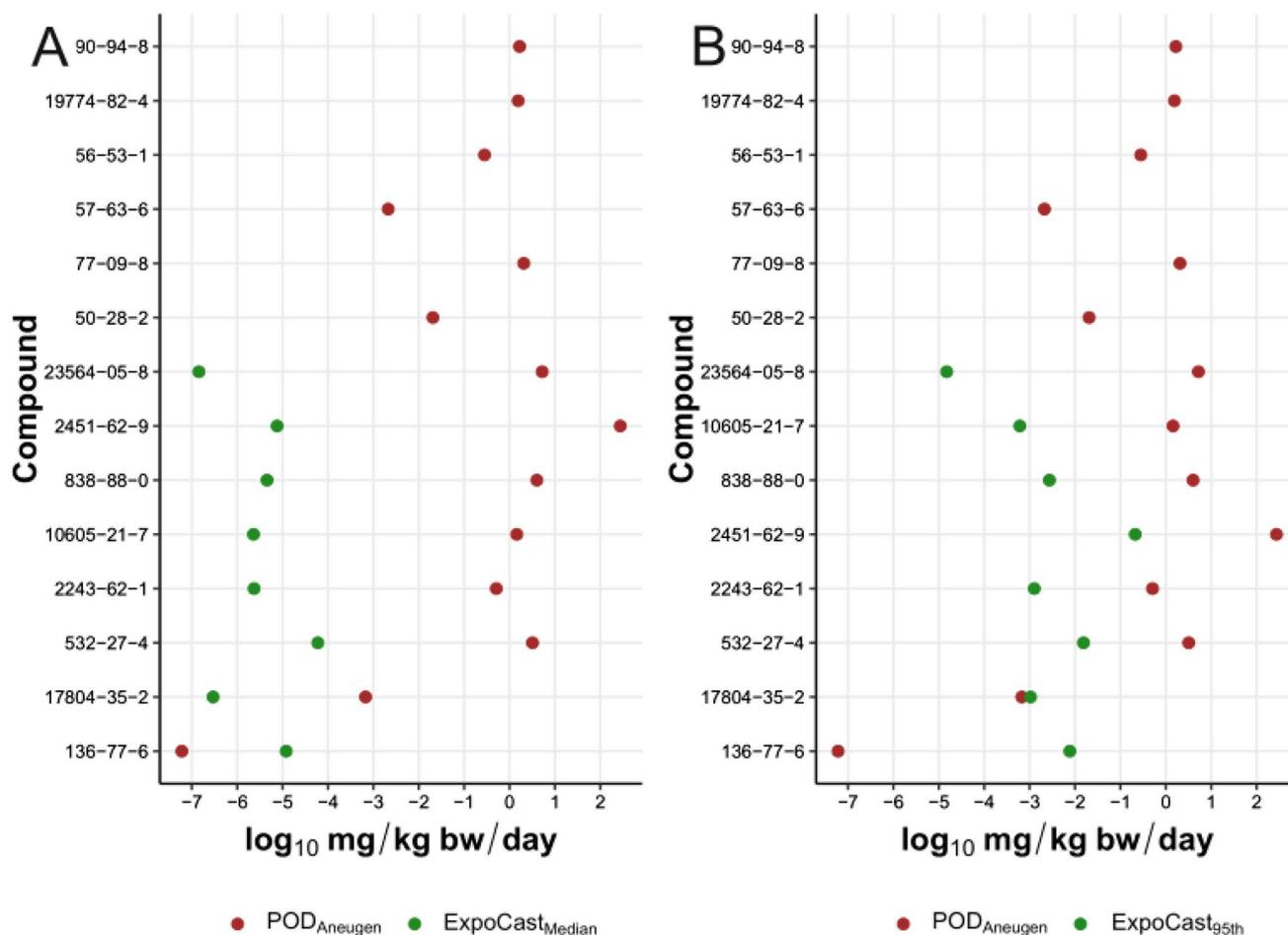


Fig. 6 Bioactivity Exposure Ratios (BERs) for compounds classified as aneugens based on the Bryce et al. (2011) approach (red). **A** ExpoCast median exposure estimates (green); **B** ExpoCast 95th percentile exposure estimates (green) (color figure online)

the genotoxicity of 292 chemicals using a modified experimental design for the high-throughput MNvit assay. Our approach consisted of four parts: (1) a decision tree to classify the chemicals based on MNvit data as positive, negative or inconclusive, (2) BMC analyses to identify a POD, (3) a toxicokinetic pipeline for IVIVE to produce AEDs for comparison to existing in vitro and traditional toxicity PODs; and (4) derivation of BERs through comparison of in vitro genotoxicity PODs with human exposure estimates for prioritization. We demonstrate how our approach can be used for hazard identification through identifying chemicals inducing significant concentration-responses, as well as for deriving PODs for each chemical. In addition, application of IVIVE enables conversion of the in vitro POD values to AEDs for comparison to in vivo responses and human exposure levels. Overall, the work describes an approach that can be used for weight-of-evidence evaluation as well as for potency comparisons and chemical prioritization.

In a typical analysis of MNvit dataset, a positive genotoxicity assessment requires a statistically significant increase

in MN frequency in at least one concentration over solvent-matched and historical controls (OECD 2016). In our datasets, there were 19 concentrations tested per chemical, enabling a robust quantitative analysis to increase the precision of hazard assessment and POD derivation. The highest tested concentration was 200 μ M, which is over the cytotoxicity threshold for most chemicals. With a few exceptions, this top concentration is also well above expected exposure levels for environmental chemicals. This design allowed us to bypass the typical pilot studies needed to establish test concentrations. A typical MN analysis might include 12–18 samples in total to investigate 3–5 concentrations, plus solvent controls. Instead, despite the lack of concentration-specific replicates, the experimental design herein includes more samples to more precisely define the concentration–response relationship. To ensure that the approach is robust, we designed a decision tree-based assessment approach that includes the use of cytotoxicity assessment, filtering to identify increases above batch solvent control, trend analyses, filtering to ensure at least one response is $> 2 \times$ the solvent control, and

concentration–response modeling to fit curves and identify BMCs for each chemical. The decision tree began with assessing the response of each concentration against the 95th percentile of batch-specific controls. We required at least two consecutive concentrations below or at the cytotoxicity threshold that were also over the 95th percentile of the controls prior to concentration–response analyses to ensure that the damage observed was not the result of cytotoxicity. In addition to requiring a concentration–response trend, filtering to ensure that the BMC was less than or equal to the highest concentration applied (i.e., below the cytotoxic threshold) further eliminated chemicals not meeting minimal positive responses within the concentration range tested. For most chemicals, the number of remaining concentrations applied in the BMC analysis following this filtering for cytotoxicity was more than ten, which provided robust BMC values.

