Environmental and Molecular Mutagenesis 61:34–41 (2020)

*Commentary*

# Mutation as a Toxicological Endpoint for Regulatory Decision-Making

**Robert H. Heflich,[1]* George E. Johnson,[2] Andreas Zeller,[3] Francesco Marchetti,[4] George R. Douglas,[4] Kristine L. Witt,[5] B. Bhaskar Gollapudi,[6] and Paul A. White[4]**

[1]*U.S. Food and Drug Administration, National Center for Toxicological Research, Jefferson, Arkansas*

[2]*Swansea University, Swansea, UK*

[3]*Roche Pharma Research and Early Development, Roche Innovation Center Basel, F. Hoffmann-La Roche Ltd, Basel, Switzerland*

[4]*Environmental Health Science and Research Bureau, Health Canada, Ottawa, Canada*

[5]*National Institutes of Health, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina*

[6]*Exponent, Inc, Alexandria, Virginia*

Mutations induced in somatic cells and germ cells are responsible for a variety of human diseases, and mutation *per se* has been considered an adverse health concern since the early part of the 20th Century. Although *in vitro* and *in vivo* somatic cell mutation data are most commonly used by regulatory agencies for hazard identification, that is, determining whether or not a substance is a potential mutagen and carcinogen, quantitative mutagenicity dose–response data are being used increasingly for risk assessments. Efforts are currently underway to both improve the measurement of mutations and to refine the computational methods used for evaluating mutation data. We recommend continuing the development of these approaches with the objective of establishing consensus regarding the value of including the quantitative analysis of mutation *per se* as a required endpoint for comprehensive assessments of toxicological risk. Environ. Mol. Mutagen. 61:34–41, 2020. 

**Key words: germ cell mutation; somatic cell mutation; somatic mosaicism; point of departure; error-corrected next-generation sequencing**

## INTRODUCTION

Mutagenicity data most often are used in a dichotomous "screen and bin" approach to identify mutagens (i.e., hazard identification or hazard ID) (Gollapudi, 2017). This paradigm is predicated on two assumptions: that mutagens are rare and that they increase mutation frequencies above the ever-present background throughout the exposure range (i.e., there are no true thresholds). It has been known for some time, however, that neither of these assumptions is necessarily true. In general, mutagens are not rare. For instance, Zeiger and Margolin (2000) estimated that about 20% of the >60,000 organic compounds in commercial use are mutagenic, and in a survey of marketed pharmaceuticals, Snyder (2009) found that 7.1% were positive in the Ames test and 26.1% positive for *in vitro* chromosome aberrations. In addition, it is now commonly accepted that mutagenicity dose–response data for agents such as directly reactive alkylating agents, polycyclic aromatic hydrocarbons, and nitroarenes can have distinct nonlinear responses. Low doses can display little or no apparent mutagenicity increases relative to the background, while higher doses show appreciable response increases, displaying at least "practical" thresholds for mutagenicity (e.g., Doak et al., 2007; reviewed in Jenkins et al., 2010; MacGregor et al., 2015; Nohmi and Tsuzuki, 2016; White and Johnson, 2016; COM, 2018).

***Correspondence to:** Robert H. Heflich, Division of Genetic and Molecular Toxicology, US FDA/NCTR, 3900 NCTR Rd., Jefferson, AR 72079. E-mail: robert.heflich@fda.hhs.gov

This presentation is solely the opinion of the authors and is not to be interpreted as the position or policy of the U.S. Food and Drug Administration, Swansea University, F. Hoffmann-La Roche, the U.S. National Institutes of Health/National Institute of Environmental Health Sciences or Exponent, Inc.

Received 20 July 2019; Revised 9 September 2019; Accepted 14 September 2019

DOI: 10.1002/em.22338

Published online 10 October 2019 in Wiley Online Library (wileyonlinelibrary.com).

Nevertheless, rather than being used quantitatively to establish an exposure limit, in most jurisdictions classification as a mutagen is used to modify the assumptions used in quantitative risk assessments for health outcomes such as cancer or reproductive damage. This Commentary advocates recognition that mutation is a *bona fide* toxicological endpoint, and recommends employing quantitative analyses of dose–response mutagenicity data, along with other available toxicity information, to determine Health-Based Guidance Values (HBGVs) and/or Margins of Exposure (MOEs) for regulatory decision-making (ECHA, 2012; Dourson et al., 2013; WHO, 2014; ICH, 2016, 2017; Hardy et al., 2017).

Mutations in germ cells have been known to cause adverse effects on human health since the early 20th Century. In this Special Issue, Marchetti et al. review germ cell mutagenesis, where a single mutation can result in a phenotype that affects the resulting offspring, for example, mutation resulting in Lesch–Nyhan syndrome, sickle cell anemia. With somatic cell mutagenesis, mutants may clonally expand to a degree that affects the phenotype of an organ or tissue, with the expanded mutant cell populations potentially contributing to a variety of diseases. Somatic cancer driver mutations confer phenotypic properties promoting cellular expansion and the malignancy of the resulting cell population (see Harris et al., 2020). Also, somatic mutations that are induced early in development can expand through normal cell division to populate a large portion of a tissue. Termed *somatic mosaicism*, these mixtures of wild-type cells and mutant cell clones are being increasingly recognized as a cause of human diseases other than cancer, for example, paroxysmal nocturnal hemoglobinuria and perhaps autism (see article by Godschalk et al., 2020).

