Critical Reviews in Toxicology

ISSN: 1040-8444 (Print) 1547-6898 (Online) Journal homepage: https://www.tandfonline.com/loi/itxc20

Hazard identification, classification, and risk assessment of carcinogens: too much or too little? – Report of an ECETOC workshop

Susan P. Felter, Alan R. Boobis, Philip A. Botham, Alice Brousse, Helmut Greim, Heli M. Hollnagel & Ursula G. Sauer

To cite this article: Susan P. Felter, Alan R. Boobis, Philip A. Botham, Alice Brousse, Helmut Greim, Heli M. Hollnagel & Ursula G. Sauer (2020) Hazard identification, classification, and risk assessment of carcinogens: too much or too little? – Report of an ECETOC workshop, Critical Reviews in Toxicology, 50:1, 72-95, DOI: 10.1080/10408444.2020.1727843

To link to this article: https://doi.org/10.1080/10408444.2020.1727843

Published online: 05 Mar 2020.

CRITICAL REVIEWS IN TOXICOLOGY
2020, VOL. 50, NO. 1, 72–95
https://doi.org/10.1080/10408444.2020.1727843

Taylor & Francis
Taylor & Francis Group

REVIEW ARTICLE

OPEN ACCESS Check for updates

Hazard identification, classification, and risk assessment of carcinogens: too much or too little? – Report of an ECETOC workshop

Susan P. Felter[a], Alan R. Boobis[b], Philip A. Botham[c], Alice Brousse[d], Helmut Greim[e], Heli M. Hollnagel[f] and Ursula G. Sauer[g]

[a]Procter & Gamble, Mason, OH, USA; [b]Imperial College London, London, UK; [c]Syngenta Jealott's Hill, Bracknell, UK; [d]European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC), Brussels, Belgium; [e]Technical University of Munich, Munich, Germany; [f]Dow Europe GmbH, Horgen, Switzerland; [g]Scientific Consultancy – Animal Welfare, Neubiberg, Germany

ABSTRACT

The European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC) organized a workshop "*Hazard Identification, Classification and Risk Assessment of Carcinogens: Too Much or Too Little?*" to explore the scientific limitations of the current binary carcinogenicity classification scheme that classifies substances as either carcinogenic or not. Classification is often based upon the rodent 2-year bioassay, which has scientific limitations and is not necessary to predict whether substances are likely human carcinogens. By contrast, tiered testing strategies founded on new approach methodologies (NAMs) followed by subchronic toxicity testing, as necessary, are useful to determine if a substance is likely carcinogenic, by which mode-of-action effects would occur and, for non-genotoxic carcinogens, the dose levels below which the key events leading to carcinogenicity are not affected. Importantly, the objective is not for NAMs to mimic high-dose effects recorded *in vivo*, as these are not relevant to human risk assessment. Carcinogenicity testing at the "maximum tolerated dose" does not reflect human exposure conditions, but causes major disturbances of homeostasis, which are very unlikely to occur at relevant human exposure levels. The evaluation of findings should consider biological relevance and not just statistical significance. Using this approach, safe exposures to non-genotoxic substances can be established.

ARTICLE HISTORY
Received 14 January 2020
Revised 3 February 2020
Accepted 5 February 2020

KEYWORDS
Non-genotoxic carcinogens; mode of action (MoA); classification and labeling (C&L); new approach methodologies (NAMs); tiered testing strategy; weight of evidence (WoE)

Table of contents

Background

Human health risk assessment of man-made and naturally occurring substances (for use as e.g. food additives, industrial chemicals, pesticides, cosmetics, or pharmaceuticals) serves to protect humans from unwanted effects caused by exposure to these substances. The internationally agreed-upon risk assessment paradigm consists of four steps: (1) hazard identification, (2) hazard characterization including dose-response assessment ((1) and (2) together = hazard assessment), (3) exposure assessment, and (4) risk characterization; followed by the implementation of risk management measures, as necessary (WHO IPCS 2004, 2010). Although this paradigm has been updated to put greater emphasis on problem

CONTACT Susan P. Felter Felter.sp@pg.com Global Product Stewardship; Procter & Gamble; 8700 Mason Montgomery Road, Mason, OH 45040, USA

formulation and understanding human exposure first (NRC 2009), it is still the *de facto* approach for risk assessment. Information gathered during the first step of risk assessment (hazard identification) may be used for the classification and labeling (C&L) of substances, e.g. in accordance with the *Globally Harmonized System (*GHS) *of Classification and Labeling of Chemicals* (United Nations 2017). Legislation, regulation, and guidance for the hazard and risk assessment of chemicals as well as for their C&L have been implemented in all major jurisdictions world-wide, e.g. within the European Union (EU) in *Regulation (EC) 1907/2006 concerning the registration, evaluation, authorization, and restriction of chemicals* (REACH; EP and Council 2006) and *Regulation (EC) 1272/2008 on classification, labeling, and packaging (CLP) of substances and mixtures* (EP and Council 2008), or based on an acceptable risk range, as in the USA, in the *Frank R. Lautenberg Chemical Safety Act of the 21st Century* (US Government 2016).

One of the human health endpoints included in the GHS is carcinogenicity (Box 1), for which it is currently classified in a binary manner; that is as either present (carcinogenic) or absent (non-carcinogenic) (Doe et al. 2019). Although there are subcategories ("known, probable, possible" carcinogens), this distinction is often lost in the regulatory translation and in public perception. By contrast, the classification schemes for a number of other human health endpoints (e.g. acute toxicity, skin and eye irritation, and corrosivity) include estimates of potency to assign a substance to the respective category (Hennes et al. 2014).

It is widely accepted that chemical carcinogenicity is a complex, multi-stage process that may involve different mechanisms or modes of action (MoAs; e.g. as reviewed by Cohen and Ellwein 1991; Boobis et al. 2006, 2009; and Boobis 2010). MoA is a biologically plausible sequence of substance-dependent key events, starting with exposure and proceeding through the interaction of the substance or its metabolites with a cell, through functional and anatomical changes leading to an observed effect (Sonich-Mullin et al. 2001; Boobis et al. 2009; Fenner-Crisp and Dellarco 2016).

Box 1. Definitions for carcinogenicity, mutagenicity, genotoxicity.

Carcinogenicity refers to the induction of cancer or an increase in the incidence of cancer occurring after exposure to a substance or mixture. Substances and mixtures which have induced benign and malignant tumours in well performed experimental studies on animals are considered also to be presumed or suspected human carcinogens unless there is strong evidence that the mechanism of tumour formation is not relevant for humans (GHS; Section 3.6.1; (United Nations 2017)).

The term *mutation* applies both to heritable genetic changes that may be manifested at the phenotypic level and to the underlying DNA modifications when known (including, for example, specific base pair changes and chromosomal translocations) (GHS; Section 3.5.1.4 (United Nations 2017)).

The more general terms genotoxic and *genotoxicity* apply to agents or processes which alter the structure, information content, or segregation of DNA, including those which cause DNA damage by interfering with normal replication processes, or which in a non-physiological manner (temporarily) alter its replication. Genotoxicity test results are usually taken as indicators for mutagenic effects (GHS; Section 3.5.1.5 (United Nations 2017)).

Broadly, there is a distinction between DNA-reactive genotoxic and non-genotoxic MoAs of carcinogenicity. Genotoxic MoAs involve direct interactions with the DNA that lead to an increased rate of DNA damage and mutations per cell division. By contrast non-genotoxic MoAs of carcinogenicity include increased cell proliferation (e.g. due to the activity of mitogens or as a response to necrosis with regenerative proliferation), altered DNA methylation, cellular growth by dysregulated epigenome, and hormonal effects (Greenfield et al. 1984; Cohen and Ellwein 1991; Goodman 1998; Goodman and Watson 2002; Cohen 2010; Timp and Feinberg 2013; Cohen et al. 2019; Kobets et al. 2019; Kobets and Williams 2019).

While a broad number of *in vitro* and *in vivo* test methods are available to assess mutagenicity and genotoxicity (Box 1), the assessment of carcinogenicity is most commonly based on the outcome of *in vivo* long-term studies. The standard test method is the rodent 2-year bioassay that has been adopted as *Organization for Economic Co-operation and Development (OECD) Test Guideline (TG) 451 Carcinogenicity study* or the *OECD TG 453 Combined chronic toxicity – carcinogenicity study* (http://www.oecd.org/chemicalsafety/testing/oecdguidelinesforthetestingofchemicals.htm [accessed 2020 Feb]). For pharmaceuticals, the 6-month bioassay in transgenic mice may be acceptable *in lieu* of a 2-year rat and/or mouse study (ICH 1997).

Concerns have been expressed that the current binary hazard identification-based classification scheme for carcinogenicity is outdated and does not serve the goal of human health protection (Hennes et al. 2014; Boobis et al. 2016; Cohen et al. 2019; Doe et al. 2019; Wolf et al. 2019). This is especially true for non-genotoxic carcinogens, which elicit effects *via* non-linear, threshold-based mechanisms, such as cell proliferation. This means that they do not elicit effects if exposure is below the respective threshold. Therefore, a binary classification scheme is particularly inadequate for such substances as it does not include, e.g. dose-response or MoA considerations (Bolt et al. 2004; Boobis 2010; Doe et al. 2019; Wolf et al. 2019). Furthermore, scientific limitations of the standard rodent 2-year bioassay may result in impaired hazard identification and hence erroneous classification of carcinogenicity. For example, the rodent 2-year bioassay includes very high doses (e.g. the maximum tolerated dose (MTD)) that rarely, if ever, reflect human exposure scenarios. It does not provide mechanistic information nor does it consider toxicokinetics, both of which are often different at high *versus* lower doses (Cohen 2017; Felter et al. 2018; Doe et al. 2019; Sauve-Ciencewicki et al. 2019). This is explicitly recognized by the US National Toxicology Program (NTP 2016), which states that its conclusions on rodent carcinogenicity are relevant only to the conditions of the bioassay under which the respective substance was tested.

In contrast to these concerns for "over-classification" of substances as carcinogens, other scientists have asked whether current practices are adequately protective. Woutersen et al. (2019) suggested that the current tonnage-based standard information requirements implemented under REACH (EP and Council 2006) do not provide sufficient information to determine whether a substance is a carcinogen based on animal evidence (i.e. Category 1B; Table 1), and

Table 1. Carcinogenicity classification schemes implemented by the IARC (2019a), the GHS (United Nations 2017), under the former EU Dangerous Substances Directive (Council 1967), by the US EPA (2005a) and by the German MAK Commission (DFG 2019).

International Agency for Research on Cancer	Globally Harmonized System of C&L of Chemicals*	EU Dangerous Substances Directive	US EPA weight-of-evidence descriptors	MAK Commission†
Group 1 Human evidence	**Category 1A** Human evidence	**Category 1** Human evidence	Carcinogenic to humans	**Category 1** Human evidence
Group 2A Limited human evidence, strong animal and mechanistic evidence	**Category 1B** Animal evidence for carcinogenicity in humans	**Category 2** Sufficient evidence for human carcinogenesis from animal data	Likely to be carcinogenic to humans	**Category 2** Animal evidence for carcinogenicity in humans
Group 2B Limited human evidence, less than sufficient animal evidence, or strong mechanistic data	**Category 2** Suspected human carcinogen	**Category 3** Some evidence from animal data, but insufficient for Category 2	Suggestive evidence of carcinogenic potential	**Category 3A** Proposed Category 4 or 5, but no MAK available
Group 3 Inadequate human and animal data for classification	–	–	Inadequate information to assess carcinogenic potential	**Category 3B** Inadequate data for classification
Group 4 (deleted 2019) No indication for carcinogenicity	–	–	Not likely to be carcinogenic to humans	**Category 4** No (or only secondary) genotoxicity; health-based MAK available **Category 5** Genotoxic carcinogen; risk-based MAK available

C&L: classification and labeling; MAK: maximum work place level (*maximale Arbeitsplatzkonzentration*); MAK commission: permanent senate commission for the investigation of health hazards of chemical compounds in the work area (German Research Foundation).

*In the EU, the Globally Harmonized System of C&L of Chemicals has been implemented in Regulation (EC) No 1272/2008 on classification, labeling and packaging of substances and mixtures (EP and Council 2008), which is applied e.g. by the Committee for Risk Assessment (RAC) of the European Chemicals Agency (ECHA).

†The scheme of the MAK Commission (that is widely concordant with the scheme used by the former EU Scientific Committee on Occupational Exposure Limits (SCOEL)) is not directly comparable to the other schemes since it considers (genotoxic *versus* non-genotoxic) MoAs in the assignment of categories.

thus that "indications of very severe hazards of substances are missed and health risks could occur" (Woutersen et al. 2019). The legal grounds underlying this view are that carcinogenicity studies are only standard information requirements for substances manufactured or imported in quantities of 1000 tonnes/year or more (Annex X of the REACH Regulation). Braakhuis et al. (2018) then posed the question of whether current risk assessment approaches for non-genotoxic carcinogens are adequately protected and suggested that a broader discussion within the scientific community is needed.

Against this background, the European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC) *organized the Worksop Hazard Identification, Classification and Risk Assessment of Carcinogens: Too Much or Too Little?* that took place on 8 September 2019 in Helsinki, Finland, and was organized as a Satellite Workshop to the EUROTOX 2019 Conference. The workshop was open to all registered participants and encompassed representatives globally from academia, industry, governments and authorities, non-governmental organizations, etc.

Susan P. Felter (*Procter & Gamble, USA*) opened the workshop welcoming the approx. 150 participants and outlined the objectives of the workshop. Overall, the workshop focused on non-genotoxic carcinogens, and it served to explore two themes.

- *Theme 1*: The basis for hazard identification and classification of non-genotoxic carcinogens: Are substances that are not carcinogenic to humans being classified as carcinogens? *Vice versa*, are substances that are carcinogenic to humans being missed due to insufficiencies of the available data or methods?
- *Theme 2*: Current methodologies for the quantitative risk assessment of non-genotoxic carcinogens: Are current methods overly conservative or insufficient to provide adequate protection?

The proceedings of the workshop, presented below, are structured accordingly. *Theme 1* encompasses presentations on "*Classification of non-genotoxic carcinogens: how and why,*" "*Classification of carcinogens: what could go wrong?*" and "*β-Myrcene: Implications for classification of this non-genotoxic carcinogen,*" followed by a Panel Discussion to further explore these topics. *Theme 2* encompasses presentations entitled "*Inconsistent evaluation and interpretation of chemical cancer risk inhibits innovation without enhancing safety,*" "*Establishing an adequate margin of protection for non-genotoxic carcinogens,*" and "*A path forward for carcinogenicity evaluation without the two-year bioassay.*" The workshop was complemented by Stakeholder Perspectives, an overarching Panel Discussion, and Concluding Thoughts. While most of the examples provided during the workshop included commercially relevant substances, the principles set forth are equally applicable to contaminants and other substances not subject to approval.

THEME 1: Cancer hazard identification for non-genotoxic substances: too much or too little?

Classification of non-genotoxic carcinogens: How and why?