The MNvit assay operates through detection of DNA fragments or whole chromosomes that manifest in the cytoplasm as micronucleus events after cell division. Thus, an increase in %MN can be the result of clastogenic or aneugenic (hypodiploidy) mechanisms. We attempted to assess both MN and hypodiploidy % with the proposed decision-tree approach. The decision tree classified 157 (53.7%) of the chemicals as MN positive, 25 (8.6%) negatives and 110 (37.7%) inconclusive. The optional two-fold filter reclassified 16 MN positive chemicals to negative and 26 to inconclusive. The same approach classified 209 (71.6%) chemicals as aneugenic positives, 27 (9.2%) negatives and 56 (19.2%) inconclusive. Conversely, the approach proposed by Bryce et al. (2011) classified 15 (5.1%) of these chemicals as positives and 277 (94.9%) as negatives. Since the number of positive calls using our initial approach was much higher than expected based on the number of known aneugens, we have more confidence in the Bryce et al. (2011) approach, indicating 5% of the chemicals are aneugens. Indeed, by manually examining the 15 classified aneugens, we found that seven and three chemicals were true or potential aneugens based on published data (Supplemental File 3) and there were no conflicting data suggesting otherwise for the others. Thus, assuming that these 15 chemicals induced MN through aneugenic mechanisms, leads us to conclude that the remaining 142 chemicals (of the 157 yielding positive calls) are potential clastogens.

Although the results from the proposed classification approaches can help prioritizing positive clastogens and aneugens for further investigation, the number of chemicals is substantial for comprehensive follow-up testing. Thus, we applied IVIVE to predict the AEDs of these clastogenic and aneugenic chemicals (i.e., $POD_{Clastogen}$ and $POD_{Aneugen}$, respectively), and compared them to exposure estimates to enable prioritization based on potential risk. To establish confidence in the AEDs, we compared them with known

genetic toxicity and cancer PODs. The closer a $POD_{Clastogen}$ or $POD_{Aneugen}$ for a compound was to its known traditional POD, the less conservative, or more aligned, the IVIVE derived POD was. For compounds where the AED was not protective of human health, follow-up investigations using higher tier physiologically based toxicokinetics models or other genetox assays may be required to understand the discrepancies with in vivo data.

Comparisons of $POD_{Clastogen}$ with traditional PODs, based on in vivo genotoxicity or cancer studies, revealed that the $POD_{Clastogen}$ was protective (i.e., lower) for most compounds. Specifically, the $POD_{Clastogen}$ was lower than the POD_{Cancer} for 83.9% of compounds, and on average was lower by approximately 14-fold. Similarly, the $POD_{Clastogen}$ was lower than the $POD_{Genetox}$ for 72.7% of compounds, on average by approximately sixfold. Previous work comparing PODs based on in vitro bioactivity data from the ToxCast battery, found that the bioactivity PODs were lower than traditional PODs from repeat-dose studies for 89% of chemicals by > 100-fold (Paul Friedman et al. 2020). Thus, MNvit is not as sensitive as the ToxCast test battery, but it is impressive this single assay provided protective PODs in a majority of cases.

Acrylamide was the only compound where the $POD_{Clastogen}$ was two orders of magnitude higher than the POD_{Cancer} . The genotoxicity of acrylamide has been extensively studied, and it is well-established that acrylamide is a clastogenic compound at relatively high doses (European Food Safety Authority 2015). In addition, the active metabolite of acrylamide, glycidamide, is a strong mutagen as well as a clastogen. The POD_{Cancer} values used in the comparison with the $POD_{Clastogen}$ were based on a two year cancer study in rats where acrylamide was administered through drinking water (Johnson et al. 1986). The lowest POD_{Cancer} was a no-observed-adverse-effect level (NOAEL) of 0.1 mg/kg for mesothelioma of the testes tunica albuginea. The other two POD_{Cancer} values were based on NOAELs of 0.5 mg/kg for mammary gland tumors or thyroid gland tumors. These tissues are hormone-sensitive, and it has been postulated that there are alternative modes of action than genotoxicity driving the formation of these tumors, and that the modes of action for scrotal mesothelioma tumor response may not be relevant to humans (Haber et al. 2009; Hogervorst et al. 2010). In addition, it has been shown that carcinogenicity occurs at lower doses than genotoxicity in both rat and mouse models, supporting alternative modes of action (Chepelev et al. 2017; Hobbs et al. 2016). Indeed, gene expression analysis of rat testis following acrylamide exposure found that the main transcriptional alterations were associated with cytoskeletal proteins and calcium signaling, with no strong evidence for the induction of genotoxicity (Recio et al. 2017). Thus, a more reliable evaluation of the $POD_{Clastogen}$

would be to compare it with PODs derived from in vivo MN studies. Examining rodent studies measuring in vivo MN induced by acrylamide (Rothfuss et al. 2010; Yener 2013) showed that the lowest $POD_{Genetox}$ is 2.0 mg/kg bw/day based on a lowest-observed-adverse-effect level (LOAEL) in polychromatic erythrocytes of male rats (Yener 2013) and a NOAEL in normochromatic erythrocytes of male mice (Zeiger et al. 2009). Moreover, Hobbs et al. (2016) found that B6C3F1 mice exposed to 6 mg/kg bw/day of acrylamide had increased frequencies of micronucleated reticulocytes and normochromatic erythrocytes. However, the same response was not identified in F344 rats at doses as high as 12 mg/kg bw/day, which is lower than the $POD_{Clastogen}$ of 12.6 mg/kg bw/day. Hence, the $POD_{Clastogen}$ more closely aligns with the $POD_{Genetox}$, with the $POD_{Clastogen}$ being 2- to 6-fold higher than the $POD_{Genetox}$ for positive results.