Around the time that the Environmental Mutagen Society was founded, the attention of the regulatory community was directed toward the relationship between somatic cell mutation and cancer, and the use of short-term *in vitro* testing to rapidly and efficiently identify mutagens and potential carcinogens (Zeiger, 2004; White and Johnson, 2016; see article by DeMarini, 2020). The resulting screen-and-bin hazard ID paradigm led to a lack of emphasis on germ cell mutation and somatic mosaicism as health concerns. This was partially due to the absence of practical *in vivo* assays for these endpoints, particularly assays able to detect treatment-induced gene mutations in humans and in the model rodent systems most relevant to health risk assessment. Although transgenic rodent (TGR) assays measure both germ cell and somatic cell mutations, these tests only became available after short-term tests for the identification of environmental mutagens had become established. Moreover, the use of TGR dose–response data for risk assessment is limited to genetically manipulated rodent models because the endpoint (transgene mutation) cannot be evaluated in humans or conventional animals.

Prior to the development of TGR assays, germ cell mutagenicity testing could be done, but it required extraordinary numbers of animals (e.g., Russell et al., 1998). With respect to somatic cell mosaicism, it has been demonstrated definitively in humans and animals (e.g., Lupski, 2013; Erickson, 2014; Mortazavi et al. 2003; McConnell et al., 2017; Meier et al., 2017; Luderer et al., 2019; Yizhak et al., 2019); however, detection and analysis of mosaicism is challenging (Campbell et al., 2015) and there are no standardized strategies to specifically assess this endpoint. Currently, germ cell mutation is considered for toxicological risk assessments (e.g., Health Canada, 1993; United Nations, 2017), and Adverse Outcome Pathways (AOPs) have been developed that treat germ cell mutation as an Adverse Outcome (Yauk et al., 2015b; Marchetti et al., 2016; see Sasaki et al. 2020). Because the phenotypes associated with a germ cell mutation have the potential to be expressed in offspring carrying the mutation, germ cell mutations conferring a phenotype are likely to cause adverse health effects. Accordingly, mutation is already considered as a *bona fide* toxicological endpoint in germ cells and can be evaluated both in qualitative and quantitative terms (reviewed by Marchetti et al., 2020).

In contrast, several, often poorly understood steps must occur in the clonal expansion of cells with a somatic cell mutation before an adverse health effect becomes evident. Furthermore, many mutations and mutant clones with the potential for causing disease exist at low levels in apparently healthy individuals (e.g., Tomasetti, 2019; Yizhak et al., 2019; Harris et al., 2020). Nevertheless, since useful methods to assess *in vivo* mutation in relevant species, including humans, now exist, and potentially better methods are on the horizon, they should be increasingly employed for generating mutation data that can be interpreted in a quantitative manner for regulatory decision-making. Because somatic mutations have a marked potential for causing adverse health effects, they, like germ cell mutations, should be considered Adverse Outcomes, as defined for the development of AOPs, and *bona fide* toxicological endpoints. With advances in our knowledge about the relationships between mutation and disease, it is conceivable that, in the future, mutation data also can be used routinely as a quantitative biomarker for cancer and other diseases associated with somatic mosaicism (see below).

While the hazard ID paradigm has clear benefits in terms of screening substances for genotoxic activity and in prioritizing research and follow-up testing, there are several compelling benefits to using mutation as a toxicological endpoint. Relying solely on hazard, ID testing effectively eliminates exposure to mutagens (and by extension, carcinogens) in regulated consumer products. An unintended consequence of eliminating exposure to any dose of a mutagen, however, might be the loss of potentially valuable substances (e.g., drugs) that are unlikely to induce mutations in humans, due to, for instance, the route or level of exposure,

effective DNA repair at low doses, and/or active detoxification at low doses (MacGregor et al., 2015). Besides the loss of valuable products, the screen-and-bin paradigm also has the potential for the unnecessary diversion of resources to address existing exposures that have no practical consequence. In fact, endogenous exposure to known mutagens already occurs in humans (e.g., Nakamura et al., 2014). In contrast, if mutation *per se* is considered a toxicological endpoint, and compensatory mechanisms (e.g., DNA repair or detoxification) acknowledged as affecting dose–responses, quantitative analysis of mutagenicity dose–response data can be used as the basis for risk assessments. More specifically, Point of Departure (PoD) metrics such as No Observed Adverse Effect Level (NOAEL) or Benchmark Dose (BMD), suitably adjusted using uncertainty/safety/adjustment factors, can be used to define human exposure limits, and such limits can be compared to actual human exposure data. This concept is not foreign to regulatory science, and indeed the following examples demonstrate that such concepts have been employed previously.