Helmut Greim (*Technical University of Munich, Germany*) described the carcinogenicity classification schemes and

approaches adopted by the following authoritative bodies:

1. The International Agency for Research on Cancer (IARC 2019a);
2. The United Nations in the GHS (United Nations 2017);
3. The EU in the former Dangerous Substances Directive (Council 1967) that has been superseded by the EU CLP Regulation (EP and Council 2008) implementing the GHS in the EU;
4. The "MAK Commission," i.e. the German Research Foundation's *Permanent Senate Commission for the Investigation of Health Hazards of Chemical Compounds in the Work Area* (DFG 2019);
5. The former EU *Scientific Committee on Occupational Exposure Limits* (SCOEL) whose work was taken over by the Committee for Risk Assessment (RAC) of the European Chemicals Agency (ECHA) in 2018; https://ec.europa.eu/social/main.jsp?catId=148&intPageId=684&langId=en [accessed 2020 Feb]
6. The United States Environmental Protection Agency (US EPA 2005a) in its *Guidelines for carcinogen risk assessment*.

The IARC (2019a) classification scheme includes the following carcinogen classification groups:

- Group 1: Classification based on human evidence
- Group 2: Classification predominantly based on animal and mechanistic evidence
  - Group 2A: Probably carcinogenic to humans; assigned if two of the following three parameters are present:
    I. limited evidence of carcinogenicity in humans
    II. sufficient evidence of carcinogenicity in experimental animals
    III. strong evidence that the agent exhibits key characteristics of carcinogens
  - Group 2B: Possibly carcinogenic to humans; assigned if one of the three parameters listed above are present
- Group 3: The available human and animal data are inadequate for classification
- Group 4: No indication for carcinogenicity. (Group 4 was deleted from the IARC (2019a) classification scheme in January 2019).

The GHS (United Nations 2017) and the former EU Dangerous Substances Directive (Council 1967), that is mentioned for historical completeness, have simplified and harmonized the IARC classification scheme (Table 1).

The carcinogenicity classification schemes of the IARC and the GHS are exclusively hazard-based. They do not consider either dose-response or exposure and hence do not inform on risk. By comparison, the MAK Commission (and similarly the former SCOEL) does consider exposure and MoAs when classifying carcinogens (DFG 2019).

Accordingly, the scheme of the MAK Commission (DFG 2019) includes a category for carcinogens exhibiting a threshold-based MoA, and it considers the potency of such substances. Specifically, non-genotoxic carcinogens are assigned as MAK Category 4, and a health-based maximum workplace concentration (MAK) value is derived for these substances (Table 1). If exposures are kept below the MAK value, no significant human health effects are to be expected. The MAK Commission has classified a number of carcinogens as Category 4 (including chloroform, 1,4-dioxane, formaldehyde, hexachlorobenzene, polyvinyl chloride, lindane, polyacrylates, titanium dioxide, and vinyl acetate), and it has established MAK values for these substances (DFG 2019).

The distinction between linear (non-threshold) MoAs and non-linear (threshold-based) MoAs for carcinogenicity is also made in guidance from ECHA (2012), the European Food Safety Authority Scientific Committee (EFSA SC 2005) and in the *US EPA Guidelines for carcinogen risk assessment* (US EPA 2005a).

The US EPA (2005a) Guidelines go beyond a mere assignment of substances to categories to allow for the complexity of the endpoint carcinogenicity. They require a weight of evidence (WoE) narrative to explain the substance's potential carcinogenicity in humans and the conditions that characterize its expression. To provide additional clarity and consistency, a set of WoE descriptors is suggested to accompany the WoE narratives (US EPA 2005a; Table 1).

Since effects elicited by non-genotoxic carcinogens are assumed to be threshold-based, no-observed adverse effect levels (NOAELs) or no-observed adverse effect concentrations (NOAECs) can be identified from the outcomes of the animal studies. From the available database, the most sensitive NOAEL or NOAEC is then used as a point-of-departure to derive health-based exposure limits, empirically selecting and applying assessment factors (AFs; also called uncertainty factors) to account for uncertainties of the hazard data. Depending on the jurisdiction and/or legislation, health-based exposure limits may include reference doses/reference concentrations, acceptable daily intake levels (ADIs), derived no-effect levels, MAK values, 15-min limit values, etc.

In accordance with the *ECHA Guidance on the application of the CLP criteria* (ECHA 2017), a conclusion of a carcinogen as having a non-linear, threshold-based MoA may lead to the downgrading of a Category 1 classification to Category 2 (ECHA 2017) and the derivation of health-based exposure limit values for this substance. This "downgrading" has regulatory consequences because for Category 2 carcinogens, in contrast to those in Category 1 A or B, consumer exposure is not restricted *a priori*.

Explanations for why specific AFs were selected for the derivation of the health-based exposure limits improve transparency of the risk management measures. For example, for 1,4-dichlorobenzene, the ECHA RAC (2013) used a 2-year inhalation toxicity study in F344 rats and BDF1 mice as the key study for its evaluation. Hepatoblastomas and histiocytic sarcomas were recorded in the mouse studies at the highest concentration only (300 ppm), but not in the rat study. There was no evidence for genotoxicity. The NOAEC was identified as 75 ppm (the second-highest concentration in the mouse study). The ECHA RAC (2013) concluded that the overall WoE points to a "low potency, non-genotoxic carcinogen which exerts its tumourigenic response *via* mitogenic MoA in mice

only" and classified 1,4-dichlorobenzene as a Category 2 carcinogen. From the NOAEC of 75 ppm, the ECHA RAC set very low derived no-effect levels of 0.6 ppm for workers and 0.11 ppm for consumers. An AF of 3 was selected to account for the "steep dose-response observed" and the uncommonness of the tumors; AFs of 5 and 10 were selected to account for intra-species differences for workers and consumers, respectively, and an AF of 2.5 to account for inter-species differences (ECHA RAC 2013).

In addition, three recent examples show how the ECHA RAC has applied the provisions from the ECHA (2017) guidance to derive occupational exposure limits (OELs) for non-genotoxic carcinogens:

For *acrylonitrile*, ECHA RAC (2018a) derived an 8-h average OEL of 1 mg/m^3 and a 15 min limit value of 4 mg/m^3, and it drew the following key conclusion:

> The critical endpoint in establishing the relevance of an OEL for acrylonitrile is its carcinogenicity. From the total WoE from both animal and human data a MoA-based threshold can be assumed for the carcinogenic effects of acrylonitrile. At acrylonitrile exposures below the resulting proposal for a limit value, no significant residual cancer risk is expected for workers. (ECHA RAC 2018a)

For *benzene*, ECHA RAC (2018b) derived an 8-h average OEL of 0.16 mg/m^3, and it drew the following key conclusions:

> A MoA-based threshold for chromosomal damage (aneugenicity and clastogenicity) in workers can, in the view of RAC, be used to establish an OEL for carcinogenicity. The limit so derived, will avoid exposures that induce chromosomal damage in workers, is considered to have no significant residual cancer risk and will also avoid other adverse effects. The leading genotoxic effects, aneugenicity and clastogenicity, are considered to be of secondary nature, i.e. acting indirectly and to follow a non-linear threshold-mechanism. Various studies show induction of adverse chromosomal damage in benzene-exposed workers from different working environments. Primary DNA reactivity of benzene and/or its metabolites seems of little importance. (ECHA RAC 2018b)

The ECHA RAC further explained why it considers primary DNA reactivity to be of little importance for benzene:

> Certainly, primary DNA reactivity of benzene or its reactive metabolites … cannot be fully ruled out, thus it is difficult to definitively exclude some remaining risk at lower exposure levels. There is however, a remarkable consistency of published cancer risk estimates based on the higher exposure levels previously encountered in occupational settings, i.e. above 1 ppm. Considering, however, that multiple thresholded MoAs likely contribute to benzene leukaemia development and in view of the overall experimental and epidemiological evidence available supporting a genotoxic-threshold for benzene, the remaining uncertainties are considered to be very low. Given this evidence, estimated excess cancer risks as derived by linear extrapolation can be seen as overly conservative. (ECHA RAC 2018b)

For *nickel and its compounds*, ECHA RAC (2018c) derived an 8-h average OEL of 0.005 mg/m^3 for respirable dust and of 0.03 mg/m^3 for inhalable dust. It concluded that the main hazard of nickel compounds is their carcinogenicity in the respiratory tract and that they are not directly mutagenic, but rather induce genotoxic effects by indirect mechanisms. Further, the ECHA RAC stated:

> The available information on the mechanisms of genotoxicity and cancer support a MoA-based threshold for carcinogenic effects. The proposed OEL therefore relies on a MoA-based threshold for the carcinogenicity of nickel compounds. In addition to the mechanistic data reviewed by RAC, data on the lack of genotoxicity in animals at inhalation doses below the levels causing inflammation and cytotoxicity support this conclusion. At exposures below the proposed limit value, no significant residual cancer risk is expected for workers. (ECHA RAC 2018c)

In conclusion, in the evaluation of carcinogens, a distinction should be made between those for which the MoA is by direct and indirect genotoxicity, and the hazard classification of non-genotoxic carcinogens should be based on an understanding of the threshold mechanism and of its dose-response. Since non-genotoxic carcinogens elicit effects by a non-linear MoA, linear extrapolations to doses below the threshold are not justifiable to estimate human cancer risk.

Classification of carcinogens: what could go wrong?

Rita S. Schoeny (*Rita Schoeny LLC, USA*) presented experiences gained at the US EPA in developing and applying a scheme for the hazard identification, classification, and risk assessment of carcinogens.

Generally, risk assessment is performed to provide support for decisions to protect human health and the environment. It summarizes the state-of-the-art science making use of the available data. Risk assessment is a complex process that sometimes also yields controversial conclusions. Depending on the scope of the risk assessment, specific questions need to be answered, and sufficient data and knowledge are needed to make an informed, rational choice. As explained in the US National Research Council "Silver Book" *Science and Decisions. Advancing Risk Assessment* (NRC 2009): "Risk assessment should be viewed as a method for evaluating the relative merits of various options for managing risk." Hence, the risk assessment must be fit-for-purpose, providing conclusions that can best inform risk management choices. Since the available database is never exhaustive, risk assessors need to deal with the inevitable uncertainties in a rational, scientifically supportable manner.

A historical milestone in the risk assessment of substances was the publication of the "Red Book" *Risk assessment in the Federal Government: Managing the process* in 1983 (NRC 1983). Therein, a four-step process for risk assessment was described: hazard identification, dose-response assessment, exposure assessment, risk characterization, leading to risk management, and risk communication. As experience has been gained with the steps, the necessity of considering them in an iterative and linked fashion has become apparent. Moreover, more recent guidance has stressed the need for dialog among risk assessors and risk managers to ensure that the risk assessment is fit-for-purpose (see, for example, US EPA 2014). The "Red Book" advised the US Federal Agencies to publish guidelines for risk assessment. Following this advice, the US EPA (1986) published, amongst other guidance, the first version of the *Guidelines for carcinogen risk assessment*.

The US EPA (1986) Guidelines encompassed strict number-based rules of evidence to conclude if there was either limited or sufficient evidence to classify a substance as a carcinogen. Overall, it was considered that there was no safe exposure to a carcinogen, i.e. that all carcinogens exhibited a linear low dose-response relationship (US EPA 1986). However, in the early 2000's, this view was challenged as increased knowledge of the mechanisms of carcinogenicity became available. Accordingly, the *US EPA Guidelines for carcinogen risk assessment* were revised in 2005 to include WoE narratives to integrate all available evidence in the overall conclusion and to address differential risks in children, amongst other issues. The identification of MoAs forms an essential part of the decisions, and linear and non-linear extrapolations are considered, as applicable (US EPA 2005a; see also above; summary of H. Greim's presentation).

A key element of the risk assessment procedure is the determination of the sufficiency of the available evidence (Box 2). The example of the insecticide lindane (gamma-hexachlorocyclohexane; CAS No. 58-89-9) shows how challenging it can be to establish the sufficiency of the evidence and to interpret the relevance of the findings with respect to the classification for carcinogenicity. In the earlier version of the *US EPA Guidelines for Cancer Risk Assessment* (US EPA 1986), rather circumscribed criteria were provided to define "sufficient" and "limited" evidence for carcinogenicity. This was generally based on the availability of one or more positive human (epidemiological) or animal studies. For example, "sufficient animal evidence" was defined thusly:

> that there is an increased incidence of malignant tumors or combined malignant and benign tumors: (a) in multiple [emphasis by the authors of the present report] species or strains; or (b) in multiple experiments (e.g. with different routes of administration or using different dose levels); or (c) to an unusual degree in a single experiment with regard to high incidence, unusual site or type of tumor, or early age at onset. (US EPA 1986)

A rodent bioassay addressing the potential carcinogenicity of lindane was published in 1977 (NCI 1977). This bioassay included female and male F344 rats and female and male B6C3F1 mice, and substance administration was *via* the feed, following the protocol of the NTP that was valid at that time using the MTD as high dose (see Boobis et al. 2016; Doe et al. 2019) for a discussion of the scientific irrelevance of the MTD). In brief, the bioassay showed no increase in the incidence of tumors in the rats or female mice and a significant increase in liver tumors in the low-dose male mice only (i.e. at 80 ppm, but not at 160 ppm). This finding was assessed as biologically not relevant as there was no dose-response relationship and the results were deemed negative for all four subsets of the study (NCI 1977). Other studies in mice have also been considered by risk assessors (see for example US EPA (1988) for summaries). Thorpe and Walker (1973) recorded liver tumors in male and female CF1 mice orally exposed to 440 ppm (52 mg/kg body weight (bw)/d) for 110 weeks; however, only 3% of the female mice and 17% of the males from these test groups survived at the end of the study. Goto et al. (1972) reported liver tumors in 5 of 10 IRC-JCL male mice orally exposed to 600 mg/kg bw/d, and Hanada et al. (1973) recorded that 1 of 3 surviving female mice and 3 of 4 surviving male mice developed liver tumors after 37- to 38-week oral exposure to 600 ppm lindane (US EPA 1988).

Box 2. Definitions for sufficient and limited evidence of carcinogenicity in the GHS (United Nations 2017).

Sufficient evidence of carcinogenicity implies that "a causal relationship has been established between the agent or mixture and an increased incidence of malignant neoplasms or of an appropriate combination of benign and malignant neoplasms in (i) two or more species of animals or (ii) in two or more independent studies in one species carried out at different times or in different laboratories or under different protocols. Exceptionally, a single study in one species might be considered to provide sufficient evidence of carcinogenicity when malignant neoplasms occur to an unusual degree with regard to incidence, site, type of tumour or age at onset."

Limited evidence of carcinogenicity implies that "a positive association has been observed between exposure to the agent, mixture or exposure circumstance and cancer for which a causal interpretation is considered by the working group to be credible, but chance, bias or confounding could not be ruled out with reasonable confidence."

The US EPA has held some differing opinions across its regulatory and research programs (and at different points in time) on the sufficiency of the evidence from the mouse studies for lindane. Points of discussion included whether the studies performed at the MTD are well-conducted and whether a standard National Cancer Institute (NCI)/NTP bioassay ought to be considered as one single study, or as two, or even four separate studies. Furthermore, there have been varying opinions as to the weight that should be given male mouse liver tumors, in general. The 1986 US EPA Cancer Guidelines provided some guidance on "discounting" these neoplasms particularly in the absence of a dose response, and the relevance of these tumors to human health risk assessment has been a subject of continuing debate.