Folic acid was the only compound with a $POD_{Clastogen}$ more than 100-fold higher than the $POD_{Genetox}$. However, the $POD_{Genetox}$ was based on folate deficiency and this is not addressed by our study. Manufactured folic acid is a dietary supplement that is used in food fortification and is converted in the body to folate, a water-soluble naturally occurring B vitamin. It is recommended that pregnant women are supplemented with an extra 0.4 mg of synthetic folic acid per day to reduce the chance of neural tube defects in their offspring, and the tolerable daily intake of folic acid is 1 mg/day (Institute of Medicine (US) Standing Committee on the Scientific Evaluation of Dietary Reference Intakes and its Panel on Folate, Other B Vitamins, and Choline 1998). There is strong evidence that folic acid deficiency is associated with increased cancer risk and elevated rates of mutations and micronuclei (LeBlanc et al. 2018; National Toxicology Program 2015), and that the response may be tissue-dependent (Diaz et al. 2021). Indeed, the folic acid $POD_{Genetox}$ of 3×10^{-3} mg/kg bw/day (i.e., intake of 0.195 mg/day in a 60 kg adult) used in the comparison with $POD_{Clastogen}$ was based on the baseline folic acid level in postmenopausal women, and folic acid levels that were one third below this level were associated with increased MN in lymphocytes and exfoliated buccal cells (i.e., representing levels of folic acid deficiency) (Titenko-Holland et al. 1998). Thus, the $POD_{Clastogen}$ should not be viewed as being non-protective of $POD_{Genetox}$, but instead the $POD_{Clastogen}$ raises questions about the relationship between excessively high levels of folic acid and micronuclei induction, which require further study. There is evidence that relatively high levels of folic acid may be weakly associated with colorectal cancer under select circumstances and through an unknown mechanism (Kim 2006, 2018). A recent study demonstrated that mice supplemented with a folic acid dose of 8 mg/day (i.e., approximately a diet of 1.6 mg/day in a human) had elevated mutation frequencies in colon but not bone marrow

(Diaz et al. 2021). The $POD_{Clastogen}$ derived in this study was 0.4 mg/kg bw/day (24 mg/day in 60 kg adult; 30 mg/day in 75 kg adult), and far exceeds the established tolerable daily intake for folic acid. Consequently, any potential risk associated with clastogenicity can be considered negligible relative to the benefits of supplementary folic acid intake, and the guidance for folic acid supplementation should be followed (Koren 2011).

We explored using MNvit assay to identify aneugenic compounds. Using the decision-tree approach, an unexpectedly high number of aneugens totaling 209 chemicals were identified. In contrast, the Bryce et al. (2011) approach classified 15 chemicals as aneugens, seven of which are known aneugens and three are likely aneugens. Specifically, three benzimidazole compounds, benomyl, carbendazim, and methyl thiophanate (parent compound of carbendazim), are well-established aneugens (Barale et al. 1993; Van Hummel et al. 1995), and were classified accordingly in this study. The naturally occurring steroid hormone 17β -estradiol and synthetic analogs, 17α -ethinylestradiol and diethylstilbestrol, are also established aneugens (Brown et al. 2008; Hernández et al. 2013; Parry et al. 2002; Rosefort et al. 2004) and our MNvit results support these findings. The other chemicals with elevated hypodiploidy levels have been previously shown to be aneugenic, such as Michler's ketone (Lafi et al. 1986) and phenolphthalein (Armstrong et al. 2000), or test positive for other chromosomal aberrations, such as triglycidyl isocyanurate (Willcocks et al. 1998). These comparisons support the utility of the Bryce et al. (2011) approach in identifying aneugenic compounds, and thus, chemicals classified as aneugenic using this approach should be considered as priorities for further scoping and risk evaluation.

Based on the BER analysis, 4-hexylresorcinol, an anti-septic and food additive, has a high potential for aneugenic concern due to its very low $POD_{Aneugen}$ and BER. The $POD_{Aneugen}$ for 4-hexylresorcinol was two and five orders of magnitude lower than the ExpoCast median and 95th percentile estimates, respectively. The low $POD_{Aneugen}$ of 6×10^{-8} mg/kg bw/day results from the low BMC (1×10^{-6} μ M) as opposed to the HHTK-modeled C_{ss} used for POD derivation. The C_{ss} is not expected to be high as in vivo studies have shown that 4-hexylresorcinol is rapidly excreted in urine and feces (Walker 2005). Interestingly, previous investigations of 4-hexylresorcinol using the MN assay found that it was negative in vitro and positive in vivo (Soeteman-Hernández et al. 2015, 2016). In our study, 4-hexylresorcinol was considered negative for clastogenicity classification with and without S9. Likewise, resorcinol is the precursor for 4-hexylresorcinol, which has been shown to be negative for MN in vivo (Natarajan and Obe 1986). Resorcinol is an ECVAM group 3 chemical, meaning it yields positive in vitro results at high concentrations, likely as a result of high levels of cytotoxicity (Kirkland et al. 2008,

2016). In contrast, there is evidence that 5-pentylresorcinol, another structural analog of 4-hexylresorcinol, may disrupt microtubule and spindle formation through a decrease in DNA synthesis and indirect induction of abnormal anaphase configurations (National Toxicology Program 1988), indicating that it is a potential aneugen. In the EFSA re-evaluation of 4-hexylresorcinol as a food additive, the panel concluded that there was no safety concern of genotoxicity or carcinogenicity for 4-hexylresorcinol based on its current use levels (European Food Safety Authority 2014). However, the panel noted that there were no reproductive and developmental toxicity studies, and that a one-generation study would be required before permitting an increase in usage level. Based on the findings of this study, an *in vivo* assessment of aneugenicity should also be considered for any assessments of 4-hexylresorcinol.

There are several reasons why the PODs derived using our computational approach may differ from traditional PODs. First, the PODs were derived using a BMC of one standard deviation (30% for micronucleus and 60% for hypodiploidy), whereas the traditional PODs are often based on NOAELs confined by study design. Second, the IVIVE model was limited to the toxicokinetic parameters of the parent compound. This is a suitable approach when S9 activation is not required, as was the case for the majority of clastogens (118/157) and aneugens (12/15). However, for compounds where the metabolite is active and drives genotoxicity, there are several assumptions that need to be made for IVIVE (Lutz et al. 2010). Specifically, we made conservative assumptions that the formation clearance rate of the active metabolite was approximately the same as the elimination clearance rate, the active metabolite is the main metabolite from biotransformation of the parent and approaches the same concentration in the plasma as the parent, and the metabolite is available to the systemic circulation. In general, this approach yielded protective PODs from the MNvit data; however, more focused research on predicting the *in vivo* disposition of metabolites is required to improve the approach. Third, the PODs were derived using the nominal (assumed) concentration in the medium as opposed to the true intracellular concentration, which can vary by orders of magnitude for certain chemical classes (Armitage et al. 2014). Overall, further refinements to the computational approach will serve to enhance the utility of this novel quantitative *in vitro* genetic toxicology framework for risk assessment applications.