## MUTATION AS A TOXICOLOGICAL ENDPOINT: EXAMPLES WHERE DOSE–RESPONSE DATA HAVE BEEN USED

In several instances, quantitative interpretation of mutagenicity dose–response data has been used for risk assessment and regulatory decision-making. For example, in 2007, a batch of the AIDS drug Viracept became contaminated with a genotoxic carcinogen, ethyl methanesulfonate (EMS; Müller and Singer, 2009). In the absence of rodent tumor data suitable for assessing the cancer risk to exposed patients, the drug manufacturer, in consultation with the European Medicines Agency, the Committee for Medicinal Products for Human Use, and other experts conducted a comprehensive toxicological evaluation of EMS that included an extensive *in vivo* mutagenicity study. Dose–response gene mutagenicity data generated using the transgenic MutaMouse, and dose–response micronucleus data from the mouse bone marrow micronucleus test, showed nonlinear patterns for EMS genotoxicity. Their data supported the existence of a "practical threshold" below which the likelihood of a response above background was negligible. The analysis indicated that the oral NOAEL for EMS-induced transgene mutation, that is, the highest dose that failed to elicit a response significantly greater than the control, was 454-fold greater than the maximum daily human intake of the EMS contaminant (Müller et al., 2009). The data were used in a risk assessment of EMS, which concluded that patients using the contaminated drug product were unlikely to be at increased risk of cancer.

In a more recent example, the Committee for Risk Assessment (RAC) of the European Chemicals Agency (ECHA) used mutagenicity data in responding to a European Commission request to evaluate the Occupational Exposure Limit (OEL) for benzene (ECHA, 2018). A combination of epidemiological and experimental data indicated that bone marrow and the hematological system are targets for benzene-induced toxicity, with acute myeloid leukemia being the major health concern from occupational exposure. The RAC concluded that the carcinogenicity of benzene was closely associated with its aneugenicity and clastogenicity and that, in their opinion, the OEL could be a dose with a negligible risk of aneugenicity and clastogenicity in the bone marrow.

In these two cases, mutation was used not only as a toxicological endpoint but also as a quantitative biomarker of an adverse outcome that is mechanistically linked to a particular disease, that is, cancer. As indicated previously, the exact mechanistic and quantitative relationship between mutant frequency and cancer outcome is affected by many factors (see Harris et al., 2020). The regulators, however, considered the multiple lines of evidence indicating that mutations produced by environmental substances are causative factors in cancer, including cancer in humans, to be sufficiently strong for conducting quantitative assessments of risk. In the Viracept-EMS case, using the available mutation data as a basis for a regulatory decision was considered preferable to waiting for the results of a cancer bioassay before deciding on how best to protect the health of the affected patients.

Finally, we cite two examples where dose–response mutation data were considered but not used for the final risk assessment. In a regulatory decision regarding formaldehyde, the Bundesinstitut für Risikobewertung (BfR) decided that mutation was not the most appropriate quantitative biomarker for cancer so other toxicities were used for the risk assessment (BfR, 2006). Thus, the cancer risk assessment conducted by the BfR was based on dose–response data for cytotoxicity and compensatory cell proliferation. Here, mutation (inferred from DNA damage) was one of the toxicological endpoints used by the BfR for conducting the risk assessment, but as is common regulatory practice, the more sensitive endpoints were chosen for calculating HBGVs for formaldehyde exposure. Similarly, in their screening assessment of natural gas condensates (NGCs) and NGC components, the Government of Canada quantitatively considered *in vivo* clastogenicity dose–response data. These genotoxicity dose–response data, however, were deemed to be inadequate relative to those for other toxicities, which were ultimately used as the basis for the risk assessment (Environment Canada and Health Canada, 2016).

## RECENT DEVELOPMENTS, ISSUES, AND POSSIBLE ADVANCEMENTS

Encouraged by the acceptance of the risk assessment conducted in response to the Viracept contamination

incident, the Genetic Toxicology Technical Committee of the Health and Environmental Sciences Institute established a program to develop computational methods for analyzing genetic toxicology dose–response data, with the objective of selecting PoD metrics (e.g., NOGEL, BMD) for use in quantitative risk assessments (Gollapudi et al., 2013; Johnson et al., 2014). Some of the authors of this Commentary have been part of this effort (see White and Johnson, 2016), and the Commentary by White et al. (2020) updates the status of these continuing efforts. Additionally, a case study described in this Special Issue outlines how these quantitative methods can be used to establish HBGVs for benzene (Luijten et al., 2020).

In the context of these case studies, the term "mutation" encompasses both large and small sequence changes including the relatively small DNA sequence alterations associated with many gene mutations, as well as larger changes such as structural and numerical chromosomal aberrations. All of these genetic changes are at least potentially heritable, and all have been associated with human disease phenotypes (e.g., Campbell et al., 2015; Stenson et al., 2017). In the absence of information that would narrow the analysis to a particular class of mutation, as was the case with the benzene occupational exposure example described above, it is important to note that a comprehensive analysis of mutation as a toxicological endpoint would include gene mutation, clastogenicity, and aneugenicity.

Measuring both gene mutation and chromosomal mutations using standard genotoxicity assays, however, may not yield a comprehensive assessment of mutagenic activity because these assays have inherent limitations and biases. With gene mutation assays, only a single or a small number of loci are typically examined, and depending on the locus, the mutations detected are generally only a small subset of all possible genetic alterations. For TGR assays, base substitutions and small insertions and deletions are measured most commonly in a very specific target gene that has a different functionality than those that lead to disease (Lambert et al., 2005; Boverhof et al., 2011; Nohmi et al., 2017). Gene mutation, however, can be caused by megabase deletions, mitotic recombination spanning a large portion of a chromosome, and even aneuploidy (e.g., Wang et al., 2009). Also, mutagens differ not only by the types of sequence changes they cause but where they cause them (Thilly, 1990; Chuang and Li, 2004; Lynch, 2010; Kucab et al., 2019); there are many examples demonstrating that mutations in both transgenes and endogenous genes differ with respect to both frequency and type (Moore et al., 1989; DeMarini et al., 1989; Monroe et al., 1998; Walker et al., 1999). Differences in locus-specific mutation frequencies appear to be related to sequence context and methylation status, target size, overall gene structure, and chromosome type (i.e., autosomal *vs*. sex-determining).