The US EPA Office of Water regulated lindane in drinking water on 30 January 1991 (US EPA 1991). They considered the evidence to be "limited" and classified the carcinogenicity to be Category C (*possible human carcinogen on account of limited evidence of carcinogenicity in animals*). At this point in time, the Office of Water was required to treat "possible carcinogens" as having a threshold; dose-response assessment was done by applying a 10-fold safety factor to a calculated reference dose for a non-cancer effect. In the same general time frame, the US EPA Office of Research and Development Integrated Risk Information System (IRIS) classified lindane as Category B2 (*probable human carcinogen with sufficient evidence from animal studies*)/Category C (*possible human carcinogen*) following the 1986 US EPA classification scheme. IRIS considered the same suite of studies as the Office of Water. An item of note is that the proposed split classification leaves open the procedure for dose-response extrapolation (linear or non-linear). IRIS did not resolve the issue of categorization of lindane, and the IRIS Chemical Assessment Summary for lindane does not include carcinogenicity assessment for lifetime exposure (US EPA 1987).

There is no discussion as to whether this omission would have caused health problems.

In 1997, the US EPA Office of Pesticide Programs reviewed some additional available data (none of it positive for a carcinogenic effect) and classified lindane as Category C only (US EPA 1997). Thus it will remain, as the last registered uses of lindane were canceled in the US in 2006. Similarly, in 1979 IARC, which does not consider dose-response relationships in its monographs, classified lindane as Category 2B, but in 2015, IARC revisited this classification and assigned it as Category 1 carcinogen (IARC 2018; see Table 1 for IARC categories).

In conclusion, the example of lindane shows that a rigid categorization scheme, that does not consider dose response, conditions of exposure or human relevance of tumor(s) observed in an animal bioassay, leads to a confusing array of judgments that are not useful as a basis for risk management.

With regard to the scenario of "limited evidence for carcinogenicity," the US EPA (2005a) Guidelines indicate that, if critical information is lacking or uncertainty is too high for other reasons, a default option as presented in Appendix A of the Guidelines should be invoked. Mostly, these default options are inferences that help use the recorded data under empirical conditions in order to estimate events and outcomes under environmental conditions (US EPA 2005a). To reduce inconsistencies in the interpretation of findings, a structured approach to risk assessment is required to ensure that all available data are collected and evaluated and that uncertainties are identified before invoking defaults. By contrast, if the evidence is considered sufficient, risk assessment can be conducted. Hazard characterization, i.e. dose-response assessment taking into account the hazard identification, is an important part of the risk assessment. Hazard characterization also considers the specific conditions under which an effect is produced, e.g. route-specific effects, the relevance of effects to humans, and the occurrence of effects at high doses only (US EPA 2014).

As compared to a simple categorization of substances, WoE narratives provide the opportunity for a detailed explanation of the substance's potential carcinogenicity and the considerations underlying the WoE conclusion. The example of the risk assessment of chloroform (US EPA 2001) illustrates the advantages of preparing a WoE narrative. As summarized in US EPA (2001), chloroform had been classified as Group B2 "probable human carcinogen" following the 1986 classification scheme based on "sufficient evidence" of carcinogenicity in animals. In the 1990s, it was classified as "likely to be carcinogenic to humans by all routes of exposure." However, this classification was based on data from high-exposure conditions that elicited cytotoxicity and regenerative hyperplasia in susceptible tissues. By contrast, chloroform is currently classified as "not likely to be carcinogenic to humans by any route of exposure" under exposure conditions that do not cause cytotoxicity and cell regeneration. The evaluation was done in the context of regulation of chloroform and other trihalomethanes in US drinking water. This chloroform effects (hazard) characterization was done according to the revised US EPA Cancer Guidelines (US EPA 2005a) that specifies a narrative classification rather than a letter/number category. The conclusion is based on these considerations: (1) oral and inhalation toxicity data indicating that sustained or repeated cytotoxicity with secondary regenerative hyperplasia precedes hepatic and renal neoplasia and is likely required for tumor development; (2) the unavailability of epidemiological data specific to chloroform and, epidemiological data related to drinking water exposures that cannot be attributed to chloroform alone; (3) an overall evaluation of the available genotoxicity data on chloroform that support the conclusion that chloroform is not mutagenic and that genotoxicity is not likely to be an MoA of chloroform carcinogenicity; (4) consideration that chloroform oral exposure is a high dose effect not relevant to drinking water exposure attributable to water disinfection (US EPA 2001).

Today, the experiences gained in performing carcinogenicity risk assessment should be used to reassess if "carcinogen" is required as a label for substances at all. Knowledge of MoAs of carcinogenicity should be applied when evaluating the available evidence within a WoE approach, and this should include the human relevance of hazard at exposure levels that are relevant to humans (further considering highly susceptible populations).

Finally, it is challenging to communicate the outcome of risk assessment to the public. For example, the US EPA Office of Pesticide Programs regularly publishes a list of chemicals that have been evaluated for carcinogenicity (US EPA 2018). This is not a "list of carcinogens." In the introduction to the list, it is emphasized that the list is not intended to be used independently of the full risk assessment for the chemical and that the simple fact of being listed does not imply that the chemical poses a significant carcinogenic risk to the public (US EPA 2018). Nevertheless, sometimes a specific substance on this list is inappropriately flagged as posing health concerns in the media. This may cause unnecessary concern in the public and possibly ultimately a ban of a useful product, while diverting the attention and resources from more urgent problems. This issue is explored in further detail in the summary of S. P. Felter's presentation below.

β-Myrcene: implications for classification of this non-genotoxic carcinogen

S. P. Felter presented a case study highlighting the regulatory and legal consequences resulting from the IARC classification of the non-genotoxic flavor substance β-myrcene as an IARC Category 2B carcinogen.

β-Myrcene is a naturally occurring monoterpene found in more than 200 plant species. It is also widely used as a flavoring agent, for which it is produced industrially by the pyrolysis of β-pinene, one of key components of turpentine. In 1965, the Flavor Extract Manufacturers Association (FEMA 1965) Expert Panel granted β-myrcene the status as of Generally Recognized as Safe, and the US Food and Drug Administration (US FDA) approved β-myrcene as a food additive. In 1974, the (Europe-wide) Council of Europe ad hoc Working Party on Natural and Artificial Flavoring Substances approved β-myrcene for use as an artificial flavor (Council of

Europe 1974). In the EU, β-myrcene is included in the list of flavoring substances established under Regulation No 872/2012 (Commission 2012). Robust safety evaluations conducted by the Joint FAO/WHO Expert Committee on Food Additives (JECFA 2015) and EFSA (2015) confirmed there were no safety concerns. Similarly, on 21 June 2018, the US FDA (2018) published an updated review of the safety of β-myrcene as a flavoring substance in food, including its potential carcinogenicity, and concluded that there were no safety concerns.

Nonetheless, in October 2018, the US FDA amended the food additive regulations to no longer authorize the use of β-myrcene (and six other flavoring agents) in food. This action came in response to a 2016 citizen petition and subsequent lawsuit from consumer and environmental groups that referred to positive results in rodent carcinogenicity studies for these flavoring agents. Food manufacturers have until October 2020 to find suitable alternatives for the seven substances. The legal foundation for the 2016 citizen petition and the US FDA's action is the so-called Delaney Clause that was enacted under the US Food, Drug, and Cosmetic Act in 1958. Thereby, the US FDA cannot allow the legal use of any food additive found to be carcinogenic in humans or animals at any dose (thus, there is no option under this legal mandate to consider the human health relevance of findings from animal studies). However, the 2018 US FDA decision does not affect the legal status of foods that contain the natural counterparts of the flavoring substances; see also https://www.fda.gov/food/cfsan-constituent-updates/fda-removes-7-synthetic-flavoring-substances-food-additives-list [accessed 2020 Feb].

The key 2-year bioassay investigating β-myrcene (NTP 2010) was conducted in F344/N rats and B6C3F1 mice, following the standard NTP cancer bioassay protocol. Dose levels were 0, 250, 500, and 1000 mg/kg bw/d (gavage administration). In the male rats and mice, there was "clear evidence" of tumor formation in the kidneys and liver, respectively. In the female rats and mice, this evidence was equivocal. The clear evidence in the male rats and mice relates to a dose-dependent increase in tumor formation at 250 and 500 mg/kg bw/d. The highest dose (1000 mg/kg bw/d) exceeded the MTD, since all of the male rats and most of the male and female mice from this test group died before the end of the study. All available genotoxicity assessments were negative (summarized in US FDA 2018).

Based upon the findings from the NTP (2010) rodent bioassay, IARC (2019b) has classified β-myrcene in Group 2B "possible human carcinogen" based on "sufficient evidence" in animals. By contrast, β-myrcene is not included in the NTP's Report on Carcinogens (NTP 2016) that lists agents, substances, mixtures, and exposure circumstances that are known or reasonably anticipated to be carcinogenic in humans.

When evaluating the human health relevance of the findings from the rodent bioassay, it should be considered that β-myrcene is a naturally occurring monoterpene present in over 200 plants and fruits, including basil, carrots, citruses, hops, lemongrass, mangoes, pomegranates, rosemary, and thyme. The structurally identical artificial flavoring chemical is used in beverages, ice cream, candy, and baked goods. Exposure to naturally occurring β-myrcene by eating plants is 50-fold higher than that from its use as an artificial flavor (Adams et al. 2011).

Further, the overall range of dietary exposure of humans to β-myrcene (approx. 1.2–4.8 μg/kg bw/d) is about five orders of magnitude lower than the lowest dose in the rodent bioassay (250 mg/kg bw/d; NTP 2010). This estimation is based upon the following data:

- US FDA (2018): Estimated daily intake of β-myrcene as a synthetic flavoring substance: 1.23 μg/kg/d for a 60-kg person.
- EFSA (2015): Maximized survey derived daily intake in the EU: 4.8 μg/kg/d for 60-kg adult; in the USA: 2.6 μg/kg/d for 60-kg adult.
- FEMA (Adams et al. 2011): Daily per capita intake (eaters only): 3 μg/kg/d for 60-kg person.

β-Myrcene was also tested in an OECD TG 408-compliant 90-d toxicity study using male and female Sprague-Dawley rats (Bastaki et al. 2018). The oral NOAEL for rats of both sexes was the highest dose tested. Based on food consumption and test substance (in-)stability in the diet, Bastaki et al. (2018) calculated the NOAEL to be 115 and 136 mg/kg bw/d for males and females, respectively. EFSA (2015), applying a "worst-case" scenario of test substance instability, calculated the NOAEL from the 90-d study as being 44 and 53 mg/kg bw/d, for male and female rats, respectively. However, irrespective of the assumptions underlying the calculation of the NOAEL, the 90-d NOAELs for both male and female rats again exceed the average human intake by many orders of magnitude, providing a margin of exposure in the range of 10,000.

Overall, the available database supports the conclusion drawn by JECFA (2015), EFSA (2015), and US FDA (2018) that β-myrcene (which is not genotoxic, acutely toxic, or toxic to reproduction/development) does not pose a human health safety concern. The tumors recorded in male rats and mice following gavage administration at very high doses for 2 years do not raise any safety concerns (for cancer or any other adverse effect) for humans exposed to very low levels in food, whether naturally-occurring or synthetic. This is also reconfirmed on the above-mentioned US FDA website announcing the ban of the seven flavoring substances:

> Although we are amending our food additive regulations for these synthetic flavoring substances in accordance with the Delaney Clause, the FDA's rigorous scientific analysis has determined that they do not pose a risk to public health under the conditions of their intended use.

S. P. Felter highlighted that, beyond the US FDA ban, the 2018 events have induced an erosion of trust in regulatory agencies (and industry), which was exacerbated by non-governmental organizations that sued the US FDA. This does not only have economic implications for food manufacturers and consumers alike, but it also has implications for consumer perception of safety.

Importantly, the Delaney Clause, adopted in 1958, addresses only if a substance was shown to be carcinogenic in humans or animals. It does not consider exposure

potential or the human health relevance of effects observed in animals. Thereby, it creates an imbalance between the legal framework and the state of the art carcinogen risk assessment. The scientific limitations of the Delaney Clause have been recognized for many decades (e.g. Roberts 1982), but it is also recognized that it will be very difficult to change the law. In the meantime, it is critical that the scientific community considers implications of new study designs and approaches aimed at "identifying" non-genotoxic carcinogens, such that human relevance is a clear focus.

In conclusion, the classification of a substance as a carcinogen should only be undertaken following a WoE evaluation of all available evidence further considering the human health relevance of findings (Boobis et al. 2016; Doe et al. 2019). For substances classified as carcinogens, the legislation implemented in many jurisdictions vastly restricts their applicability and precludes certain uses or prescribes minimization of human exposure. Notwithstanding, the carcinogenicity classification schemes implemented by different bodies (see above; summary of H. Greim's presentation) have not been harmonized, and the legal provisions for classification are not necessarily founded on the state-of-the-science, as the β-myrcene case study shows. Finally, the focus on classification alone in the public perception impairs risk communication and makes it difficult to explain to the public what constitutes a real risk.

Scientific and political activities to advance the legal and regulatory provisions underlying current toxicology and risk assessment should address the following questions:

- Can we (and should we) mandate a change to testing protocols and/or study interpretations that require consideration of MoAs, human relevance, including consideration of exposure (and potential threshold)?
- Can we (and should we) affect a change to hazard identification and classification of non-genotoxic carcinogens such that it must be applied in the context of human exposure (and for substances already classified, that exposure must be considered when assessing whether a particular user should be allowed or banned)?
- How do we better inform consumers of what is real *versus* perceived risks?

Panel discussion

Moderator: **Alan R. Boobis** (Imperial College London, UK)

Panel members: **Warren Casey** (National Institute for Environmental Health Sciences (NIEHS)/NTP, USA); **Raffaella Corvi** (EU Commission, Joint Research Centre (JRC), Italy); **Wolfgang Dekant** (University of Würzburg, Germany); **S. P. Felter**; **H. Greim**; **R. S. Schoeny**

In further pursuing the questions raised by S. P. Felter (see above), the panelists deliberated about whether scenarios might exist where the classification of a chemical as carcinogenic truly serves human health protection. There was agreement that, historically, the establishment of carcinogenicity classification schemes was beneficial because it formed a starting point for the structured collation and evaluation of the available evidence. However, as knowledge of the mechanisms of carcinogenicity has evolved, the binary approach to carcinogenicity classification implemented more than three decades ago is no longer appropriate. Specifically, the classification of non-genotoxic carcinogens causes more confusion and misperceptions to the public than it has benefits.

Beyond the endpoint carcinogenicity, while the classification of substances may be pragmatic tools for risk management, it runs the risk of being overly simplistic. A classification scheme that exclusively focuses on hazard identification without considering the context (e.g. exposure potential) results in misleading perceptions. This is important because classification of a substance as a carcinogen can have regulatory consequences (e.g. leading to its ban or highly restrictive use) such that the other three steps of the risk assessment paradigm (i.e. dose-response relationship, exposure assessment and risk characterization; WHO IPCS 2004, 2010) are disregarded and an evaluation of the actual risk to human health is not carried out. Without an estimate of actual risk, effective risk management is not possible, as this requires consideration not only of risk but also of other legitimate factors, such as benefit and socioeconomic impact.

Hence, substance classification is not only a scientific issue, but also a political one. The general approach to substance classification, as such, has been agreed upon at the international level, e.g. in the GHS, and there are no initiatives to abandon it. To enhance the scientific relevance of substance classification for human health protection, it should always be applied together with risk assessment, thereby also considering dose-response relationships and exposure potential. Currently, however, there is usually a time gap between C&L and risk assessment. Hence, even if the risk assessment concludes that the substance does not pose a risk to humans, the respective substance might already have been taken off the market e.g. because downstream users do not want to use a substance labeled as carcinogenic.