Another important consideration is that there are several modes of action underlying genotoxicity and carcinogenicity, and one assay alone is often not sufficient for hazard characterization of genotoxicants. The development of a validated panel of complementary higher-throughput and higher-content *in vitro* assays could greatly enhance the

mechanistic information and concomitantly improve genotoxicity assessment. Several higher-throughput assays have recently been developed for assessing genotoxicity and most of these assays are functionally related to well-established assays with existing test guidelines. Another complementary assay for detecting gene mutations is the *in vitro* MutaMouse assay that detects mutations in the *lacZ* reporter gene in FE1 cells (Maertens et al. 2017; White et al. 2003). Furthermore, there are novel assays that detect biomarkers of DNA damage response using flow cytometry (MultiFlow) (Bemis and Heflich 2019; Bryce et al. 2018) or using reporter cell lines (ToxTracker) (Hendriks et al. 2012). Higher content assays, such as the TGx-DDI biomarker (Li et al. 2015, 2017) provide a robust classification system for identifying DNA-damaging agents, and aid in better understanding a chemical's mode of action. In addition, the choice of cell line and the efficacy of the S9 metabolic activation, and other factors, may influence the PODs derived from the *in vitro* assays. Thus, application of a combination of these *in vitro* assays in a variety of cell lines, with an integrated computational approach, could serve to identify compounds with the highest potential for genotoxic concern as priorities for further work.

In summary, we propose a framework for the use of a modified test paradigm of MNvit datasets for genotoxic hazard identification, potency analysis and compound prioritization. This study is aligned with recent efforts to advance the use of genotoxicity data, which has traditionally been used exclusively for hazard identification, toward quantitative evaluation and application in risk assessment to support regulatory decision making (Heflich et al. 2020; White et al. 2020). We applied IVIVE to extrapolate the BMC values of the positive clastogens and aneugens to derive AEDs, enabling comparison with traditional genotoxicity and cancer PODs. The majority of the determined AEDs were protective relative to $POD_{Genetox}$ and POD_{Cancer} . Overall, our proposed approach enables prioritization of compounds for further evaluation on the basis of BERs and the use of our decision-tree pipeline supports weight-of-evidence based decision making in the context of effective risk assessment.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00204-022-03286-2>.

Acknowledgements We thank Dr. Russell Thomas of the US EPA for initiating the collaboration for this project. We are grateful for the help provided for the selection of test chemicals from Drs. Ann Richard and Chihae Yang. CLY thanks the Canada Research Chairs Program for partial funding on this project. Finally, we also like to acknowledge the support of Health Canada's Chemicals Management Plan for this project.

Declarations

Conflict of interest The authors report no conflict of interest and have no competing interests to declare.

Disclaimer This article has been reviewed by the U.S. Environmental Protection Agency, Office of Research and Development and approved for publication. Approval does not signify that the contents reflect the views of the Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Armitage JM, Wania F, Arnot JA (2014) Application of mass balance models and the chemical activity concept to facilitate the use of in vitro toxicity data for risk assessment. *Environ Sci Technol* 48:9770–9779. <https://doi.org/10.1021/es501955g>
- Armstrong MJ, Gara JP, Gealy R, Greenwood SK, Hilliard CA, Laws GM, Galloway SM (2000) Induction of chromosome aberrations in vitro by phenolphthalein: mechanistic studies. *Mutat Res* 457:15–30. [https://doi.org/10.1016/s0027-5107\(00\)00119-6](https://doi.org/10.1016/s0027-5107(00)00119-6)
- Avlasevich SL, Bryce SM, Cairns SE, Dertinger SD (2006) In vitro micronucleus scoring by flow cytometry: differential staining of micronuclei versus apoptotic and necrotic chromatin enhances assay reliability. *Environ Mol Mutagen* 47:56–66. <https://doi.org/10.1002/em.20170>
- Barale R, Scapoli C, Meli C, Casini D, Minunni M, Marrazzini A, Loprieno N, Barrai I (1993) Cytogenetic effects of benzimidazoles in mouse bone marrow. *Mutat Res* 300:15–28. [https://doi.org/10.1016/0165-1218\(93\)90135-z](https://doi.org/10.1016/0165-1218(93)90135-z)