For chromosome mutations, the peripheral blood micronucleus assay measures the frequency of an induced response that is not heritable, although the mechanism involved in micronucleus formation is believed also to result in heritable structural and numerical chromosomal alterations. Likewise, the structural chromosome aberration assay directly measures known heritable aberrations only if special techniques are used (e.g., chromosome banding or painting).

The TGR assay, and assays that can be conducted in both humans and rodents like the *PIG-A*/*Pig-a* gene mutation assay and the peripheral blood chromosome aberration and MN assays, measure the appropriate endpoints in the appropriate species for assessing potential adverse outcomes. These assays, however, have limitations in terms of what they measure and the tissues and animal models that can be evaluated, and, thus, have the potential to give imprecise measurements of mutation frequency. The justification for the quantitative interpretation of dose–responses from these assays can be summarized as follows. What is measured, both in humans and in animal models, is sufficiently related to the mutations actually caused by the exposure that they provide useful, if imprecise data for assessing and managing potential adverse health outcomes.

We anticipate, however, that the newer error-corrected next-generation sequencing (EC-NGS) technologies that directly measure mutations at any location in the genome or over the entire genome (Dong et al., 2017; Revollo et al., 2018; Guo et al. 2018; Zhang et al., 2019; see Salk and Kennedy, 2020) will: (1) provide a more comprehensive evaluation of smaller gene mutations and hence more appropriate estimates of mutation frequency; and, (2) be able to evaluate gene mutation in any tissue of any animal without the need for using transgene reporters. With the use of longer-read NGS technologies, EC-NGS also may be able to evaluate chromosomal mutation. Although these novel methods have great promise for comprehensively quantifying mutations induced by exogenous agents, they require further development and validation before they can be used for regulatory applications.

A final issue concerning the use of mutation as a toxicological endpoint is that quantitative evaluations of mutations, including the examples cited above, generally have used mutation in the context of cancer as the sole human health concern. While demonstrating the value of evaluating mutation quantitatively, this restricted focus may have limited how the analysis was performed. The International Workshops on Genotoxicity Testing committee report of Yauk et al. (2015a) indicates that there are quantitative differences between the induction of germ cell and somatic cell mutations, and in some cases, higher frequencies are detected for germ cell than somatic cell mutation. For example, 1,1,-dimethylhydrazine, dimethylnitrosamine, diethylnitrosamine, and beta-propiolactone produced negative results for bone marrow micronucleus induction, an endpoint commonly used for cancer hazard assessment, but increased micronucleus frequencies in male rodent germline cells (Cliet et al., 1993; Yauk et al., 2015a).

Although some of these differences are undoubtedly due to differences in metabolism and the exposure of the tissues used for the assays, these observations suggest that it may not be health-protective to assume that regulating on somatic cell effects, specifically those involved with cancer, also protects against germ cell mutation (for further discussion on this topic, see Marchetti et al., 2020). In addition, it may be sufficiently health-protective to measure somatic mutant frequencies without confirming clonal expansion, but it also may be necessary to evaluate transplacental mutagenesis to accurately quantify induced mutations involved in somatic mosaicism. The value of evaluating mutation following transplacental exposure was recently demonstrated by Meier et al. (2017) who found that tissues like brain, which are resistant to mutation induction in adult animals, maybe readily mutagenized by *in utero* exposure to a potent mutagen.

All these issues make it a major challenge to evaluate the different types of mutations that occur in different tissues, at different life stages, and with differing manifestation times in a manner that is consistent with global initiatives to conserve animal resources. The current deliberations to revise the OECD Test Guideline for the TGR gene mutation assay (TG 488; OECD, 2013) indicate that it may be possible to devise a compromise protocol that is capable of measuring the induction of both somatic cell and germ cell mutation in a single set of animals with acceptable sensitivity (Marchetti et al., 2018a,b). Ideally, if the mutation is to be used as a *bona fide* toxicological endpoint, its analysis should be integrated into standard repeat-dose toxicology studies (e.g., OECD TG 407, 415, and 422). It also may be possible to integrate mutagenicity assays into reproductive toxicity studies so that the induction of somatic mosaicism due to exposure during early developmental stages can be assessed. Integration into general toxicity or reproductive toxicity testing also may enable comparison of mutation data with other toxicity endpoints to establish the most sensitive endpoint to use in establishing a safe level of exposure, a common regulatory practice. There is currently resistance to using transgenic animals in general toxicology assays primarily due to a lack of adequate comparative data with non-mutation endpoints in wild-type rodent strains. But, as noted above, ECNGS may overcome this problem by allowing mutation analysis in any animal.