Substance classification and risk assessment should always consider the human health relevance of effects observed in animals and take account of other legitimate factors such as socio-economic impact, direct and indirect befits (US EPA 1984), as is generally undertaken for pharmaceuticals and is also reflected in the socio-economic analysis mandated by the REACH Regulation (EP and Council 2006). In establishing the human health relevance of effects observed in animal studies, not only species differences in mechanisms, but also the dose-dependency of effects should be considered. If findings in the animal studies are only observed at very high doses that by far exceed any likely (or possible) human exposure, such effects should not be considered relevant for human health risk assessment.

Further, all risk assessments should consider information on the mechanisms of effects, integrating all available relevant and reliable evidence from e.g. *in silico*, *in vitro*, *in vivo,* and human studies as well as grouping and read-across in a WoE approach. The adoption of NAMs (e.g. *in silico* and *in vitro* tools, grouping into chemical categories and read-across to similar substances) to replace traditional animal toxicity test methods also needs to consider how well the

outcomes of such methodologies predict effects in humans at relevant concentrations. Data interpretation procedures are required to ensure that the results of the NAMs are predictive of effects in humans and applicable for the setting of human health exposure limits (OECD 2016). Further, establishing the fitness-for-purpose of NAMs serves to ensure their relevance and reliability. Both scientific and political efforts are required to meet these goals.

For the legal situation in the USA, there was agreement that the Delaney Clause needs to be amended to include assessments of human health relevance and the setting of human health exposure limits. In 1996, passing of the Food Quality Protection Act (US Government 1996) in effect modified the Delaney Clause as it applied to pesticides to accept a practical *de minimis* excess cancer risk of 1×10^{-6} for pesticides. Notwithstanding, it is unlikely that any elected representative will initiate substantial amendments to the Delaney Clause given the public perception of carcinogenicity.

Overall, communication is the key to advance the current carcinogenicity classification paradigm. Communication with the public should highlight the difference between hazard identification and classification on the one hand and risk assessment and risk management on the other hand. Notwithstanding, it is difficult to gain public attention, or attention in the media, to these issues, and engaging in constructive communication is challenging. Often, the intricacy and complexity of topics related to substance classification and risk assessment are almost impossible to communicate clearly to the general public, or even to academics with expertise in other areas. Further discussions are also required within the scientific community to agree on if and how carcinogenicity classification should be advanced, also considering that classification is used for different purposes by different bodies and organizations. In jurisdictions with evolving chemical legislation or where the chemical legislation is undergoing fundamental revision, the hazard and risk assessment of carcinogens should be based on the state-of-the-science. Regardless of the difficulties encountered in engaging in productive communication, a multi-stakeholder dialog is indispensable to overcome the current limitations of classification of non-genotoxic carcinogens.

THEME 2: Quantitative risk assessment of non-genotoxic carcinogens: are current methods adequately protective?

Inconsistent evaluation and interpretation of chemical cancer risk inhibits innovation while not enhancing safety

Douglas C. Wolf (Syngenta, USA) provided examples for how inconsistent evaluation of cancer risk inhibits innovation of active substances, plant protection products, and biocidal products, without enhancing safety.

In the EU, active substances, as well as any plant protection products or biocidal products containing them, must be classified following the provisions of the EU CLP Regulation (EP and Council 2008). It is the aim of C&L to identify the hazardous properties of a substance or mixture by applying specific classification criteria to the available hazard data and then to provide appropriate hazard labeling and information on safety measures. Hence, the EU CLP classification scheme (just as the GHS; both generally follow the IARC approach) is exclusively hazard-based and does not take exposure into consideration (see also above; summary of H. Greim's presentation).

Classification of a substance as Category 1A or 1B carcinogenic, mutagenic, or reprotoxic generally precludes its approval as an active substance, safener, or synergist in plant protection products or biocidal products (EP and Council 2009, 2012). Derogations are possible under the Plant Protection Products Regulation (EP and Council 2009) if the use of the active substance is indispensible (for limited periods) to control a serious danger to plant health or if human exposure is negligible under realistic proposed conditions of use. Hence, substances and mixtures may still have to be classified even when placed on the market in forms that are not hazardous. Further, any metabolites of the substance that have the potential to enter the groundwater may not be more toxic than the parent substance based on the key adverse effect. Therefore, drivers for the development of new active substances planned to be marketed in Europe focus on C&L to prevent Category 1A/1B classification. Toxicological studies are designed and evaluated to comply with the respective cut-off criteria for substances and those metabolites that can enter the groundwater. Thereby, the evaluations disregard dose–response relationships or exposure considerations, i.e. parameters which are pivotal in ensuring human health and environmental safety.

An overview of active substances that were all classified as Category 2 carcinogens under the EU CLP Regulation and the diversity of conclusions regarding carcinogenicity risk assessment drawn by the *US EPA Office of Pesticide Programs Cancer Assessment Review Committee* (US EPA CARC) for these same compounds illustrates the limitations of an oversimplified, hazard identification-based carcinogenicity classification system (Table 2). Usage of the US EPA (2005a) WoE descriptors (Table 1) provides much more relevant information than the EU CLP categorization. Also, more than one WoE descriptor can be used following the US EPA Guidelines when a substance's effects differ by dose or exposure route. For example, the substance could be "likely to be carcinogenic" above a specified dose, but "not likely to be carcinogenic" below that dose because a key event in tumor formation does not occur.

In this regard, WoE evaluations should consider that carcinogenicity is not an inherent property of any chemical, but a stochastic process. Tumors can develop when permanent errors occur in the DNA. However, multiple errors are necessary for the DNA before a tumor can evolve and the errors need to accumulate in one single pluripotent stem cell in the respective tissue (Doe et al. 2019; Wolf et al. 2019).

The following case study shows how mechanisms of carcinogenicity should be considered during substance classification. For a specific active ingredient (the identity of which is proprietary business information), uterine tumors were observed in female rats (at high doses only) and liver and thyroid tumors were observed in male rats and/or mice. The

Table 2. Active ingredients: EU CLP classification as compared to decision of the US EPA Cancer Assessment Review Committee.

Active ingredient	Indication	EU CLP Carc. Category	Decision of the US EPA Cancer Assessment Review Committee
Acetochlor	Herbicide	2	Suggestive evidence of carcinogenic potential
Chlordimeform	Insecticide	2	Group B – probable human carcinogen
Fenoxycarb	Insecticide	2	Likely to be carcinogenic to humans
Fluometuron	Herbicide	2	Group C – possible human carcinogen (use of low dose extrapolation model)*
Isopyrazam	Fungicide	2	Likely to be carcinogenic to humans (linear low-dose extrapolation approach to estimate human cancer risk†)
Molinate	Herbicide	2	Suggestive evidence of carcinogenicity, but not sufficient to assess human carcinogenic potential
Nicosulfuron	Herbicide	2	Group E – evidence of non-carcinogenicity for humans; NOAEL from dog study used for chronic effects
Pirimicarb	Insecticide	2	Likely to be carcinogenic to humans (linear low-dose extrapolation approach to estimate human cancer risk‡)
Propazine	Herbicide	2	Not likely to be carcinogenic to humans
Pymetrozine	Insecticide	2	Likely human carcinogen (but used LOAEL from developmental neurotoxicity study + uncertainty factors for chronic reference dose¶)
Simazine	Herbicide	2	Not likely to be carcinogenic to humans
Tralkoxydim	Herbicide	2	Suggestive evidence of carcinogenicity, but not sufficient to assess human carcinogenic potential

EU CLP classification as per ECHA dissemination portal (https://echa.europa.eu); US EPA Cancer Assessment as per US EPA (2018); unless otherwise noted, see below.
*Source: US EPA (2005b).
†Source: US EPA (2017).
‡Source: US EPA (2005c).
¶Source: US EPA (2000).

US EPA CARC concluded that the active ingredient should be classified as having "suggestive evidence of carcinogenic potential" since the available evidence was not considered sufficient to support the company's proposed non-genotoxic MoA for the uterine tumor. Nevertheless, the US EPA CARC also concluded that quantification of cancer risk using a non-linear approach would adequately account for all chronic toxicity, including carcinogenicity, that could result from exposure to the active ingredient. Hence, even though the MoA was not fully defined, the most sensitive NOAEL could be identified and health-based exposure limits (such as references doses) could be derived to protect humans from potential carcinogenicity.

This conclusion from the US EPA CARC was founded on a risk-driven evaluation of the data showing that the liver and thyroid tumor response occurred only in male rats and/or mice exposed to the active ingredient, but not in the female rats or mice. The liver tumor response in male rats occurred late in the course of treatment and comprised mainly adenomas, whereas no carcinomas were observed. It was considered to be weak evidence of a treatment-related effect. The male mouse liver tumors included adenomas and combined adenomas and/or carcinomas. All non-neoplastic histopathological findings were considered background findings associated with the age and strain of mice. The thyroid tumor response in male rats comprised mainly adenomas; however, there was also an increase in combined adenomas and/or carcinomas. It was concluded that thyroid tumor incidence provided weak evidence of a treatment-related effect. There was no concern for mutagenicity. The available data were considered sufficient to support the proposed non-genotoxic MoA for liver tumors in male rats and mice and thyroid tumors in male rats, but they were not considered sufficient to support the proposed MoA for female rat uterine tumors, which only occurred in the high-dose group.

Taken together, the GHS (and hence also EU CLP) approach to carcinogenicity classification, developed in the 1970s and 1980s should be fundamentally revised. Classification based on the presence of tumors alone could potentially eliminate a substance's commercial utility even though there would be no risk of carcinogenicity to humans. Additionally, the standard rodent 2-year bioassay is a poor model for predicting human carcinogenicity. It includes doses that do not reflect human exposure scenarios and provides only limited information on dose-response relationships. It does not provide mechanistic information and the procedures to evaluate the findings from the 2-year bioassay are not harmonized (either across governments or, in some cases, across different agencies within specific governments). The 2-year bioassay has a high false-positive rate and can produce tumors that are irrelevant to humans. It is time- and resource-consuming and uses large numbers of animals (Cohen 2017; Felter et al. 2018; Doe et al. 2019; Sauve-Ciencewicki et al. 2019).

For these reasons, the science of carcinogenicity risk assessment should be advanced to facilitate the use of modern tools and technologies that allow identifying the context in which a substance could elicit an adverse effect in humans to ensure that appropriate risk management measures are taken.

However, the standard rodent 2-year bioassay is often legally mandated and typically used by authorities to address required information on the potential for chemicals to cause cancer. Therefore, it is being performed, even if its findings are not needed to protect human health, and often it is not helpful for hazard assessment. A first step toward abandoning the 2-year bioassay would be to establish accepted guidance for waiving it. Opportunities for waiving the 2-year bioassay are most appropriate where a strong argument for low exposure can be provided for the given substance or group of substances, an abundance of information is available, and all potential MoAs can be identified in less-than-lifetime studies.

Options to specify carcinogenicity waiver requests were presented at the 58th Annual Meeting of the Society of Toxicology (Hilton et al. 2019). Within the "Rethinking Carcinogenicity Assessment for Agrochemicals Project" (ReCAAP), Hilton and co-workers assessed (1) valinamide

carbamates (a group of fungicides) and (2) sulfonamides (herbicide safeners that increase the rate of herbicide metabolism) using a WoE decision-tree for carcinogenicity assessment. First, all available data were considered except for those from the 2-year bioassay. Next, the outcome of the WoE evaluation was consolidated by comparing it to the outcome of the 2-year bioassay. Following the provisions of the EU Plant Protection Products Regulation (EP and Council 2009; Commission 2013) and the US Food Quality Protection Act of 1996 that amended the Federal Insecticide, Fungicide and Rodenticide Act and the Federal Food, Drug, and Cosmetic Act (US Government 1996), the database for these substances includes intended use indication and class of chemistry, metabolic profile and *in vitro* and *in vivo* toxicological data (also from grouping and read-across) on short-term and subchronic toxicity, genotoxicity, hormonal perturbation, and immune suppression.

In the WoE evaluation of valinamide carbamates, Hilton et al. (2019) investigated if the potential carcinogenicity of one member of this group could also be established by read-across from data available for two other valinamide carbamates. Generally, human exposure to valinamide carbamates is such that it is below the level of concern, based on the respective margins of exposure. Upon oral exposure, valinamide carbamates are rapidly absorbed, widely distributed, extensively metabolized, and rapidly eliminated. They are not genotoxic. The subchronic target organs are the liver and the thyroid gland and there is limited evidence for hormone perturbation. With respect to potential for immunotoxicity, a 90-d dog study yielded lymphatic edema and atrophy of femoral and sternum bone marrow. As regards carcinogenic MoA, data available from rat studies for the two valinamide carbamates provided insufficient evidence to conclude on thyroid-pituitary homeostasis or the development of liver foci. There was no significant tumor formation in mice. The overall WoE conclusion was that there were insufficient data to extrapolate long-term effects and that further mechanistic information was needed to understand tumor formation in rats, but that further testing in mice should be waived. By comparison, the 2-year bioassays available for the two valinamide carbamates indicated that they were "likely to be carcinogenic to humans" based on combined tumor rates for female rat thyroid follicular cell adenoma and carcinoma (Hilton et al. 2019).

In the WoE evaluation of sulfonamides, Hilton et al. (2019) recorded that human exposure is such that it is below the level of concern, based on the respective margins of exposure. The sulfonamide under investigation was rapidly absorbed and rapidly excreted mostly in urine, primarily unchanged. There was no evidence for genotoxicity; the subchronic target organ was the urinary tract. The available database indicated no concern for hormone perturbation or immunotoxicity. As regards mechanisms of carcinogenicity, tumors associated with cell proliferation and formation of urinary tract crystals/calculi had been recorded, typically at high-doses. Read-across indicated no additional concerns for chronic effects or tumor formation. The overall WoE conclusion was that sufficient data were available to extrapolate long-term effects and that waiving both the rat and mouse 2-year bioassay was possible. By comparison, the conclusion from an existing 2-year bioassay was that the sulfonamide was "not likely to be carcinogenic to humans" at doses that do not cause urothelium cytotoxicity. Hence, the WoE conclusion was consistent with the conclusion from this 2-year bioassay (Hilton et al. 2019).

The WoE evaluations conducted by Hilton et al. (2019) show that new approaches for a risk-based evaluation strategy serve to avoid the 2-year bioassay without impairing human health protection. *In vitro* assays and subacute and subchronic *in vivo* studies can be used to evaluate potential carcinogenicity. The WoE approach allows identifying primary effects that lead to DNA damage or increased cell proliferation. Mechanisms of carcinogenicity are identified by determining hazardous properties including mutagenicity, genotoxicity, and target organ toxicity. Thereby, the WoE approach allows protecting humans against adverse long-term effects, including carcinogenicity. It includes dose-response considerations and allows setting exposure limits that prevent the primary effects, and, therefore, also the long-term effects.

A separate classification for carcinogenicity provides no additional human protection. On the contrary, cut-off-based hazard evaluation strategies, as currently implemented under the EU CLP Regulation (EP and Council 2008), inhibit the application of innovative approaches and delay or deny innovation. Additionally, observations of how tumors (that are often irrelevant to humans) form in rats are not helpful or required to reliably protect humans from carcinogenicity risks. Abandoning the rodent 2-year bioassay avoids waste of resources, time, and animals without impairing human health protection. D. C. Wolf concluded his presentation with a quote from the British economist John Maynard Keynes:

> The difficulty lies, not in the new ideas, but in escaping from the old ones. (Keynes 1936)

Establishing an adequate margin of protection for non-genotoxic carcinogens

A. R. Boobis pursued the question of how current C&L approaches could and should be advanced beyond mere hazard identification to allow establishing thresholds of effects and adequate margins of safety. These questions not only considered classification for carcinogenicity, but also the general risk assessment of chemicals.