- Barter ZE, Bayliss MK, Beaune PH, Boobis AR, Carlile DJ, Edwards RJ, Houston JB, Lake BG, Lipscomb JC, Pelkonen OR, Tucker GT, Rostami-Hodjegan A (2007) Scaling factors for the extrapolation of in vivo metabolic drug clearance from in vitro data: reaching a consensus on values of human microsomal protein and hepatocellularity per gram of liver. *Curr Drug Metab* 8:33–45. <https://doi.org/10.2174/138920007779315053>
- Beal MA, Gagne M, Kulkarni SA, Patlewicz G, Thomas RS, Barton-Maclaren TS (2022) Implementing in vitro bioactivity data to modernize priority setting of chemical inventories. *Altex* 39:123–139. <https://doi.org/10.14573/altex.2106171>
- Bemis JC, Heflich RH (2019) In vitro mammalian cell mutation assays based on the Pig-a gene: a report of the 7th International Workshop on Genotoxicity Testing (IWGT) Workgroup. *Mutat Res* 847:403028
- Benigni R, Bossa C, Jeliaskova N, Netzeva T, Worth A (2008) The Benigni / Bossa rulebase for mutagenicity and carcinogenicity - a module of toxtree. *EUR 23241 EN*. Luxembourg (Luxembourg): OPOCE; 2008. JRC43157
- Benz RD, Irausquin H (1991) Priority-based assessment of food additives database of the U.S. food and drug administration center for food safety and applied nutrition. *Environ Health Perspect* 96:85–89
- Brendt J, Lackmann C, Heger S, Velki M, Crawford SE, Xiao H, Thalmann B, Schiwy A, Hollert H (2021) Using a high-throughput method in the micronucleus assay to compare animal-free with rat-derived S9. *Sci Total Environ*. <https://doi.org/10.1016/j.scitotenv.2020.142269>
- Brown KH, Schultz IR, Cloud JG, Nagler JJ (2008) Aneuploid sperm formation in rainbow trout exposed to the environmental estrogen 17{alpha}-ethynylestradiol. *Proc Natl Acad Sci USA* 105:19786–19791. <https://doi.org/10.1073/pnas.0808333105>
- Bryce SM, Avlasevich SL, Bemis JC, Dertinger SD (2011) Miniaturized flow cytometry-based CHO-K1 micronucleus assay discriminates aneugenic and clastogenic modes of action. *Environ Mol Mutagen* 52:280–286. <https://doi.org/10.1002/em.20618>
- Bryce SM, Bernacki DT, Smith-Roe SL, Witt KL, Bemis JC, Dertinger SD (2018) Investigating the generalizability of the MultiFlow © DNA damage assay and several companion machine learning models with a set of 103 diverse test chemicals. *Toxicol Sci* 162:146–166. <https://doi.org/10.1093/toxsci/kfx235>
- Chepelev NL, Gagné R, Maynor T, Kuo B, Hobbs CA, Recio L, Yauk CL (2017) Transcriptional profiling of male F344 rats suggests the involvement of calcium signaling in the mode of action of acrylamide-induced thyroid cancer. *Food Chem Toxicol* 107:186–200. <https://doi.org/10.1016/j.fct.2017.06.019>
- Cohen Hubal EA, Richard A, Aylward L, Edwards S, Gallagher J, Goldsmith M, Isukapalli S, Tornero-Velez R, Weber E, Kavlock R (2010) Advancing exposure characterization for chemical evaluation and risk assessment. *J Toxicol Environ Health B Crit Rev* 13:299–313. <https://doi.org/10.1080/10937404.2010.483947>
- Decordier I, Kirsch-Volders M (2006) The in vitro micronucleus test: from past to future. *Mutat Res* 607:2–4. <https://doi.org/10.1016/j.mrgentox.2006.04.008>
- Diaz GS, LeBlanc DP, Gagné R, Behan NA, Wong A, Marchetti F, MacFarlane AJ (2021) Folate intake alters mutation frequency and profiles in a tissue- and dose-specific manner in mutamouse male mice. *J Nutr* 151:800–809. <https://doi.org/10.1093/jn/nxaa402>
- Dix DJ, Houck KA, Martin MT, Richard AM, Setzer RW, Kavlock RJ (2007) The ToxCast program for prioritizing toxicity testing of environmental chemicals. *Toxicol Sci* 95:5–12. <https://doi.org/10.1093/toxsci/kfl103>
- EFSA (2009) Guidance of the Scientific Committee on Use of the benchmark dose approach in risk assessment [1] | European Food Safety Authority. <http://www.efsa.europa.eu/en/efsajournal/pub/1150>
- EFSA (2017) Update: use of the benchmark dose approach in risk assessment. *EFSA J* 15:4658. <https://doi.org/10.2903/j.efsa.2017.4658>
- European Food Safety Authority (2014) Re-evaluation of 4-hexylresorcinol (E 586) as a food additive. *EFSA J* 12:3643. <https://doi.org/10.2903/j.efsa.2014.3643>
- European Food Safety Authority (2015) Scientific opinion on acrylamide in food. *EFSA J* 3:4104 <https://doi.org/10.2903/j.efsa.2015.4104>
- Fenech M (2000) The in vitro micronucleus technique. *Mutat Res* 455:81–95. [https://doi.org/10.1016/s0027-5107\(00\)00065-8](https://doi.org/10.1016/s0027-5107(00)00065-8)
- Gannon AM, Moreau M, Farmahin R, Thomas RS, Barton-Maclaren TS, Nong A, Curran I, Yauk CL (2019) Hexabromocyclododecane (HBCD): a case study applying tiered testing for human health risk assessment. *Food Chem Toxicol*. <https://doi.org/10.1016/j.fct.2019.110581>
- Haber LT, Maier A, Kroner OL, Kohrman MJ (2009) Evaluation of human relevance and mode of action for tunica vaginalis