## NEXT STEPS

If increases in mutation induced by exogenous substances is a *bona fide* adverse health effect, as extensive scientific data indicate, we contend that regulatory risk assessments should treat mutation as a toxicological endpoint and employ quantitative analyses to determine exposure levels below which the risk of adverse outcomes (in this case mutation) can be deemed negligible (i.e., establish HBGVs). However, because our knowledge of how to use mutation data to quantitatively assess disease outcomes requires further refinement, we are not currently recommending the next logical step, that quantitative mutation analysis routinely replace the measurement of tumors and other adverse health effects in making regulatory decisions. We do, however, suggest that this should be a long-term goal for the field of genetic toxicology. For the time being, however, regulators who may be considering the use of mutation as a toxicological endpoint will benefit from a critical review of case studies that compare HBGVs and/or MOEs derived from mutagenicity studies with those derived from carcinogenicity studies or studies of other disease-endpoints. In addition, the following is recommended:

1. Where possible, samples should be collected from ongoing cancer bioassays and other repeat-dose toxicity studies to enable parallel and/or future mutational analyses, for example, for micronucleus, *Pig-a* mutation, and error-corrected NGS analysis (as done in Mittelstaedt et al., 2019). This recommendation applies to routine toxicology testing that is not normally conducted in TGRs.
2. Advanced DNA sequencing methods should be validated for somatic and germ cell analyses so that they can be used to comprehensively measure induced mutations directly.
3. Protocols and guidelines for designing and conducting regulatory mutation studies should be developed that enhance the utility of dose–response data for quantitative analysis and determination of HBGVs and/or MOEs. For example, assay guidelines should include recommendations regarding tissue selection, sampling times, number of doses, study design and animal allocation, acceptable data quality and background frequencies, possible assay integration, and computational approaches for dose–response assessment. Current recommendations in OECD TGs should be adjusted to derive maximum value from dose–response data.
4. The uncertainty factors also referred to as extrapolation or adjustment or safety factors that are employed to calculate HBGVs and/or MOEs, should be scrutinized with respect to their applicability to mutagenicity. This complex issue, and the issues surrounding the quantitative interpretation of genotoxicity PoD values more generally, is more thoroughly discussed by White et al. (2020).

## CONCLUSIONS

Mutations take different forms that require different analytical methods to detect, and their induction can differ quantitatively among tissues and particularly for different life stages: all of which make a comprehensive analysis of mutagenesis extremely challenging. However, we argue that practical methods for considering mutation as a

toxicological endpoint exist today and that such methods already have been used for regulatory decisions. The well-established TGR assay can measure gene mutation in any tissue of transgenic rats and mice, and the MN and chromosome aberration assays, as well as the *HPRT*/*hprt* and emerging *PIG-A*/*Pig-a* gene mutation assays, can be used to measure chromosome and gene mutations in both rodent models and humans. Information focusing the analysis on mutations associated with a particular health concern in specific tissues has been helpful in conducting these studies.

Additional methods to comprehensively measure mutation and generate mutagenicity dose–response data in both somatic and germ cells are developing rapidly, potentially increasing the value of mutation as an endpoint for regulating new and existing substances. In the not-too-distant future, EC-NGS may overcome limitations of the current assays in terms of transferability between species and tissues and the range of genetic changes detected.

Implementing quantitative methods for risk assessment will make better use of mutation data, and will be of value in mitigating the negative health outcomes of environmental exposures. Nevertheless, in light of current regulatory mandates, hazard ID approaches for interpretation of mutagenicity data will continue to be important for screening commercial products under development and for conducting regulatory decision-making; hazard ID data also can be used for prioritizing testing and designing follow-up studies. Integrating mutation analysis with other toxicological endpoints not only optimizes the utilization of resources (e.g., personnel and animals), but also benefits the mutation-based risk assessment by providing the contextual data (e.g., data on absorption, distribution, metabolism and excretion, and non-neoplastic effects) that are gathered routinely in general toxicology testing. By establishing HBGVs and/or MOEs for mutation induction, procedures that are both pragmatic and beneficial to public health can be implemented for effectively managing the risk of mutations, and after appropriate validation, the incidence of diseases associated with mutations. Importantly, the uncertainty factors employed to routinely calculate HGBVs will need to be carefully considered prior to routine application of mutagenicity dose–response analysis in regulatory risk assessments.

## AUTHOR CONTRIBUTIONS

All authors contributed to the writing of the manuscript and all authors agree to its publication.

## REFERENCES

BfR (Bundesinstitut für Risikobewertung). 2006. Toxicological Assessment of Formaldehyde; Opinion of BfR No. 023/2006 of 30 March 2006. Available at: http://m.bfr-meal-studie.de/cm/349/toxicological_assessment_of_formaldehyde.pdf.

Boverhof DR, Chamberlain MP, Elcombe CR, Gonzalez FJ, Heflich RH, Hernandez LG, Jacobs AC, Jacobson-Kram D, Luijten M, Maggi A, et al. 2011. Transgenic animal models in toxicology: Historical perspectives and future outlook. Toxicol Sci 121:207–233.

Campbell IM, Shaw CA, Stankiewicz P, Lupski JR. 2015. Somatic mosaicism: Implications for disease and transmission genetics. Trends Genet 31:382–392.

Chuang JH, Li H. 2004. Functional bias and spatial organization of genes in mutational hot and cold regions in the human genome. PLoS Biol 2:253–263.