When the risk assessment is performed to meet human health protection goals, its outcome is used to establish reference doses (human health exposure limits, e.g. ADI) for risk management and risk communication. These reference doses are based on the most sensitive relevant adverse effect observed in the available experimental studies, i.e. typically the lowest NOAEL, further making use of AFs to account for e.g. interindividual variability and species differences between humans and rodents (see also above; summary of H. Greim's presentation).

Since non-genotoxic carcinogens elicit effects by a non-linear threshold-based MoA, a NOAEL can be established for these substances from which to derive a reference dose.

Braakhuis et al. (2018) recorded that usually the NOAEL from subchronic repeated-dose toxicity studies is used as point-of-departure for the assessment of non-genotoxic carcinogens under REACH (EP and Council 2006) since carcinogenicity studies may generally only be performed if a specific concern is present. For 44 known non-genotoxic carcinogens for which both data from subchronic toxicity studies and carcinogenicity studies were available, Braakhuis et al. reported that the NOAELs from the two types of studies were similar. Braakhuis et al. also reported that the subchronic toxicity and carcinogenicity NOAELs were, on average, associated with a cancer risk of approx. 1% in rodents. They concluded that the carcinogenic risk in rodents at the NOAEL was not nil (i.e. approx. 1%) and that the derived health-based guidance values might not fully preclude a risk for tumor development in humans (Braakhuis et al. 2018). However, this line of arguments disregards that the definition for NOAEL does not require absence of adversity: "The NOAEL is the highest exposure at which there is no statistically or biologically significant increase in the frequency of an adverse effect when compared with a control group" (NRC 1994). Once more, this example shows the limitations of a binary, exclusively hazard identification-based approach to carcinogenicity risk assessment that disregards the threshold-based MoAs of non-genotoxic carcinogenicity, dose-response relationships, the human health relevance of effects, and relevant human exposure levels and exposure durations.

Depending on the MoA, carcinogenicity evolves by a sequence of key events and each of the intermediate key events has its own quantitative response-response relationship (Garcia-Reyero and Murphy 2018). The subsequent key event is only triggered if the effect caused by the preceding key event is of sufficient magnitude. For the non-systemic fungicide chlorothalonil, an MoA for renal carcinogenicity has been suggested that has glutathione conjugation and further metabolism to cysteine conjugates as the molecular initiating event, followed by active uptake of the cysteine conjugates by proximal convoluted tubular cells that may ultimately lead to renal cell regenerative proliferation (JMPR 2011). The resulting adverse outcome in rats is increased renal adenoma and carcinoma (Wilkinson and Killeen 1996; JMPR 2011).

As summarized by the 2009 Joint FAO/WHO Meeting on Pesticide Residues (JMPR 2011), different studies have indicated that chlorothalonil is nephrotoxic in rats. The following findings were reported:

- *28-d study*: Increased renal weight (lowest-observed adverse effect level (LOAEL) 80 mg/kg bw/d).
- *90-d study*: Increased blood urea nitrogen (NOAEL 40 mg/kg bw/d) and increased renal weight, renal hyperplasia, and karyomegaly in kidneys (LOAEL 40 mg/kg bw/d).
- *90-d study*: Increased renal weight (NOAEL 1.5 mg/kg bw/d), hyperplasia of the epithelium of proximal convoluted tubules (NOAEL 10 mg/kg bw/d).
- *2-year study*: Increased blood urea nitrogen, serum creatinine, and renal weight (NOAEL 3.8 mg/kg bw/d); hyperplasia of the epithelium of proximal convoluted tubules (NOAEL 1.8 mg/kg bw/d); renal adenomas and carcinomas in male and female rats (NOAEL 3.8 mg/kg bw/d).

The JMPR concluded that the formation of kidney tumors was the result of prolonged renal cytotoxicity and regenerative cell proliferation (similar to the chloroform example described in the summaries of the presentations by R.S. Schoeny (see above) and S. M. Cohen (see below)), and was consistent with a threshold phenomenon (since chlorothalonil further showed no evidence for genotoxicity). The JMPR established an ADI of 0–0.02 mg/kg bw for chlorothalonil based on the NOAEL of 1.8 mg/kg bw/d for kidney toxicity observed in the 2-year bioassay and using an AF of 100. (The ADI reflects the amount of a chemical to which an individual can be exposed daily over a lifetime without appreciable health risk.) The ADI for chlorothalonil provides a margin of 200 for the induction of renal tumors in rats (NOAEL 3.8 mg/kg bw/d). Based on the MoA, the JMPR concluded that, while it is plausible that humans are less sensitive to the renal effects of chlorothalonil, it was not possible to dismiss relevance to humans on quantitative grounds, nor was it possible to quantify any difference in sensitivity (JMPR 2011). (Notably, given the species differences in the β-lyase bioactivation pathway (Iyer and Anders 1996), the ADI is likely to be conservative.).

For cadmium, the MoA for nephrotoxicity is reported to be initiated when cadmium accumulates in the kidneys thereby eliciting thiol group activation. This leads to cytotoxicity which results in proximal tubular damage (identified by measuring beta-2-microglobulin levels in the urine). The proximal tubular damage leads to the nephrotoxic adverse outcome (Branca et al. 2018).

The potential renal carcinogenicity of cadmium was assessed in a standard 2-year bioassay using Noble NBL/Cr rats receiving 25, 50, 100, and 200 ppm cadmium chloride in the drinking water (Waalkes et al. 1999). The rats showed a low incidence of tumors at the two higher doses which was statistically significant by a trend test, albeit there were no significant differences in pairwise comparisons with the control group. By comparison, statistically significant increases in renal hyperplasia were recorded in all test groups (Waalkes et al. 1999). Hence, this 2-year bioassay provided limited evidence in rats that cadmium-induced nephrotoxicity might progress to tumor formation. For humans, the primary health concern related to oral uptake of cadmium is its accumulation in the proximal tubular cells where it may cause renal dysfunction (EFSA 2009).

Exemplarily, the *EFSA Panel on Contaminants in the Food Chain (CONTAM) scientific opinion on cadmium in food* (EFSA 2009) is presented to explore the existence of population thresholds for precursor effects of carcinogenicity, in this case, elevated beta-2-microglobulin as a measure of proximal tubular damage, indicating the potential renal carcinogenicity of cadmium. In its risk assessment of cadmium, EFSA CONTAM calculated the 5% lower bound of the benchmark dose confidence interval (BMDL5) using beta-2-microglobulin data available for 30 000 human subjects. The BMDL5 reflects the lower 95% confidence limit on a 5% extra risk (defined as the absolute change in risk divided by the non-affected fraction in the control population). The calculated BMDL5 of 4 μg Cd/g creatinine[1] was divided by a chemical-specific adjustment factor of 3.9, to account for the inter-individual

variation of urinary cadmium within the study populations, leading to a critical urinary level of 1.0 μg Cd/g cretonne. One-compartment modeling was applied to identify the dietary cadmium exposure that corresponds to the critical urinary level. The modeling indicated that, in order to remain below the critical urinary level in 95% of the population by the age of 50 years, the average weekly dietary cadmium intake should not exceed 2.52 μg Cd/kg bw (EFSA 2009).

This example shows that there was a clear population threshold for cadmium-induced proximal tubular damage, as indicated by elevations of beta-2-microglobulin, in a pooled analysis of data from 30 000 subjects. Hence, the health-based guidance value established by CONTAM to ensure that sensitive sub-populations will be protected from the nephrotoxicity of cadmium, will also ensure the protection of the population from any risk of renal carcinogenicity.

Carcinogenicity resulting from a non-genotoxic MoA will exhibit a true biological threshold. Whilst human sensitivity will vary, it will not be distributed infinitely, but rather will be subject to biologically-determined limits. If the respective exposure level necessary to elicit the molecular initiating event and all subsequent key events are not reached, the given adverse outcome will not occur. Preferably, the risk assessment should be based on mechanistic considerations and the identification of biological thresholds rather than statistically-based defaults. This is especially important for carcinogenicity that evolves through a non-genotoxic, threshold-based MoA.

A path forward for carcinogenicity evaluation without the 2-year bioassay

Samuel M. Cohen (*University of Nebraska Medical Center, USA*) presented opportunities to conduct carcinogenicity evaluations without the rodent 2-year bioassay.

The basic assumptions underlying the use of rodent 2-year bioassays are (1) that carcinogenic effects observed at the high doses used in the bioassay will also occur at the lower doses humans usually are exposed to (dose extrapolation); and (2) that chemicals that cause tumors in rodents also cause tumors in humans (species extrapolation). However, often these two assumptions are not valid. The 2-year bioassay has proven to be a poor predictor of human carcinogenicity. It generally does not provide mechanistic information and does not allow evaluating the human relevance of effects. Further, many non-genotoxic carcinogens elicit effects in rodents that are irrelevant to humans. Such effects include: (1) tumors in rodent organs that have no human counterpart (Zymbal's gland, Harderian gland, and forestomach); (2) rodent tumors without human analog (e.g. splenic mononuclear cell leukemia, mouse submucosal mesenchymal lesion of bladder); and (3) rodent tumors where the underlying MoA is not predictive of human risk (e.g. rat pancreas, mouse lymphoma, different tumors of endocrine and reproductive organs) (Alison et al. 1994; Cohen 2004).

These scientific limitations of the rodent 2-year bioassay provide a strong incentive to abandon its use. To meet this goal, a general approach for carcinogenicity assessment is needed that allows addressing all relevant genotoxic and non-genotoxic MoAs of human carcinogenicity.

Generally, direct genotoxic MoAs of carcinogenicity always involve DNA reactivity. Non-genotoxic MoAs of carcinogenicity always involve increased numbers of cell divisions (i.e. increased cell proliferation) as a key event; they involve a precursor non-cancer key event (e.g. immunosuppression, estrogenicity, or cytotoxicity and regeneration); and they involve a threshold. Further, carcinogenesis is a stochastic process that may evolve if a sequence of key events occurs (see also above; summary of D. C. Wolf's presentation). This implies that if the molecular initiating event takes place there is a certain probability that the adverse outcome will evolve. Hence, protecting against the respective precursor non-cancer key events will also protect against non-genotoxic carcinogenicity.

Against this background, a general approach for carcinogenicity assessment without the 2-year bioassay should begin by screening for DNA reactivity, immunosuppression, and estrogenic activity. Further screening should aim at identifying organ-specific effects, i.e. any evidence for increased proliferation, such as organ hyperplasia, cytotoxicity, etc. The findings from the *in vitro* and *in vivo* screening tests, and in silico modeling, should be evaluated together to identify any likely MoA and to establish the dose-response relationship and human relevance of the findings. The *World Health Organization (WHO)/International Programme on Chemical Safety (IPCS) Human Relevance Framework* (Boobis et al. 2006) provides a scientific framework to evaluate early changes that increase the probability of carcinogenicity and the human relevance of MoAs of carcinogenicity.

The example of hepatocellular carcinogenesis shows how the general approach for carcinogenicity assessment can be devised. MoAs for hepatocellular carcinogenesis include DNA reactivity and increased cell proliferation (that can either be receptor-mediated or non-receptor mediated). Receptor-mediated increases in cell proliferation can be initiated by peroxisome proliferator-activated receptor-alpha (PPARα), by nuclear receptors (e.g. constitutive androstane receptor, pregnane X receptor), by the aryl hydrocarbon receptor, by estrogen or by statins. Non-receptor mediated increases in cell proliferation can be initiated by e.g. cytotoxicity, viral processes, iron overload, or increased apoptosis.

PPARα agonists induce the formation of liver tumors by a sequence of key events including metabolic activation (if required) and PPARα activation (peroxisome proliferation and oxidative damage), which then leads to increased cell proliferation. Hence, the data supporting PPARα activation as a carcinogenic MoA include these key events: PPARα activation and hepatocellular hypertrophy; increased induction of peroxisomal enzymes and confirmation of increased peroxisomes by transmission electron microscopy; and increased cell proliferation (that should be reversible upon discontinuance of the respective treatment).

Criteria for establishing the non-receptor mediated key event of cytotoxicity include non-DNA reactivity of the substance, histopathological indications of cytotoxicity (i.e. necrosis and/or increased apoptosis), the induction of serum enzymes indicating toxicity, and increased labeling index,

and/or increased numbers of hepatocytes as indicators of increased cell proliferation. Further, there should be a parallel dose-response for cytotoxicity and carcinogenicity, effects should be reversible, and other MoAs should be ruled out, e.g. using a WoE approach.

A comparison of the key events of the MoA of carcinogenicity of chloroform in rodents *versus* humans (see above; summary of R. S. Schoeny's presentation) shows that, also when key events are concordant, the MoA in rodents – leading to an adverse outcome – is not necessarily relevant to predict effects in humans, since human exposures are often below that required to produce chronic toxicity. In both rodents and humans, the molecular initiating event involves generation of phosgene/HCl by CYP2E1 which then leads to cytotoxicity. In rodents, the cytotoxicity has been shown to lead to cell proliferation and regeneration, which may ultimately lead to the formation of liver tumors. By comparison, there are no human data indicating such cell proliferation and regeneration, and the human data related to tumor formation are inadequate. Therefore, based upon the available data alone, tumor formation in humans cannot be ruled out. Nevertheless, the MoA for chloroform carcinogenicity in rodents – leading to an adverse outcome – is not relevant for human exposures: it requires high-dose sustained lifetime exposures in the animal studies (that further indicate dose-related and temporal thresholds of effects). By comparison, humans are not exposed to chloroform at high doses over any considerable periods of time.

Allen et al. (2004) assessed if findings from available rodent 90-d studies (i.e. hepatocellular necrosis, hypertrophy, cytomegaly and/or increased liver weight) correlated with liver tumors in the 2-year bioassay (mice: 83 compounds; rats 87 compounds). All substances that elicited liver tumors in the 2-year bioassay showed at least one of the findings in the 90-d study. There were no false negatives, but a large number of false positives (Allen et al. 2004). This outcome showed that the 90-d study can be used to screen for hepatotoxicity. If none of the above-mentioned hepatic parameters are altered in the 90-d study, the substance is not a hepatocarcinogen. If one of the parameters is altered, it should be investigated if the substance is DNA reactive and its toxicokinetic properties and potential for metabolic activation should be established. If the substance is not DNA reactive, mechanistic screens should aim at identifying key events of potential non-genotoxic MoAs for carcinogenicity. Based thereupon, specific follow-up studies then serve to substantiate the likely MoA.

Figure 1 presents the decision-tree for a general screen for carcinogenesis that abstains from using the 2-year bioassay. Instead, the screen includes detailed subacute and subchronic repeated-dose toxicity studies (addressing organ weights, histopathology, cell proliferation, blood and urine chemistries, DNA labeling indices) and specialized studies, as relevant (e.g. immunohistochemistry, 'omics). The focus of the screen should be to predict non-cancer toxicity of non-genotoxic substances, since protecting against their non-cancer toxicity also protects against cancer toxicity. This short-term assay approach has identified all known human carcinogens, which is not the case for the standard full 2-year rodent bioassay (e.g. inorganic arsenic (Cohen et al. 2013; Cohen 2018) and cigarette smoking (Witchi 2005)).