- mesotheliomas resulting from oral exposure to acrylamide. *Regul Toxicol Pharmacol* 53:134–149. <https://doi.org/10.1016/j.yrtph.2008.12.008>
- Heflich RH, Johnson GE, Zeller A, Marchetti F, Douglas GR, Witt KL, Gollapudi BB, White PA (2020) Mutation as a toxicological endpoint for regulatory decision-making. *Environ Mol Mutagen* 61:34–41. <https://doi.org/10.1002/em.22338>
- Health Canada (2021) Science approach document—bioactivity exposure ratio: application in priority setting and risk assessment
- Hendriks G, Atallah M, Morolli B, Calléja F, Ras-Verloop N, Huijskens I, Raamsman M, van de Water B, Vrieling H (2012) The ToxTracker assay: novel GFP reporter systems that provide mechanistic insight into the genotoxic properties of chemicals. *Toxicol Sci* 125:285–298. <https://doi.org/10.1093/toxsci/kfr281>
- Hernández LG, van Benthem J, Johnson GE (2013) A mode-of-action approach for the identification of genotoxic carcinogens. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0064532>
- Hobbs CA, Davis J, Shepard K, Chepelev N, Friedman M, Marroni D, Recio L (2016) Differential genotoxicity of acrylamide in the micronucleus and Pig-a gene mutation assays in F344 rats and B6C3F1 mice. *Mutagenesis* 31:617–626. <https://doi.org/10.1093/mutage/gew028>
- Hogervorst JGF, Baars B, Schouten LJ, Konings EJM, Goldbohm RA, van den Brandt PA (2010) The carcinogenicity of dietary acrylamide intake: a comparative discussion of epidemiological and experimental animal research. *Crit Rev Toxicol* 40:485–512. <https://doi.org/10.3109/10408440903524254>
- Honda GS, Pearce RG, Pham LL, Setzer RW, Wetmore BA, Sipes NS, Gilbert J, Franz B, Thomas RS, Wambaugh JF (2019) Using the concordance of in vitro and in vivo data to evaluate extrapolation assumptions. *PLoS One* 14:e0217564. <https://doi.org/10.1371/journal.pone.0217564>
- Hsieh J, Smith-Roe SL, Huang R, Sedykh A, Shockley KR, Auerbach SS, Merrick BA, Xia M, Tice RR, Witt KL (2019) Identifying compounds with genotoxicity potential using Tox21 high-throughput screening assays. *Chem Res Toxicol* 32:1384–1401. <https://doi.org/10.1021/acs.chemrestox.9b00053>
- Institute of Medicine (US) Standing Committee on the Scientific Evaluation of Dietary Reference Intakes and its Panel on Folate, Other B Vitamins, and Choline (1998) Dietary reference intakes for thiamin, riboflavin, niacin, vitamin B6, folate, vitamin B12, pantothenic acid, biotin, and choline. National Academies Press (US), Washington (DC)
- Johnson KA, Gorzinski SJ, Bodner KM, Campbell RA, Wolf CH, Friedman MA, Mast RW (1986) Chronic toxicity and oncogenicity study on acrylamide incorporated in the drinking water of Fischer 344 rats. *Toxicol Appl Pharmacol* 85:154–168. [https://doi.org/10.1016/0041-008x\(86\)90109-2](https://doi.org/10.1016/0041-008x(86)90109-2)
- Kavlock R, Chandler K, Houck K, Hunter S, Judson R, Kleinstreuer N, Knudsen T, Martin M, Padilla S, Reif D, Richard A, Rotroff D, Sipes N, Dix D (2012) Update on EPA's ToxCast program: providing high throughput decision support tools for chemical risk management. *Chem Res Toxicol* 25:1287–1302. <https://doi.org/10.1021/tx3000939>
- Kim Y (2006) Folate: a magic bullet or a double edged sword for colorectal cancer prevention? *Gut* 55:1387–1389. <https://doi.org/10.1136/gut.2006.095463>
- Kim Y (2018) Folate and cancer: a tale of Dr. Jekyll and Mr. Hyde? *Am J Clin Nutr* 107:139–142. <https://doi.org/10.1093/ajcn/nqx076>
- Kirkland D, Kasper P, Müller L, Corvi R, Speit G (2008) Recommended lists of genotoxic and non-genotoxic chemicals for assessment of the performance of new or improved genotoxicity tests: a follow-up to an ECVAM workshop. *Mutat Res* 653:99–108. <https://doi.org/10.1016/j.mrgentox.2008.03.008>
- Kirkland D, Kasper P, Martus H, Müller L, van Benthem J, Madia F, Corvi R (2016) Updated recommended lists of genotoxic and non-genotoxic chemicals for assessment of the performance of new or improved genotoxicity tests. *Mutat Res Genet Toxicol Environ Mutagen* 795:7–30. <https://doi.org/10.1016/j.mrgentox.2015.10.006>
- Kirsch-Volders M, Plas G, Elhajouji A, Lukamowicz M, Gonzalez L, Vande Loock K, Decordier I (2011) The in vitro MN assay in 2011: origin and fate, biological significance, protocols, high throughput methodologies and toxicological relevance. *Arch Toxicol* 85:873–899. <https://doi.org/10.1007/s00204-011-0691-4>
- Kligerman AD, Young RR, Stankowski LF, Pant K, Lawlor T, Aardema MJ, Houck KA (2015) An evaluation of 25 selected ToxCast chemicals in medium-throughput assays to detect genotoxicity. *Environ Mol Mutagen* 56:468–476. <https://doi.org/10.1002/em.21934>
- Knight AW, Little S, Houck K, Dix D, Judson R, Richard A, McCarrroll N, Akerman G, Yang C, Birrell L, Walmsley RM (2009) Evaluation of high-throughput genotoxicity assays used in profiling the US EPA ToxCast chemicals. *Regul Toxicol Pharmacol* 55:188–199. <https://doi.org/10.1016/j.yrtph.2009.07.004>
- Koren G (2011) Folic acid and colorectal cancer: unwarranted fears. *Can Fam Physician* 57:889–890
- Krewski D, Andersen ME, Tyshenko MG, Krishnan K, Hartung T, Boekelheide K, Wambaugh JF, Jones D, Whelan M, Thomas R, Yauk C, Barton-Maclaren T, Cote I (2020) Toxicity testing in the 21st century: progress in the past decade and future perspectives. *Arch Toxicol* 94:1–58. <https://doi.org/10.1007/s00204-019-02613-4>
- Lafi A, Parry JM, Parry EM (1986) The effect of Michler's ketone on cell division, chromosome number and structure in cultured Chinese hamster cells. *Mutagenesis* 1:17–20. <https://doi.org/10.1093/mutage/1.1.17>
- LeBlanc DP, Behan NA, O'Brien JM, Marchetti F, MacFarlane AJ (2018) Folate deficiency increases chromosomal damage and mutations in hematopoietic cells in the transgenic mutamouse model. *Environ Mol Mutagen* 59:366–374. <https://doi.org/10.1002/em.22190>
- Li H, Hyde DR, Chen R, Heard P, Yauk CL, Aubrecht J, Fornace AJ (2015) Development of a toxicogenomics signature for genotoxicity using a dose-optimization and informatics strategy in human cells. *Environ Mol Mutagen* 56:505–519. <https://doi.org/10.1002/em.21941>
- Li H, Chen R, Hyde DR, Williams A, Frötschl R, Ellinger-Ziegelbauer H, O'Lone R, Yauk CL, Aubrecht J, Fornace AJ (2017) Development and validation of a high-throughput transcriptomic biomarker to address 21st century genetic toxicology needs. *Proc Natl Acad Sci U S A* 114:E10881–E10889. <https://doi.org/10.1073/pnas.1714109114>
- Liu Z, Huang R, Roberts R, Tong W (2019) Toxicogenomics: a 2020 vision. *Trends Pharmacol Sci* 40:92–103. <https://doi.org/10.1016/j.tips.2018.12.001>
- Lutz JD, Fujioka Y, Isoherranen N (2010) Rationalization and prediction of in vivo metabolite exposures: the role of metabolite kinetics, clearance predictions and in vitro parameters. *Expert Opin Drug Metab Toxicol* 6:1095–1109. <https://doi.org/10.1517/17425255.2010.497487>
- MacGregor JT, Frötschl R, White PA, Crump KS, Eastmond DA, Fukushima S, Guérard M, Hayashi M, Soeteman-Hernández LG, Johnson GE, Kasamatsu T, Levy DD, Morita T, Müller L, Schoeny R, Schuler MJ, Thybaud V (2015) IWGT report on quantitative approaches to genotoxicity risk assessment II. Use of point-of-departure (PoD) metrics in defining acceptable exposure limits and assessing human risk. *Mutat Res Genet Toxicol Environ Mutagen* 783:66–78. <https://doi.org/10.1016/j.mrgentox.2014.10.008>
- Maertens RM, Long AS, White PA (2017) Performance of the in vitro transgene mutation assay in MutaMouse FE1 cells: evaluation of