Cliet I, Melcion C, Cordier A. 1993. Lack of predictivity of bone marrow micronucleus test versus testis micronucleus test: Comparison with four carcinogens. Mutat Res 292:105–111.

COM (Committee on Mutagenicity of Chemicals in Food, Consumer Products and the Environment). 2018. Statement 2018/S1: Statement on the quantitative approaches to the assessment of genotoxicity data. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/704462/COM_statement_on_quantitative_approaches.pdf.

DeMarini DM. 2020. The mutagenesis moonshot: The propitious beginnings of the Environmental Mutagenesis and Genomics Society. Environ Mol Mutagen 61:8–24.

DeMarini DM, Brockman HE, de Serres FJ, Evans HH, Stankowski LF Jr, Hsie AW. 1989. Specific-locus mutations induced in eukaryotes (especially mammalian cells) by radiation and chemicals: A perspective. Mutat Res 220:11–29.

Doak SH, Jenkins GJS, Johnson GE, Quick E, Parry EM, Parry JM. 2007. Mechanistic influences for mutation induction curves after exposure to DNA-reactive carcinogens. Cancer Res 67:3904–3911.

Dong X, Zhang L, Milholland B, Lee M, Maslov AY, Wang T, Vijg J. 2017. Accurate identification of single-nucleotide variants in whole-genome-amplified single cells. Nat Methods 14:491–493.

Dourson M, Becker RA, Haber LT, Pottenger LH, Bredfeldt T, Fenner-Crisp PA. 2013. Advancing human health risk assessment: Integrating recent advisory committee recommendations. Crit Rev Toxicol 43:467–492.

ECHA (European Chemicals Agency). 2012. Guidance on information requirements and chemical safety assessment. Chapter R.8: Characterisation of dose [concentration]-response for human health. Available at: https://echa.europa.eu/documents/10162/13632/information_requirements_r8_en.pdf/e153243a-03f0-44c5-8808-88af66223258.

ECHA (European Chemicals Agency). 2018. Committee for risk assessment RAC: Opinion on scientific evaluation of occupational exposure limits for benzene. Available at: https://echa.europa.eu/documents/10162/13641/benzene_opinion_en.pdf/4fec9aac-9ed5-2aae-7b70-5226705358c7.

Environment Canada and Health Canada. 2016. Screening assessment, petroleum sector stream approach, natural gas condensates. Government of Canada, Ottawa. Available at: http://www.ec.gc.ca/ese-ees/default.asp?lang=En&n=7933A3C7-1#toc010.

Erickson RP. 2014. Recent advances in the study of somatic mosaicism and diseases other than cancer. Curr Opin Genet Dev 26:73–78.

Godschalk RWL, Yauk CL, van Benthem J, Douglas GR, Marchetti F. 2020. *In utero* exposure to genotoxins leading to genetic mosaicism: An overlooked window of susceptibility in genetic toxicology testing? Environ Mol Mutagen 61:55–65.

Gollapudi BB. 2017. An ongoing journey toward a risk-based testing in genetic toxicology. Curr Opin Toxicol 3:71–74.

Gollapudi BB, Johnson GE, Hernandez LG, Pottenger LH, Dearfield KL, Jeffrey AM, Julien E, Kim JH, Lovell DP, MacGregor JT, et al.

2013. Quantitative approaches for assessing dose-response relationships in genetic toxicology studies. Environ Mol Mutagen 54:8–18.

Guo X, Pan B, Seo JE, Chen Y, Yan J, Mei N, Chen T. 2018. Whole genome sequencing analysis of small and large colony mutants from the mouse lymphoma assay. Arch Toxicol 92:3585–3595.

Hardy A, Benford D, Halldorsson T, Jeger MJ, Knutsen KH, More S, Mortensen A, Naegeli H, Noteborn H, Ockleford C, et al. 2017. Update: Use of the benchmark dose approach in risk assessment. EFSA J 15(1):e04658.

Harris KL, Myers MB, McKim KL, Elespuru RK, Parsons BL. 2020. Rationale and roadmap for developing panels of hotspot cancer driver gene mutations as biomarkers of cancer risk. Environ Mol Mutagen 61:152–175.

Health Canada. 1993. The assessment of mutagenicity, Health Canada Health Protection Branch mutagenicity guidelines. Environ Mol Mutagen 21:15–37.

ICH (International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use). 2016. Q3C(R6): Impurities: Guideline for Residual Solvents. Geneva, Switzerland: ICH. pp. 1–34.

ICH (International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use). 2017. M7(R1): Assessment and Control of DNA Reactive (Mutagenic) Impurities in Pharmaceuticals to Limit Potential Carcinogenic Risk. Geneva, Switzerland: ICH. pp. 1–116.

Jenkins GJS, Zair Z, Johnson GE, Doak SH. 2010. Genotoxic thresholds, DNA repair, and susceptibility in human populations. Toxicology 278:305–310.

Johnson GE, Soeteman-Hernández LG, Gollapudi BB, Bodger OG, Dearfield KL, Heflich RH, Hixon JG, Lovell DP, MacGregor JT, Pottenger LH, et al. 2014. Derivation of point of departure (PoD) estimates in genetic toxicology studies and their potential applications in risk assessment. Environ Mol Mutagen 55:609–623.