S. M. Cohen concluded his presentation by invoking that it is time to stop doing rodent 2-year bioassays. An innovative approach to carcinogenicity hazard and risk assessment should be implemented, which, however, is a political challenge that often requires changes in laws and guidelines.

Stakeholder perspectives

W. Casey (*NIEHS/NTP, USA*) summarized ongoing activities within the NTP to re-design carcinogenicity assessment by the implementation of a translational toxicology pipeline. These activities are founded on the NTP vision to improve public health through the development of data and knowledge that is translatable, predictive and timely (NIEHS 2018). The need to move away from the rodent 2-year bioassay is substantiated by the fact that it does not provide timely information (taking several years from start to finish) and has limited translational relevance for protecting human health due to unrealistic exposure scenarios and lack of cancer site

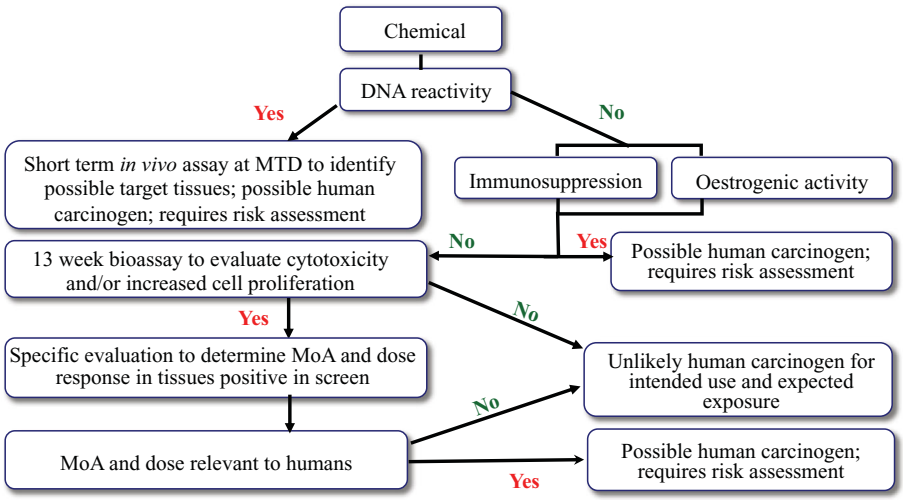

Figure 1. Decision-tree for a general screen for carcinogenesis that abstains from using the 2-year bioassay (modified from Cohen (2004); reproduced with permission of the author).

concordance. By comparison, the planned translational toxicology pipeline is expected to become an efficient and impactful tool for human carcinogenicity assessment. It includes a hypothesis-driven tiered approach covering data mining, computational toxicology, bioactivity screening, *in vitro* studies, short-term *in vivo* studies, chronic *in vivo* studies, and finally, knowledge integration. Due to the tiered structure of the translational toxicology pipeline, less reliance on chronic *in vivo* studies is anticipated.

The development of an innovative carcinogenicity assessment strategy is challenging since "cancer" is a collection of over 100 related diseases that may evolve by different pathways. To improve human relevance of the translational toxicology pipeline, it should include NAMs that are based on human data (Morgan et al. 2016). Different programs are underway to expand resources for data from the human population, including the *National Institutes of Health: All of Us* research program, a historic effort to gather data from one million or more US residents to accelerate research and improve health (https://allofus.nih.gov [accessed 2020 Feb]), and the *Million Veteran Program*, a national voluntary research program to learn how genes, lifestyle and military exposures affect health and illness, that is funded by the Department of Veterans Affairs (https://www.research.va.gov/mvp/ [accessed 2020 Feb]).

Next steps planned by the NTP to re-envision carcinogenicity assessment include further collaborations with stakeholders to understand e.g. their specific requirements, data needs, and contexts of use; endeavors to streamline access to NTP data; the identification of initial human cancer areas of focus; and the development of a communication plan.

Federica Crivellente (EFSA, Italy) explained that the standard data requirements for active substances in plant protection products, monitored by EFSA, include carcinogenicity testing in two species. The two species are mostly rats and mice, but other species should be considered depending on comparative metabolism assessments. Scientifically validated alternative carcinogenicity models may be used instead of the second species, and data on the MoA shall be provided for non-genotoxic carcinogens. Waiving of carcinogenicity testing is justifiable e.g. for natural products, highly toxic compounds, and microorganisms. Waiving is also possible if the expected human exposure is not long-lasting (e.g. in the case of ornamental plants) or if there is no evidence of exposure.

As stated in the Commission Regulation setting out the data requirements for active substances (Commission 2013), "the results of the carcinogenicity study taken together with other relevant data and information on the active substance shall be sufficient to permit the evaluation of hazards in humans." However, while the rodent 2-year bioassay was developed to explore carcinogenicity, it is most likely not relevant to predict carcinogenicity in humans.

As regards the exploration of carcinogenicity, the findings from chronic toxicity/carcinogenicity studies are used for the derivation of ADI levels and for carcinogenicity classification and they might also contribute to the assessment of endocrine disruption. EFSA recognizes that a mechanistic shift should be pursued with the aim to enhance the mechanistic understanding of carcinogenicity and to improve the predictivity of carcinogenicity assessments. For non-DNA reactive carcinogens that are either mitogenic or cytotoxic with regenerative proliferation, key events (e.g. target organ toxicity, hormonal activity, and immunosuppression) might be addressed in 90-d studies. In order to advance the carcinogenicity classification scheme, risk managers need to be convinced to accept radically different models for C&L. These models need to consider potency, and they should include *in vitro* MoA hazard classes to predict likely safe exposures for specific toxicity pathways rather than organ toxicity *per se*.

R. Corvi (*European Commission, JRC, Italy*) explained that the toolbox of novel approaches that have solid scientific bases and make use of innovative tools is growing. This offers the opportunity to make more informed decisions, which are mechanistic-based and human relevant and attain better levels of protection. Moreover, using animal-free methods also has positive ethical implications in line with *Directive 2010/63/EU on the protection of animals used for scientific purposes* (EP and Council 2010). However, it remains challenging to define how much information is really necessary for hazard and risk assessment, how to integrate the available data, and how to identify which new tools are needed to generate such data. To bring mechanistic data to bear in decision-making, there is a need to adopt some level of harmonization and standardization in the way data are generated, synthesized, and presented for regulatory use. As R. Corvi highlighted, "a wild-west mentality might get data published, but it is unlikely to prove credible to decision-makers in the long run." Therefore, the JRC is highly committed and engaged in activities to harmonize NAMs e.g. on the level of the OECD, the International Cooperation on Cosmetics Regulation (https://iccr-cosmetics.org [accessed 2020 Feb]) or the International Cooperation on Alternative Test Methods (ICATM[2]), and it fosters collaboration across such initiatives. To enhance the use of NAMs, scientific activities are required to continuously improve the mechanistic assays and political initiatives to update the relevant information requirements. To date, GHS classification criteria for non-animal methods have only been implemented for local toxicity endpoints skin corrosion and irritation, while work is ongoing for eye damage and irritation. It is hoped that these activities will facilitate the update of criteria for systemic toxicity endpoints.

The JRC is also engaged in activities that are specifically related to non-genotoxic carcinogens and that aim at moving away from the 2-year bioassay. For example, the JRC is co-leading the European Partnership for Alternative Approaches to Animal Testing (EPAA) activity on *Mechanism-based approach to cancer risk assessment of agrochemicals incorporating 3Rs principles* (https://ec.europa.eu/docsroom/documents/36296 [accessed 2020 Feb]) and is contributing to the OECD working group on the development of an integrated approach for testing and assessment (IATA) for non-genotoxic carcinogens (Jacobs et al. 2016). Moreover, the JRC is taking a broader look into an approach to carcinogenicity assessment which keeps pace with cancer burden and evolving chemical environment and considers more human data, related to etiology and links with other diseases, most

common types of cancer and biomarkers of exposure (Corvi et al. 2017; Madia et al. 2019). The reduction or abandonment of the 2-year bioassay requires in the short-term new approaches to be implemented within the current regulatory frameworks and long-term initiatives to redesign the hazard and risk assessment paradigm as such.

Jonas Nygren (*ECHA, Finland*) highlighted that attempts to change the current carcinogenicity hazard and risk assessment paradigm would require revising the applicable legislation, as well as the GHS. As carcinogenicity hazard assessment serves to identify if a substance might be carcinogenic in any organ, MoA data alone for a particular cancer type may not be very useful, unless showing a positive result. Still, it would currently be difficult to classify using MoA data alone without revision of the legislation. J. Nygren pointed out that the database to inform on carcinogenicity may be sparse: Under Annex X of the REACH Regulation (EP and Council 2006), a carcinogenicity study may be requested for high-production volume substances with widespread dispersive use or frequent or long-term human exposure if there is a specific concern for carcinogenicity e.g. from the genotoxicity tests or repeated-dose toxicity studies. Due to these restrictions, ECHA has so far requested fewer than 10 carcinogenicity studies (Karamertzanis et al. 2019). The inclusion of substances to the list of harmonized C&L of hazardous substances within Annex VI of the CLP Regulation (EP and Council 2008) has an impact on restricting the use of substances. J. Nygren mentioned that the REACH Regulation takes socioeconomic implications and exposure scenarios into account in the restriction and authorization processes. This is a complex and time-consuming process. For one substance, there can be several restrictions for different uses. This made the concept of coupling risk assessment to classification unclear as there would often be a need for multiple risk assessments. *Disclaimer*: The views expressed by Jonas Nygren are his own and may not represent the view of ECHA.

Marina Pereira (*Humane Society International, United Kingdom*) presented a perspective on the replacement of animals in the carcinogenicity assessment of chemicals as mandated by the 3Rs principle (Russell and Burch 1959) and the Directive 2010/63/EU (EP and Council 2010).

In vitro and short-term *in vivo* approaches are available to assess genotoxic carcinogens. With respect to the assessment of non-genotoxic carcinogens, an abundance of evidence confirms the irrelevance and poor reproducibility of the rodent 2-year bioassay (Ennever and Lave 2003; Knight et al. 2006). If the 2-year bioassay is performed, addition of a second species, i.e. mice, does not add to the safety evaluation (Billington et al. 2010). In the short-term, data from subchronic toxicity studies should be used to predict (the absence of) carcinogenic potential in humans (van der Laan et al. 2016; Woutersen et al. 2016). Ultimately, scientific and political activities should aim at establishing an approach for carcinogenicity testing that begins by *in vitro* investigations of pathways of carcinogenicity followed by IATAs to inform on the human health relevance of the pathways and the carcinogenic potential and potency of the given substance in humans.

Panel discussion

Moderator: **Philip A. Botham** (Syngenta, UK)
Panel members: **Diane Benford** (independent toxicological risk assessor, UK), **A. R. Boobis**, **S. M. Cohen**, **W. Dekant**, **D. C. Wolf**

The scientific limitations of the current binary hazard identification-based carcinogenicity classification scheme and those of the rodent 2-year bioassay are closely interlinked as the classification scheme builds upon the findings from the bioassay. It does not serve human health protection to know that a substance is carcinogenic in rodents upon lifetime exposure to very high doses. Instead, assessments should aim at identifying a substance's potential to be carcinogenic in humans, i.e. the sequence of key events associated with the development of tumors and whether those can occur in humans and under what exposure conditions. Importantly, this information is only relevant for hazard assessment if it occurs at (*in vitro* or *in vivo*) doses that reflect realistic exposure scenarios in humans. Hazard identification without concordant establishment of dose-response relationships and exposure assessment does not serve human health protection. The dialog should be sought within the scientific community to enhance the understanding that non-genotoxic carcinogenicity – the focus of the present workshop – is not an inherent property, i.e. specific feature, of any given substance, but depends upon dose levels and exposure durations.

NAMs are available and being assembled in IATAs to assess the different sequences of key events of the different non-genotoxic (and also genotoxic) MoAs that may ultimately result in tumor formation. When such NAMs are applied, dose-response relationships should be established for each of the key event relationships to address the probability that the sequence of key events will continue up until the adverse outcome. Sound *in vitro* to *in vivo* extrapolations are pivotal to ensure the relevance of findings from the lower tiers. In the higher tiers, the IATAs should include subacute and subchronic *in vivo* testing to identify precursor non-cancer toxicity, again considering relevant doses and dose-response relationships. If observed effects are below the threshold of toxicological concern, they are not relevant. *Vice versa*, if effects are only observed at dose levels that by far exceed realistic human exposure scenarios, they are also not relevant. (If a substance is not genotoxic and/or carcinogenic, information is still needed on the other systemic toxicological endpoints, such as reproductive toxicity.).

The overall outline of the IATA depends upon the purpose of the hazard assessment, i.e. if it follows the bottom-up approach to predict carcinogenicity (as is done under the REACH Regulation) or the top-down approach to obtain mechanistic information on a known carcinogen (as is done in the area of food contaminant safety evaluation).

The evaluation of findings obtained in the respective tiers of the IATA should address not only the statistical significance, but most importantly, also their biological relevance. For example, if cell proliferation is observed in a short-term *in vivo* study, this does not by itself allow the conclusion that the substance is a carcinogen, but only that the likelihood of

carcinogenicity is increased. The further evaluation should consider the dose at which effects are observed to determine if the concern needs to be followed up by higher-tier testing or by the implementation of appropriate risk management measures to ensure that human exposures only take place below the level of concern.

This further highlights the critical role of exposure assessment. All initiatives to advance the current carcinogenicity risk assessment paradigm should not only include hazard assessment, but equally exposure assessment. Under the REACH Regulation, production volume is used as a surrogate for exposure to stimulating the need for carcinogenicity testing, but the two are not necessarily interrelated. Many high production volume chemicals are used only in closed-system manufacturing settings, so that exposure potential is minimized and takes on a lesser role in the risk assessment process.

The assessment of a substance's exposure potential needs to consider both external exposure (e.g. dietary exposure, dermal exposure) and internal exposure (i.e. systemic bioavailability in the organism or human body). In the last 15 years, models have become available to predict external exposure, further considering the environmental fate of the given substance. Similarly, physiologically-based kinetic models that integrate structural and toxicokinetic data have become available to characterize internal exposure. Such tools are applied to different extents by different industry sectors. Cross-sector communication and collaboration will facilitate the harmonized usage of exposure models, just as data sharing between different consortia will facilitate the further development of the models (Laroche et al. 2018).

While exposure assessment most commonly refers to human exposure (including route(s), frequency, magnitude, and duration), it is equally important to consider exposure conditions associated with the respective biological test. If tumor incidence is increased in the rodent bioassay following high-dose administration of a non-genotoxic substance, this does not mean that there is a cancer risk associated with lower exposures, and raises questions whether the substance should be classified at all. (Likewise, care should be taken to conduct *in vitro* assays at doses reflecting relevant exposure scenarios.).

In this regard, it is noteworthy that substances registered under REACH were not designed to have a biological effect (contrary to e.g. pharmaceuticals or active substances in plant protection products). For the majority of the substances registered under REACH, the hazard assessment serves to prove the absence of findings. This is much more time- and resource-intensive than providing evidence for the presence of an effect, and it may include the need to use large numbers of animals. The ECHA dissemination portal comprising the information from the REACH registration dossiers, that largely encompass negative data, should be used for the development and improvement of predictive tools to determine quantitative structure-activity relationships and to perform exposure modeling for substances with low or negligible toxicity potential.