- nine misleading (“False”) positive chemicals. *Environ Mol Mutagen* 58:582–591. <https://doi.org/10.1002/em.22125>
- Natarajan AT, Obe G (1986) How do in vivo mammalian assays compare to in vitro assays in their ability to detect mutagens? *Mutat Res* 167:189–201. [https://doi.org/10.1016/0165-1110\(86\)90029-1](https://doi.org/10.1016/0165-1110(86)90029-1)
- National Toxicology Program (1988) NTP Toxicology and Carcinogenesis Studies of 4-Hexylresorcinol (CAS No. 136–77-6) in F344/N Rats and B6C3F1 Mice (Gavage Studies). *Natl Toxicol Program Tech Rep Ser* 330:1–166
- National Toxicology Program (2015) Identifying research needs for assessing safe use of high intakes of folic acid
- O’Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR (2011) Open babel: an open chemical toolbox. *J Cheminform*. <https://doi.org/10.1186/1758-2946-3-33>
- OECD (2016) Test no. 487: in vitro mammalian cell micronucleus test
- Parry EM, Parry JM, Corso C, Doherty A, Haddad F, Hermine TF, Johnson G, Kayani M, Quick E, Warr T, Williamson J (2002) Detection and characterization of mechanisms of action of aneugenic chemicals. *Mutagenesis* 17:509–521. <https://doi.org/10.1093/mutage/17.6.509>
- Paul Friedman K, Gagne M, Loo L, Karamertzanis P, Netzeva T, Sobanski T, Franzosa JA, Richard AM, Lougee RR, Gissi A, Lee JJ, Angrish M, Dorne JL, Foster S, Raffaele K, Bahadori T, Gwinn MR, Lambert J, Whelan M, Rasenberg M, Barton-Maclaren T, Thomas RS (2020) Utility of in vitro bioactivity as a lower bound estimate of in vivo adverse effect levels and in risk-based prioritization. *Toxicol Sci* 173:202–225. <https://doi.org/10.1093/toxsci/kfz201>
- Pearce RG, Setzer RW, Strobe CL, Wambaugh JF, Sipes NS (2017) htk: R package for high-throughput toxicokinetics. *J Stat Softw* 79:1–26. <https://doi.org/10.18637/jss.v079.i04>
- Pradeep P, Patlewicz G, Pearce R, Wambaugh J, Wetmore BA, Judson R (2020) Using chemical structure information to develop predictive models for in vitro toxicokinetic parameters to inform high-throughput risk-assessment. *Comput Toxicol*. <https://doi.org/10.1016/j.comtox.2020.100136>
- Rocio L, Friedman M, Marroni D, Maynor T, Chepelev NL (2017) Impact of acrylamide on calcium signaling and cytoskeletal filaments in testes from F344 rat. *Int J Toxicol* 36:124–132. <https://doi.org/10.1177/1091581817697696>
- Richard AM, Judson RS, Houck KA, Grulke CM, Volarath P, Thillainadarajah I, Yang C, Rathman J, Martin MT, Wambaugh JF, Knudsen TB, Kancherla J, Mansouri K, Patlewicz G, Williams AJ, Little SB, Crofton KM, Thomas RS (2016) ToxCast chemical landscape: paving the road to 21st century toxicology. *Chem Res Toxicol* 29:1225–1251. <https://doi.org/10.1021/acs.chemrestox.6b00135>
- Rosefort C, Fauth E, Zankl H (2004) Micronuclei induced by aneugens and clastogens in mononucleate and binucleate cells using the cytokinesis block assay. *Mutagenesis* 19:277–284. <https://doi.org/10.1093/mutage/geh028>
- Rothfuss A, O’Donovan M, De Boeck M, Brault D, Czich A, Custer L, Hamada S, Plappert-Helbig U, Hayashi M, Howe J, Kraynak AR, van der Leede B, Nakajima M, Priestley C, Thybaud V, Saigo K, Sawant S, Shi J, Storer R, Struwe M, Vock E, Galloway S (2010) Collaborative study on fifteen compounds in the rat-liver Comet assay integrated into 2- and 4-week repeat-dose studies. *Mutat Res* 702:40–69. <https://doi.org/10.1016/j.mrgentox.2010.07.006>
- Rotroff DM, Wetmore BA, Dix DJ, Ferguson SS, Clewell HJ, Houck KA, Lecluyse EL, Andersen ME, Judson RS, Smith CM, Sochaski MA, Kavlock RJ, Boellmann F, Martin MT, Reif DM, Wambaugh JF, Thomas RS (2010) Incorporating human dosimetry and exposure into high-throughput in vitro toxicity screening. *Toxicol Sci* 117:348–358. <https://doi.org/10.1093/toxsci/kfq220>
- Scholz S, Sela E, Blaha L, Braunbeck T, Galay-Burgos M, García-Franco M, Guinea J, Klüver N, Schirmer K, Tanneberger K, Tobor-Kaplon M, Witters H, Belanger S, Benfenati E, Creton S, Cronin MTD, Eggen RIL, Embry M, Ekman D, Gourmelon A, Halder M, Hardy B, Hartung T, Hubesch B, Jungmann D, Lampi MA, Lee L, Léonard M, Küster E, Lillicrap A, Luckenbach T, Murk AJ, Navas JM, Peijnenburg W, Repetto G, Salinas E, Schüürmann G, Spielmann H, Tollefsen KE, Walter-Rohde S, Whale G, Wheeler JR, Winter MJ (2013) A European perspective on alternatives to animal testing for environmental hazard identification and risk assessment. *Regul Toxicol Pharmacol* 67:506–530. <https://doi.org/10.1016/j.yrtph.2013.10.003>
- Seshan VE (2018) *clinfun: Clinical Trial Design and Data Analysis Functions*. 1.0.15 ed.
- Shibai-Ogata A, Kakinuma C, Hioki T, Kasahara T (2011) Evaluation of high-throughput screening for in vitro micronucleus test using fluorescence-based cell imaging. *Mutagenesis* 26:709–719. <https://doi.org/10.1093/mutage/ger037>
- Sipes NS, Wambaugh JF, Pearce R, Auerbach SS, Wetmore BA, Hsieh J, Shapiro AJ, Svoboda D, DeVito MJ, Ferguson SS (2017) An intuitive approach for predicting potential human health risk with the Tox21 10k library. *Environ Sci Technol* 51:10786–10796. <https://doi.org/10.1021/acs.est.7b00650>
- Slob W, Moerbeek M, Rauniomaa E, Piersma AH (2005) A statistical evaluation of toxicity study designs for the estimation of the benchmark dose in continuous endpoints. *Toxicol Sci* 84:167–185
- Soeteman-Hernández LG, Fellows MD, Johnson GE, Slob W (2015) Correlation of in vivo versus in vitro benchmark doses (BMDs) derived from micronucleus test data: a proof of concept study. *Toxicol Sci* 148:355–367. <https://doi.org/10.1093/toxsci/kfv189>
- Soeteman-Hernández LG, Johnson GE, Slob W (2016) Estimating the carcinogenic potency of chemicals from the in vivo micronucleus test. *Mutagenesis* 31:347–358. <https://doi.org/10.1093/mutage/gev043>
- Thomas RS, Philbert MA, Auerbach SS, Wetmore BA, Devito MJ, Cote I, Rowlands JC, Whelan MP, Hays SM, Andersen ME, Meek MEB, Reiter LW, Lambert JC, Clewell HJ, Stephens ML, Zhao QJ, Wesselkamper SC, Flowers L, Carney EW, Pastoor TP, Petersen DD, Yauk CL, Nong A (2013) Incorporating new technologies into toxicity testing and risk assessment: moving from 21st century vision to a data-driven framework. *Toxicol Sci* 136:4–18. <https://doi.org/10.1093/toxsci/kft178>
- Thomas RS, Bahadori T, Buckley TJ, Cowden J, Deisenroth C, Dionisio KL, Frithsen JB, Grulke CM, Gwinn MR, Harrill JA, Higuchi M, Houck KA, Hughes MF, Hunter ES, Isaacs KK, Judson RS, Knudsen TB, Lambert JC, Linnenbrink M, Martin TM, Newton SR, Padilla S, Patlewicz G, Paul-Friedman K, Phillips KA, Richard AM, Sams R, Shafer TJ, Setzer RW, Shah I, Simmons JE, Simmons SO, Singh A, Sobus JR, Strynar M, Swank A, Tornerovalez R, Ulrich EM, Villeneuve DL, Wambaugh JF, Wetmore BA, Williams AJ (2019) The next generation blueprint of computational toxicology at the U.S. Environ Protect Agency Toxicol Sci 169:317–332. <https://doi.org/10.1093/toxsci/kfz058>
- Titenko-Holland N, Jacob RA, Shang N, Balaraman A, Smith MT (1998) Micronuclei in lymphocytes and exfoliated buccal cells of postmenopausal women with dietary changes in folate. *Mutat Res* 417:101–114. [https://doi.org/10.1016/s1383-5718\(98\)00104-1](https://doi.org/10.1016/s1383-5718(98)00104-1)
- Van Hummelen P, Elhajouji A, Kirsch-Volders M (1995) Clastogenic and aneugenic effects of three benzimidazole derivatives in the in vitro micronucleus test using human lymphocytes. *Mutagenesis* 10:23–29. <https://doi.org/10.1093/mutage/10.1.23>
- Walker R (2005) 835 hexylresorcinol, 4- (WHO food additives series 35)
- Wallace Hayes A, Muriana A, Alzualde A, Fernandez DB, Iskandar A, Peitsch MC, Kuczaj A, Hoeng J (2020) Alternatives to animal use in risk assessment of mixtures. *Int J Toxicol* 39:165–172. <https://doi.org/10.1177/1091581820905088>