Kucab JE, Zou X, Morganella S, Joel M, Nanda AS, Nagy E, Gomez C, Degasperi A, Harris R, Jackson SP, et al. 2019. A compendium of mutational signatures of environmental agents. Cell 177:821–836.

Lambert IB, Singer TM, Boucher SE, Douglas GR. 2005. Detailed review of transgenic rodent mutation assays. Mutat Res 590:1–280.

Luijten M, Ball NS, Dearfield KL, Gollapudi BB, Johnson GE, Madia F, Peel L, Pfuhler S, Settivari RS, ter Burg W, van Benthem J. 2020. Utility of a next generation framework for assessment of genomic damage: A case study using the industrial chemical benzene. Environ Mol Mutagen 61:94–113.

Lupski JR. 2013. Genome mosaicism—One human, multiple genomes. Science 341:358–359.

Luderer U, Meier MJ, Lawson GW, Beal MA, Yauk CL, Marchetti F. 2019. *In utero* exposure to benzo[*a*]pyrene induces ovarian mutations at doses that deplete ovarian follicles in mice. Environ Mol Mutagen 60:410–420.

Lynch M. 2010. Rate, molecular spectrum, and consequences of human mutation. Proc Natl Acad Sci USA 107:961–968.

MacGregor JT, Frötschl R, White PA, Crump KS, Eastmond DA, Fukushima S, Guérard M, Hayashi M, Soeteman-Hernández LG, Johnson GE, et al. 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 783:66–78.

Marchetti F, Douglas GR, Yauk CL. 2020. A return to the origin of the EMGS: Rejuvenating the quest for human germ cell mutagens and determining the risk to future generations. Environ Mol Mutagen 61:42–54.

Marchetti F, Massarotti A, Yauk CL, Pacchierotti F, Russo A. 2016. The adverse outcome pathway (AOP) for chemical binding to tubulin in oocytes leading to aneuploid offspring. Environ Mol Mutagen 57:87–113.

Marchetti F, Aardema MJ, Beevers C, van Benthem J, Godschalk R, Williams A, Yauk CL, Young R, Douglas GR. 2018a. Identifying germ cell mutagens using OECD Test Guideline 488 (transgenic rodent somatic and germ cell gene mutation assays) and integration with somatic cell testing. Mutat Res 832–833:7–18.

Marchetti F, Aardema M, Beevers C, van Benthem J, Douglas GR, Godschalk R, Yauk CL, Young R, Williams A. 2018b. Simulation of mouse and rat spermatogenesis to inform genotoxicity testing using OECD Test Guideline 488. Mutat Res 832–833:19–28.

McConnell MJ, Moran JV, Abyzov A, Akbarian S, Bae T, Cortes-Ciriano I, Erwin JA, Fasching L, Flasch DA, Freed D, et al. 2017. Intersection of diverse neuronal genomes and neuropsychiatric disease: The Brain Somatic Mosaicism Network. Science 356: eaal1641.

Meier MJ, O'Brien JM, Beal MA, Allen B, Yauk CL, Marchetti F. 2017. *In utero* exposure to benzo[*a*]pyrene increases mutation burden in the soma and sperm of adult mice. Environ Helth Perspect 125: 82–88.

Mittelstaedt RA, Dobrovolsky VN, Revollo JR, Pearce MG, Wang Y, Dad A, McKinzie PB, Rosenfeldt H, Yucesoy B, Yeager R, et al. 2019. Evaluation of 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) mutagenicity using *in vitro* and *in vivo Pig-a* assays. Mutat Res 837:65–72.

Monroe JJ, Kort KL, Miller JE, Marino DR, Skopek TR. 1998. A comparative study of in vivo mutation assays: Analysis of *hprt*, *lacI*, and *cII/cI* as mutational targets for *N*-nitroso-*N*-methylurea and benzo[*a*]pyrene in Big Blue™ mice. Mutat Res 421:121–136.

Moore MM, Harrington-Brock K, Doerr CL, Dearfield KL. 1989. Differential mutant quantitation at the mouse lymphoma *tk* and CHO *hgprt* loci. Mutagenesis 4:394–403.

Mortazavi Y, Merk B, McIntosh J, Marsh JCW, Schrezenmeier H, Rutherford TR. 2003. The spectrum of *PIG-A* gene mutations in aplastic anemia/paroxysmal nocturnal hemoglobinuria (AA/PNH): A high incidence of multiple mutations and evidence of a mutational hot spot. Blood 101:2833–2841.

Müller L, Singer T. 2009. EMS in Viracept—The course of events in 2007 and 2008 from the non-clinical safety point of view. Toxicol Lett 190:243–247.

Müller L, Gocke E, Lavé T, Pfister T. 2009. Ethyl methanesulfonate toxicity in Viracept—A comprehensive human risk assessment based on threshold data for genotoxicity. Toxicol Lett 190:317–329.

Nakamura J, Mutlu E, Sharma V, Collins L, Bodnar W, Yu R, Lai Y, Moeller B, Lu K, Swenberg J. 2014. The endogenous exposome. DNA Repair 19:3–13.

Nohmi T, Tsuzuki T. 2016. Possible mechanisms underlying genotoxic thresholds: DNA repair and translesion DNA synthesis. In: Nohmi T, Fukushima S, editors. Thresholds of Genotoxic Carcinogens, from Mechanisms to Regulation. Amsterdam: Elsevier. pp. 49–66.