When aiming to replace the rodent 2-year bioassay with an IATA that provides at least the same level of safety protection, it should not be the goal to reproduce the outcome of the bioassay, which is poorly predictive and yields large numbers of false positives and false negatives. A science-based IATA that makes use of increasingly complex assessments of key events of MoAs of carcinogenicity and that incorporates dose-response relationships and information on organ- or tissue-specific potential cytotoxicity, is better suited to predict carcinogenicity in humans than the 2-year bioassay. Such an approach also calls into question the relevance of requesting a carcinogenicity classification for non-genotoxic carcinogens at all, considering their hazard potential is driven by non-carcinogenic endpoints and the hazard assessment is founded on non-carcinogenic endpoints alone.

Application of IATAs, and the NAMs that are incorporated therein, in a regulatory setting requires changing the underlying legislation. For this purpose, the validity of the IATAs and NAMs must be established. However, laying out the way forward for how to advance IATAs and NAMs up until regulatory acceptance exceeded the scope of this workshop. An efficient way forward is to perform a prospective assessment, as is currently undertaken by the US EPA, to show that the 2-year bioassay is not needed for carcinogenicity hazard and risk assessment. However, changes in the relevant legislation alone will not be sufficient to ensure that innovative approaches are truly applied. Many NAMs require expert knowledge both for their performance and for the evaluation of the test results. Therefore, education is key to facilitate the adoption and usage of IATAs including NAMs for carcinogenicity assessments.

Finally, whenever the C&L of a substance is mandated, e.g. for legal purposes, it should be ensured that it is applied together with risk assessment, i.e. the consideration of dose-response relationships and exposure assessment.

Concluding thoughts

P. A. Botham summarized that the ECETOC workshop "*Hazard Identification, Classification and Risk Assessment of Carcinogens: Too Much or Too Little?*" had served to explore both the scientific limitations of the current, outdated binary hazard identification-based carcinogenicity classification scheme and those of the rodent 2-year bioassay (see also Table 3 for highlights from the workshop). The carcinogenicity classification system implemented in most jurisdictions does not include dose-response considerations or information on the mechanisms of carcinogenicity. Therefore, it is inadequate for the classification of non-genotoxic carcinogens that elicit effects by a threshold-based MoA. Further, classification for carcinogenicity is often based upon data from the 2-year bioassay which is a poor model for predicting human carcinogenicity. Depending on the doses applied, 50% of the substances tested in the 2-year bioassay are found to be rodent carcinogens (Cohen et al. 2019; Doe et al. 2019). Nevertheless, many of these are likely to be false positives because high doses, such as the MTD, may disturb cellular functions, whereas this process would not occur at lower doses relevant for human exposures (Cohen et al. 2019; Doe

Table 3. Highlights from the ECETOC Workshop "Hazard Identification, Classification and Risk Assessment of Carcinogens: Too Much or Too Little?"

Impediment	Challenge	Way forward	Reconciliation
		Focus: non-genotoxic carcinogens	
Scientific limitations of the rodent 2-year bioassay Binary carcinogenicity classification system	Further develop and build confidence in IATAs utilizing NAMs Change risk assessors' and managers' perceptions of hazard and risk assessment	Research and development Capacity-building Education/communication, also to inform on value of mechanistic information Changes in legislative frameworks; agreement of internationally harmonized framework	*Short and medium term*: modify existing approaches to enhance consideration of exposure, better use of existing knowledge on the biology of cancer in interpreting less-than-lifetime studies, MoAs (IATAs, NAMs) during hazard and risk assessment and to discontinue use of rodent 2-year bioassay *Long-term*: discontinue C&L for carcinogenicity

et al. 2019). Similarly, simply by increasing the number of animals per dose group from 50 to 200, an estimated 92% of substances would show statistically significant dose-response trends for carcinogenicity at one or more tissue sites in either sex of rats or mice (Gaylor 2005).

For some sectors of industry, such as the pesticide industry, the erroneous classification of an innovative substance as a carcinogen generally eliminates its use even though there would be no risk to humans. Importantly, in the public perception (and that of the media), a substance classified as carcinogenic is viewed as posing a risk. It is very difficult to communicate that the current hazard identification-based classification scheme disregards exposure potential and the dose-response relationships of non-genotoxic carcinogenicity. While the scientific evidence clearly indicates the need to update the outdated approach to carcinogenicity testing and classification, "cancer" is a highly emotive subject. Therefore, politicians are disinclined to update the corresponding legal provisions with huge economic implications for both industry and consumers.

The workshop served to identify opportunities to overcome these limitations, taking into account the specificities of different legislation implemented in the EU and the USA. While the scope of the workshop specifically addressed non-genotoxic carcinogens, many of the topics discussed are also of relevance for genotoxic carcinogens.

Carcinogenicity hazard and risk assessment should focus on those substances that truly pose a human health concern and on the identification of such substances using NAMs, to either end their production or implement adequate risk management measures to prevent human exposure. At the same time, it must be ensured that substances that are safe are not subjected to unnecessary regulatory restrictions. Usage of the rodent 2-year bioassay should be abandoned. As the workshop presentations showed, methods and approaches are available to investigate if a substance may be carcinogenic, including NAMs, and by which MoA the effect will occur. Non-genotoxic carcinogens exhibit a threshold-based MoA so that safe levels of use can be established.

However, the available NAMs should be used skillfully. For example, high-dose effects recorded in the rodent 2-year bioassay should not be mimicked in the NAMs (just as NAMs should not be validated against the bioassay). *In vitro* dosimetry considerations are pivotal to ensure that *in vitro* doses reflect internal exposure in humans. Dialog should be reinforced within the scientific community that carcinogenicity testing at the MTD may not yield data that are relevant for effects in humans, both because of the irrelevance of such high doses and since they cause disturbances of homeostasis, which may not occur at lower doses. Therefore, the evaluation of findings should not be based upon statistical significance alone, but should also consider biological relevance. P. A. Botham asked if the toxicology community could agree on the opportunities that NAMs bring rather than creating further challenges, e.g. by demanding that bioassays are conducted at doses that even exceed the MTD (Slob 2014).

Finally, the discussions and presentations of the workshop, and the attendance of 150 participants in the workshop, are very encouraging. They are not only testimony to the relevance of the topic, but also re-emphasize the importance of engaging in science-based communication to re-design the hazard and risk assessment and C&L of non-genotoxic carcinogens.

Notes

1. The expression of data as a function of creatinine is the most common approach to determine levels of, e.g. heavy metals in the urine.
2. https://ec.europa.eu/jrc/en/eurl/ecvam/alternative-methods-toxicity-testing/advisory-bodies/icatm [accessed 2020 Feb].

Acknowledgements

All presenters and panellists are thanked for their valuable contributions: Diane Benford (independent toxicological risk assessor, UK), Warren Casey (NIEHS / NTP, USA); Samuel M. Cohen (University of Nebraska Medical Center, USA), Raffaella Corvi (EU Commission, JRC, Italy); Federica Crivellente (EFSA, Italy), Wolfgang Dekant (University of Würzburg, Germany), Jonas Nygren (ECHA, Finland), Marina Pereira (Humane Society International, United Kingdom), Rita S. Schoeny (Rita Schoeny LLC, USA), Douglas C. Wolf (Syngenta, USA). All workshop participants are thanked for their constructive input to the workshop discussions. Francesca Uguccioni (ECETOC, Belgium) is thanked for her support in organising the workshop.

Declaration of financial interests

This manuscript relates the proceedings of a one-day workshop "*Hazard Identification, Classification and Risk Assessment of Carcinogens: Too Much or Too Little?*" held in Helsinki, Finland, on September 8, 2019 as a satellite meeting to the EUROTOX 2019 Conference. The workshop was sponsored by the European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC; see: www.ecetoc.org). ECETOC is a scientific organisation which provides a collaborative space for scientists from industry, academia and governments. Its mission is to develop and promote practical, trusted and sustainable solutions to scientific challenges which are valuable to industry, as well as to the regulatory community and society in general. ECETOC is financed by its membership, which are the leading companies with interests in the manufacture and use of chemicals, biomaterials and pharmaceuticals. The list of member companies is available at: http://www.ecetoc.org/ecetoc-membership/member-companies/.

ECETOC provided travel and accommodations support (including economy air fare) for non-industry speakers and panellists on request. This offer was taken up by Samuel M. Cohen (University of Nebraska Medical Center, USA) and Marina Pereira (Humane Society International, United Kingdom). Further, ECETOC covered the economy air fares of Alan R. Boobis and Helmut Greim to attend the workshop (see also Declaration of Competing Interests).

All co-authors (with the exception of UGS), speakers and panellists participated at the workshop and contributed to the preparation of the manuscript without compensation. As described further in the Declaration of Competing Interests, UGS – a freelance scientific writer – did receive compensation to draft the manuscript.

Declaration of competing interests

All workshop speakers and panellists (that included representatives from academia, regulators / authorities, industry and non-governmental organisations; as listed in the Acknowledgements) and the sponsor of the workshop were provided the opportunity to review a draft of this manuscript, with all comments being addressed. The review by the sponsor of the workshop, via the ECETOC Scientific Committee consisting of representatives of academia and industry (for complete list of members, see http://www.ecetoc.org/about-ecetoc/scientific-committee/), yielded few minor comments only, all of which relating to the coherence of the manuscript. All workshop speakers and panellists read the final manuscript and agreed that it was ready for submission.

The co-authors consist of the Organising Committee for the workshop (SPF, ARB, PAB, AB, HG, HMH) and the scientific writer who drafted the manuscript (UGS). The views expressed in this article are solely those of the co-authors and may not represent those of the sponsoring organisations.

SPF is employed by Procter & Gamble (P&G), a consumer products corporation that produces and markets many products globally. Although P&G does not manufacture any chemicals discussed in this workshop, the issue of hazard identification and risk assessment of non-genotoxic carcinogens, and how these are assessed by regulatory/other agencies, does impact substances of interest to the Company. This includes β-myrcene, which was discussed as a case study in the workshop. This is a flavour chemical that is also used in perfumes (not directly added, but present in natural extracts). The responsibilities of SPF include development of the company's methods for cancer risk assessment across all product categories and geographies. SPF is a member of several scientific advisory boards, each of which includes aspects relevant to the testing and/or interpretation of data on chemical carcinogenesis: The Science Advisory Board of the U.S. Food and Drug Administration's National Center for Toxicological Research; the Chartered Science Advisory Board for the U.S. Environmental Protection Agency; and the Board of Scientific Counselors for the National Toxicology Program. SPF received no funding in cash or kind for her contribution to this manuscript; her travel expenses to participate in this workshop were covered entirely by P&G. An in-house review of this manuscript was conducted by P&G scientists, which resulted in only minor editorial changes.

ARB was employed by Imperial College London, UK, where he was involved in academic research and teaching in the fields of biomedicine and toxicology until he retired in June 2017. He was then employed 4 hours per week by the College to complete an EU Horizon 2020 project on mixture toxicology (EuroMix), until May 2019. He has no current employment. ARB is collaborating or has collaborated on a number of activities on chemical carcinogenesis through FAO/WHO JECFA, FAO/WHO JMPR, WHO-IPCS, WHO Tob-Reg, EFSA, ILSI HESI, ILSI Europe, ILSI Research Foundation, and UK Committees on Toxicity (COT), Carcinogenicity (COC), and the Medical Effects of Air Pollutants (COMEAP). ARB is a member of several scientific advisory boards, but none of these specifically involves chemical carcinogenesis. None of these collaborative activities is or was remunerated and no research funding was received. ECETOC covered his economy air fare to attend the workshop. ARB received no funding in cash or kind for his contribution to this manuscript.

PAB is employed by Syngenta, an international agribusiness that markets crop protection chemicals and seeds. Syngenta markets products (or previously marketed products) containing some of the chemicals used as case studies in this paper, including the anonymised active ingredient cited by D. C. Wolf as well as chlorothalonil, valinamide carbamates and sulfonamide herbicides. This manuscript was subjected to the usual internal peer-review process in Syngenta, in this case by the Global Head of Product Safety, but no changes were requested. PAB's current responsibilities are to provide strategic scientific advice on product safety issues to the company's Product Safety, Business Sustainability and Crop Protection Development organisations. PAB also acts as a science scientific advisor in human safety to the European Crop Protection Association (ECPA) and to Crop Life International (CLI). PAB's expenses for the attendance at the workshop were paid by Syngenta; he received no funding in cash or kind for his contribution to this manuscript.

AB is a scientific consultant employed by Peter-Fisk Associates Brussels (PFA-B) since 2016. She has been contracted since January 2017 through PFA-B to support and manage the ECETOC Programme on Human Health and Exposure Sciences. This workshop activity has been presented to and approved by the ECETOC Scientific Committee in 2018. Since then, AB has been assisting the Organising Committee together with the administration staff at ECETOC as well as liaising with the EUROTOX Secretariat. AB received no funding in cash or kind for her contribution to the activity. Further, as she did not attend the workshop, no expenses for travel or accommodation accrued for AB.

HG is Professor emeritus of the Technical University Munich, formerly being Director of the Institute of Toxicology and Environmental Hygiene. His research experience has included drug metabolism, toxicokinetics, mechanisms of carcinogenic agents, *in vitro* test systems. HG has published over 500 papers in toxicology and risk assessment and has lectured on these subjects in Europe and abroad. Besides many contributions to textbooks, HG has edited and published two textbooks in toxicology, one in German, the other by Wiley, London (*HG and R. Snyder: Toxicology and Risk Assessment. A comprehensive Introduction, 2nd ed., 2016*). In June 2012, the book "*The cellular response to the genotoxic insult: the question of threshold for genotoxic carcinogens*" (HG and R. Albertini) has been published by the Royal Society of Chemistry, London UK. Due to his long experience in toxicological research and risk assessment, HG has been member or chair of numerous national and international scientific committees. In recent years he has been invited to public hearings of the German and European Parliament on topics such as glyphosate, Diesel emissions and endocrine disrupters (recently on classification and labelling of endocrine disrupters, European Commission in Brussels, November 8, 2019). ECETOC covered his economy air fare to attend the workshop. HG received no funding in cash or kind for his contribution to this manuscript.

HMH is employed by Dow Europe GmbH, a subsidiary of the global chemical corporation The Dow Chemical Company. The issue of hazard identification and risk assessment of non-genotoxic carcinogens, and how these are assessed by regulatory/other agencies, does impact substances of interest to the corporation. HMH's role includes the harmonisation, scientific advancement and conduct of risk assessments, including the endpoint carcinogenicity, across multiple product categories and geographies. HMH received no funding in cash or kind for her contribution to this manuscript. An in-house review of this manuscript was conducted by additional scientists of The Dow Chemical Company but no changes were requested.

UGS has been a freelance scientific consultant and scientific writer since 2007. Her clients have included academia, non-governmental organisations, authorities and industry (companies, unions and related organisations). UGS was hired by ECETOC to assist in the preparation of the present manuscript. This included payment of working hours and reimbursement of accommodation and travel expenses (economy air fare) for participation at the workshop (purpose: note taking).