- Wambaugh JF, Setzer RW, Reif DM, Gangwal S, Mitchell-Blackwood J, Arnot JA, Joliet O, Frame A, Rabinowitz J, Knudsen TB, Judson RS, Egeghy P, Vallero D, Cohen Hubal EA (2013) High-throughput models for exposure-based chemical prioritization in the ExpoCast project. *Environ Sci Technol* 47:8479–8488. <https://doi.org/10.1021/es400482g>
- Wetmore BA (2015) Quantitative in vitro-to-in vivo extrapolation in a high-throughput environment. *Toxicology* 332:94–101. <https://doi.org/10.1016/j.tox.2014.05.012>
- Wetmore BA, Wambaugh JF, Ferguson SS, Sochaski MA, Rotroff DM, Freeman K, Clewell HJ, Dix DJ, Andersen ME, Houck KA, Allen B, Judson RS, Singh R, Kavlock RJ, Richard AM, Thomas RS (2012) Integration of dosimetry, exposure, and high-throughput screening data in chemical toxicity assessment. *Toxicol Sci* 125:157–174. <https://doi.org/10.1093/toxsci/kfr254>
- Wetmore BA, Wambaugh JF, Allen B, Ferguson SS, Sochaski MA, Setzer RW, Houck KA, Strobe CL, Cantwell K, Judson RS, LeCluyse E, Clewell HJ, Thomas RS, Andersen ME (2015) Incorporating high-throughput exposure predictions with dosimetry-adjusted In Vitro bioactivity to inform chemical toxicity testing. *Toxicol Sci* 148:121–136. <https://doi.org/10.1093/toxsci/kfv171>
- White PA, Johnson GE (2016) Genetic toxicology at the crossroads—from qualitative hazard evaluation to quantitative risk assessment. *Mutagenesis* 31:233–237. <https://doi.org/10.1093/mutage/gew011>
- White PA, Douglas GR, Gingerich J, Parfett C, Shwed P, Seligy V, Soper L, Berndt L, Bayley J, Wagner S, Pound K, Blakey D (2003) Development and characterization of a stable epithelial cell line from Muta Mouse lung. *Environ Mol Mutagen* 42:166–184. <https://doi.org/10.1002/em.10185>
- White PA, Long AS, Johnson GE (2020) Quantitative interpretation of genetic toxicity dose-response data for risk assessment and regulatory decision-making: current status and emerging priorities. *Environ Mol Mutagen* 61:66–83. <https://doi.org/10.1002/em.22351>
- Willcocks D, Onyon L, Diver B, Organization WH, Safety IPoC (1998) Triglycidyl isocyanurate. World Health Organization, Geneva
- Williams AJ, Grulke CM, Edwards J, McEachran AD, Mansouri K, Baker NC, Patlewicz G, Shah I, Wambaugh JF, Judson RS, Richard AM (2017) The CompTox Chemistry Dashboard: a community data resource for environmental chemistry. *J Cheminform*. <https://doi.org/10.1186/s13321-017-0247-6>
- Woutersen RA, Jonker D, Stevenson H, te Biesebeek JD, Slob W (2001) The benchmark approach applied to a 28 day toxicity study with Rhodorsil Silane in rats. the impact of increasing the number of dose groups. *Food Chem Toxicol* 39:697–707. [https://doi.org/10.1016/S0278-6915\(01\)00015-1](https://doi.org/10.1016/S0278-6915(01)00015-1)
- Yamada M, Honma M (2018) Summarized data of genotoxicity tests for designated food additives in Japan. *Genes Environ*. <https://doi.org/10.1186/s41021-018-0115-2>
- Yener Y (2013) Effects of long term low dose acrylamide exposure on rat bone marrow polychromatic erythrocytes. *Biotech Histochem* 88:356–360. <https://doi.org/10.3109/10520295.2013.790561>
- Zeiger E, Recio L, Fennell TR, Haseman JK, Snyder RW, Friedman M (2009) Investigation of the low-dose response in the in vivo induction of micronuclei and adducts by acrylamide. *Toxicol Sci* 107:247–257. <https://doi.org/10.1093/toxsci/kfn214>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Byron Kuo¹ · Marc A. Beal² · John W. Wills^{1,3} · Paul A. White¹ · Francesco Marchetti¹ · Andy Nong¹ · Tara S. Barton-Maclaren⁴ · Keith Houck⁵ · Carole L. Yauk^{1,6} 

¹ Environmental Health Science and Research Bureau, Environmental and Radiation Health Sciences Directorate, Healthy Environment and Consumer Safety Branch, Health Canada, Ottawa, ON, Canada

² Bureau of Chemical Safety, Food Directorate, Health Products and Food Branch, Health Canada, Ottawa, ON, Canada

³ Biominerals Research, Department of Veterinary Medicine, University of Cambridge, Cambridge CB3 0ES, UK

⁴ Existing Substances Risk Assessment Bureau, Safe Environments Directorate, Healthy Environment

and Consumer Safety Branch, Health Canada, Ottawa, ON, Canada

⁵ National Center for Computational Toxicology, Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, NC 27711, USA

⁶ Present Address: Department of Biology, University of Ottawa, 30 Marie Curie Private, Room 269, Ottawa, ON K1N 6N5, Canada