Nohmi T, Masumura K, Toyoda-Hokaiwado N. 2017. Transgenic rat models for mutagenesis and carcinogenesis. Genes Environ 39: 11–42.

OECD (Organisation for Economic Cooperation and Development). 2013. OECD Guidelines for the testing of chemicals: Transgenic rodent somatic and germ cell gene mutations assays. Test Guideline 488. Available at: http://www.oecd-library.org/docserver/download/9713251e.pdf?expires=1392571255&id=id&accname=guest&checksum=C7400BE10E1B8C223781FCB41BBB876B.

Revollo JR, Dad A, McDaniel LP, Pearce MG, Dobrovolsky VN. 2018. Genome-wide mutation detection by interclonal genetic variation. Mutat Res 829–830:61–69.

Russell WL, Bangham JW, Russell LB. 1998. Differential response of mouse male germ-cell stages to radiation-induced specific-locus and dominant mutations. Genetics 148:1567–1578.

Salk J, Kennedy SR. 2020. Next-generation genotoxicity: Using modern sequencing technologies to assess somatic mutagenesis and cancer risk. Environ Mol Mutagen 61:131–151.

Sasaki JC, Allemang A, Bryce SM, Custer L, Dearfield KL, Dietz Y, Elhajouji A, Escobar PA, Fornace AJ Jr, Froetschl R, Galloway S, Hemmann U, Hendriks G, Li H-H, Luijten M, Ouedraogo G, Peel L, Pfuhler S, Roberts D, Thybaud V, van Benthem J, Yauk CL, Schuler M. 2020. Application of the Adverse Outcome Pathway (AOP) Framework to genotoxic modes of action (MOA). Environ Mol Mutagen 61:114–134.

Snyder RD. 2009. An update on the genotoxicity and carcinogenicity of marketed pharmaceuticals with reference to in silico predictivity. Environ Mol Mutagen 50:235–250.

Stenson PD, Mort M, Ball EV, Evans K, Hayden M, Heywood S, Hussain M, Phillips AD, Cooper DN. 2017. The Human Gene Mutation Database: Towards a comprehensive repository of inherited mutation data for medical research, genetic diagnosis, and next-generation sequencing studies. Hum Genet 136:665–677.

Thilly WG. 1990. Mutational spectrometry in animal toxicity testing. Ann Rev Pharmacol Toxicol 30:369–385.

Tomasetti C. 2019. Mutated clones are the new normal. Science 364: 938–939.

United Nations. 2017. Globally harmonized system of classification and labelling of chemicals (GHS), Part 3: Health Hazards. pp 115–216. Available at: https://www.unece.org/trans/danger/publi/ghs/ghs_rev07/07files_e.html.

Walker VE, Andrews JL, Upton PB, Skopek TR, deBoer JG, Walker DM, Shi X, Sussman HE, Gorelick NJ. 1999. Detection of cyclophosphamide-induced mutations at the Hprt but not the lacI locus in splenic lymphocytes of exposed mice. Environ Mol Mutagen 34:167–181.

Wang J, Sawyer JR, Chen L, Chen T, Honma M, Mei N, Moore MM. 2009. The mouse lymphoma assay detects recombination, deletion, and aneuploidy. Toxicol Sci 109:96–105.

White PA, Johnson GE. 2016. Genetic toxicology at the crossroads—From qualitative hazard evaluation to quantitative risk assessment. Mutagenesis 31:233–237.

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.

WHO (World Health Organization). 2014. Guidance document on evaluating and expressing uncertainty in hazard characterization. IPCS Harmonization Project Document 11. WHO, Geneva, Switzerland. Available at: http://www.inchem.org/documents/harmproj/harmproj/harmproj11.pdf.

Yauk CL, Aardema MJ, van Benthem J, Bishop JB, Dearfield KL, DeMarini DM, Dubrova YE, Honma M, Lupski JR, Marchetti F, et al. 2015a. Approaches for identifying germ cell mutagens: Report of the 2013 IWGT workshop on germ cell assays. Mutat Res 783:36–54.

Yauk CL, Lambert IB, Meek ME, Douglas GR, Marchetti F. 2015b. Development of the adverse outcome pathway "Alkylation of DNA in male premeiotic germ cells leading to heritable mutations" using the OECD's users' handbook supplement. Environ Mol Mutagen 56:724–750.

Yizhak K, Aguet F, Kim J, Hess JM, Kübler K, Grimsby J, Frazer R, Zhang H, Haradhvala NJ, Rosebrock D, et al. 2019. RNA sequence analysis reveals macroscopic somatic clonal expansion across normal tissues. Science 364:eaaw0726.

Zeiger E. 2004. History and rationale of genetic toxicity testing: An impersonal, and sometimes personal, view. Environ Mol Mutagen 44: 363–371.

Zeiger E, Margolin BH. 2000. The proportions of mutagens among chemical in commerce. Regul Toxicol Pharmacol 32:219–225.

Zhang L, Dong X, Lee M, Maslov AY, Wang T, Vijg J. 2019. Single-cell whole-genome sequencing reveals the functional landscape of somatic mutations in B lymphocytes across the human lifespan. Proc Natl Acad Sci USA 116:9014–9019.

*Accepted by*—
C. Yauk