References

Adams TB, Gavin CL, McGowen MM, Waddell WJ, Cohen SM, Feron VJ, Marnett LJ, Munro IC, Portoghese PS, Rietjens IM. et al. 2011. The

FEMA GRAS assessment of aliphatic and aromatic terpene hydrocarbons used as flavour ingredients. Food Chem Toxicol. 49:2471–2494.
Alison RH, Capen CC, Prentice DE. 1994. Neoplastic lesions of questionable significance to humans. Toxicol Pathol. 22:179–186.
Allen DG, Pearse G, Haseman JK, Maronpot RR. 2004. Prediction of rodent carcinogenesis: an evaluation of prechronic liver lesions as forecasters of liver tumors in NTP carcinogenicity studies. Toxicol Pathol. 32: 393–401.
Bastaki M, Aubanel M, Bauter M, Cachet T, Demyttenaere J, Diop MM, Harman CL, Hayashi SM, Krammer G, Li X, et al. 2018. Absence of renal adverse effects from β-myrcene dietary administration in OECD guideline-compliant subchronic toxicity study. Food Chem Toxicol. 120:222–229.
Billington R, Lewis RW, Mehta JM, Dewhurst I. 2010. The mouse carcinogenicity study is no longer a scientifically justifiable core data requirement for the safety assessment of pesticides. Crit Rev Toxicol. 40: 35–49.
Bolt HM, Foth H, Hengstler JG, Degen GH. 2004. Carcinogenicity categorization of chemicals – new aspects to be considered in a European perspective. Toxicol Lett. 151:29–41.
Boobis AR. 2010. Mode of action considerations in the quantitative assessment of tumour responses in the liver. Basic Clin Pharmacol Toxicol. 106:173–179.
Boobis AR, Cohen SM, Dellarco VL, Doe JE, Fenner-Crisp PA, Moretto A, Pastoor TP, Schoeny RS, Seed JG, Wolf DC. 2016. Classification schemes for carcinogenicity based on hazard-identification have become outmoded and serve neither science nor society. Regul Toxicol Pharmacol. 82:158–166.
Boobis AR, Cohen SM, Dellarco V, McGregor D, Meek ME, Vickers C, Willcocks D, Farland W. 2006. IPCS framework for analyzing the relevance of a cancer mode of action for humans. Crit Rev Toxicol. 36: 781–792.
Boobis AR, Daston GP, Preston RJ, Olin S.S. 2009. Application of key events analysis to chemical carcinogens and noncarcinogens. Crit Rev Food Sci Nutr. 49:690–707.
Braakhuis HM, Slob W, Olthof ED, Wolterink G, Zwart EP, Gremmer ER, Rorije E, van Benthem J, Woutersen R, van der Laan JW, et al. 2018. Is current risk assessment of non-genotoxic carcinogens protective? Crit Rev Toxicol. 48:500–511.
Branca JJV, Morucci G, Pacini A. 2018. Cadmium-induced neurotoxicity: still much ado. Neural Regen Res. 13:1879–1882.
Cohen SM. 2004. Human carcinogenic risk evaluation: an alternative approach to the two-year rodent bioassay. Toxicol Sci. 80:225–229.
Cohen SM. 2010. An enhanced 13-week bioassay: an alternative to the 2-year bioassay to screen for human carcinogenesis. Exp Toxicol Pathol. 62:497–502.
Cohen SM. 2017. The relevance of experimental carcinogenicity studies to human safety. Cur Opin Toxicol. 3:6–11.
Cohen SM. 2018. Screening for human urinary bladder carcinogens: two-year bioassay is unnecessary. Toxicol Res. 7:565–575.
Cohen SM, Arnold LL, Beck BD, Lewis AS, Eldan M. 2013. Evaluation of the carcinogenicity of inorganic arsenic. Crit Rev Toxicol. 43:711–752.
Cohen SM, Boobis AR, Dellarco VL, Doe JE, Fenner-Crisp PA, Moretto A, Pastoor TP, Schoeny RS, Seed JG, Wolf DC. 2019. Chemical carcinogenicity revisited 3: risk assessment of carcinogenic potential based on the current state of knowledge of carcinogenesis in humans. Regul Toxicol Pharmacol. 103:100–105.
Cohen SM, Ellwein LB. 1991. Genetic errors, cell proliferation, and carcinogenesis. Cancer Res. 51:6493–6505.
Commission. 2012. Commission Implementing Regulation (EU) No 872/2012 of 1 October 2012 adopting the list of flavouring substances provided for by Regulation (EC) No 2232/96 of the EP and of the Council, introducing it in Annex I to Regulation (EC) No 1334/2008 of the EP and of the Council and repealing Commission Regulation (EC) No 1565/2000 and Commission Decision 1999/217/EC Text with EEA relevance. OJ L267:1–161.
Commission. 2013. Commission Regulation (EU) No 283/2013 of 1 March 2013 setting out the data requirements for active substances, in accordance with Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market. OJ L93:1–84.
Corvi R, Madia F, Guyton KZ, Kasper P, Rudel R, Colacci A, Kleinjans J, Jennings P. 2017. Moving forward in carcinogenicity assessment: report of an EURL ECVAM/ESTIV workshop. Toxicol In Vitro. 45: 278–286.
Council. 1967. Council Directive 67/548/EEC of 27 June 1967 on the approximation of laws, regulations and administrative provisions relating to the classification, packaging and labelling of dangerous substances. OJ L196: 1–98.
Council of Europe. 1974. Natural flavouring substances, their sources, and added artificial flavouring substances. Maisonneuve (France): Council of Europe.
[DFG] Deutsche Forschungsgemeinschaft (German Research Foundation). 2019. List of MAK and BAT values 2019. Report 55 of the Permanent Senate Commission for the Investigation of Health Hazards of Chemical Compounds in the Work Area. Germany: Wiley-VCH. [accessed 2019 Sep]. https://onlinelibrary.wiley.com/doi/pdf/10.1002/9783527826155.
Doe JE, Boobis AR, Dellarco V, Fenner-Crisp PA, Moretto A, Pastoor TP, Schoeny RS, Seed JG, Wolf DC. 2019. Chemical carcinogenicity revisited 2: current knowledge of carcinogenesis shows that categorization as a carcinogen or non-carcinogen is not scientifically credible. Regul Toxicol Pharmacol. 103:124–129.
[ECHA] European Chemicals Agency. 2012. Guidance on information requirements and chemical safety assessment. Chapter R.8: characterisation of dose [concentration]-response for human health. Version 2.1; ECHA-2010-G-19-EN, November 2012. Helsinki (Finland): European Chemicals Agency.
ECHA. 2017. Guidance on the application of the CLP criteria. Guidance to Regulation (EC) No 1272/2008 on classification, labelling and packaging (CLP) of substances and mixtures. Version 5.0; ECHA-17-G-21-EN, July 2017. Helsinki (Finland): European Chemicals Agency.
[ECHA RAC] European Chemicals Agency Committee for Risk Assessment. 2013. Opinion on an Annex XV dossier proposing restrictions on 1,4-dichorobenzene (ECHA/RAC/RES-O-0000003486-69-01/F); adopted 8 March 2013. [accessed 2020 Feb]. https://echa.europa.eu/documents/10162/7e4b19c5-8bba-f575-84bd-9fb1d918b7ad.
ECHA RAC. 2018a. Opinion on scientific evaluation of occupational exposure limits for acrylonitrile (ECHA/RAC/O-0000001412-86-188/F); adopted 9 March 2018. [accessed 2020 Feb]. https://echa.europa.eu/documents/10162/13641/acrylonitrile_opinion_en.pdf/102477c9-a961-2c96-5c4d-76fcd856ac19.
ECHA RAC. 2018b. Opinion on scientific evaluation of occupational exposure limits for benzene (ECHA/RAC/O-000000-1412-86-187/F); adopted 9 March 2018. [accessed 2020 Feb]. https://echa.europa.eu/documents/10162/13641/benzene_opinion_en.pdf/4fec9aac-9ed5-2aae-7b70-5226705358c7.
ECHA RAC. 2018c. Opinion on scientific evaluation of occupational exposure limits for nickel and its compounds (ECHA/RAC/O-000000-1412-86-189/F); adopted 9 March 2018. [accessed 2020 Feb]. https://echa.europa.eu/documents/10162/13641/nickel_opinion_en.pdf/9e050da5-b45c-c8e5-9e5e-a1a2ce908335.
[EFSA] European Food Safety Authority. 2009. Scientific opinion of the Panel on Contaminants in the Food Chain. Cadmium in food (question no EFSA-Q-2007-138), adopted 30 January 2009. EFSA J. 980:1–139.
EFSA. 2015. Panel on Food Contact Materials, Enzymes, Flavourings and Processing Aids. Scientific opinion on flavouring group evaluation 25, revision 3 (FGE.25Rev3): aliphatic hydrocarbons from chemical group 31. EFSA J. 13:4069.
[EFSA SC] European Food Safety Authority Scientific Committee. 2005. Opinion of the Scientific Committee on a request from EFSA related to a harmonised approach for risk assessment of substances which are both genotoxic and carcinogenic. (request no EFSA-Q-2004-020). EFSA J. 282:1–31.
Ennever FK, Lave LB. 2003. Implications of the lack of accuracy of the lifetime rodent bioassay for predicting human carcinogenicity. Regul Toxicol Pharmacol. 38:52–57.
EP and Council. 2006. Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending

Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. OJ L396:1.

EP and Council. 2008. Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures, amending and repealing Directives 67/548/EEC and 1999/45/EC, and amending Regulation (EC) No 1907/2006. OJ L353:1–1355.

EP and Council. 2009. Regulation (EC) No 1107/2009 of the European Parliament and of the Council of 21 October 2009 concerning the placing of plant protection products on the market. OJ L309:1–50.

EP and Council. 2010. Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes. OJ L276:33.

EP and Council. 2012. Regulation (EU) No 528/2012 of the European Parliament and of the Council of 22 May 2012 concerning the making available on the market and use of biocidal products. OJ L167:27.

Felter SP, Foreman JE, Boobis A, Corton JC, Doi AM, Flowers L, Goodman J, Haber LT, Jacobs A, Klaunig JE, et al. 2018. Human relevance of rodent liver tumors: key insights from a Toxicology Forum workshop on nongenotoxic modes of action. Regul Toxicol Pharmacol. 92:1–7.

[FEMA] Flavor Extract Manufacturers Association. 1965. Recent progress in the consideration of flavouring ingredients under the Food Additives Amendment. III. GRAS substances. In: Hall RL, Oser BL, editors. Food Technol. 19(2) Part 2; https://www.femaflavor.org/sites/default/files/3.%20GRAS%20Substances%282001-3124%29_0.pdf.

Fenner-Crisp PA, Dellarco VL. 2016. Key elements for judging the quality of a risk assessment. Environ Health Perspect. 124:1127–1135.

Garcia-Reyero N, Murphy CA. 2018. A systems biology approach to advancing adverse outcome pathways for risk assessment. Switzerland: Springer International Publishing AG.

Gaylor DW. 2005. Are tumor incidence rates from chronic bioassays telling us what we need to know about carcinogens? Regul Toxicol Pharmacol. 41:128–133.

Goodman JI. 1998. The traditional toxicological paradigm is correct: dose influences mechanism. Environ Health Persp. 106:285–288.

Goodman JI, Watson RE. 2002. Alterations in DNA methylation may play a variety of roles in carcinogenesis. Cell. 83:13–15.

Goto M, Hattori M, Miyagawa T, Enomoto M. 1972. Ecological contribution on chemistry. II. Formation of hepatoma in mice after ingestion of HCH isomers in high doses. Chemosphere. 1:279–282.

Greenfield RE, Ellwein LB, Cohen SM. 1984. A general probabilistic model of carcinogenesis: analysis of experimental urinary bladder cancer. Carcinogenesis. 5:437–445.

Hanada M, Yutani C, Miyaji T. 1973. Induction of hepatoma in mice by benzene hexachloride. Gan. 64:511–513.

Hennes C, Batke M, Bomann W, Duhayon S, Kosemund K, Politano V, Stinchcombe S, Doe J. 2014. Incorporating potency into EU classification for carcinogenicity and reproductive toxicity. Regul Toxicol Pharmacol. 70:457–467.

Hilton GM, Clippinger AJ, Papieni S, Buesen R, Melching-Kollmuss S, Kormos T, Ryan N, Wolf DC, Peffer RC, Hayes A.W. 2019. Better ways than the bioassay: weight-of-evidence approach to assess carcinogenicity potential of food-use pesticides. Abstract 2784. Poster session presented at: The 58th Annual Meeting and ToxExpo of the Society of Toxicology; Mar 10–14. [accessed 2019 Sep]. https://www.piscltd.org.uk/wp-content/uploads/2019/03/SOT-2019-Poster_Agchem-Bioassay-Waiver_Final_GMH.pdf.

[IARC] World Health Organization - International Agency for Research on Cancer. 2018. Monographs on the evaluation of carcinogenic risks to humans. Vol. 113. DDT, lindane, and 2,4-D. Lyon (France). [accessed 2020 Feb]. https://monographs.iarc.fr/wp-content/uploads/2018/07/mono113.pdf.

IARC. 2019a. Preamble to the IARC monographs on the identification of carcinogenic hazards to humans. Lyon (France). [amended 2019 Jan; accessed 2019 Sep]. https://monographs.iarc.fr/wp-content/uploads/2019/07/Preamble-2019.pdf.

IARC. 2019b. Monographs on the Evaluation of Carcinogenic Risks to Humans. Vol 119. Some chemicals that cause tumours of the urinary tract in rodents. Lyon (France). [accessed 2019 Sep]. http://publications.iarc.fr/575.

[ICH] International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. 1997. ICH harmonised tripartite guideline. Testing for carcinogenicity of pharmaceuticals S1B. Current Step 4 version; 16 July 1997. Geneva (Switzerland): ICH.

Iyer RA, Anders MW. 1996. Cysteine conjugate beta-lyase-dependent biotransformation of the cysteine S-conjugates of the sevoflurane degradation product compound A in human, nonhuman primate, and rat kidney cytosol and mitochondria. Anesthesiology. 85:1454–1461.

Jacobs M, Colacci A, Louekari K, Luijten M, Hakkert B, Paparella M, Vasseur P. 2016. International regulatory needs for development of an IATA for non-genotoxic carcinogenic chemical substances. ALTEX. 33:359–392.

[JECFA] Joint FAO/WHO Expert Committee on Food Additives. 2015. WHO food additives series: 70. Prepared by the 97th meeting of the joint FAO/WHO expert committee on food additives (JECFA). Safety evaluation of certain food additives. Geneva (Switzerland): WHO.

[JMPR] Joint FAO/WHO Meeting on Pesticides Residues. 2011. Pesticide residues in food 2009. Evaluations 2009: part II – toxicological, WHO/FAO of the United Nations, Rome. [accessed 2019 Sep]. https://apps.who.int/iris/bitstream/handle/10665/44522/9789241665254_eng.pdf;jsessionid=01CB356CE4980F3C35DC425455529221?sequence=1.

Karamertzanis PG, Atlason P, Nathanail AV, Provoost J, Karhu E, Rasenberg M. 2019. The impact on classifications for carcinogenicity, mutagenicity, reproductive and specific target organ toxicity after repeated exposure in the first ten years of the REACH regulation. Regul Toxicol Pharmacol. 106:303–315.

Keynes JM. 1936. The general theory of employment, interest and money. London: Palgrave Macmillan.

Knight A, Bailey J, Balcombe J. 2006. Animal carcinogenicity studies: 1. Poor human predictivity. Altern Lab Anim. 34:19–27.

Kobets T, Iatropoulos MJ, Williams GM. 2019. Mechanisms of DNA-reactive and epigenetic chemical carcinogens: applications to carcinogenicity testing and risk assessment. Toxicol Res (Camb). 8:123–145.

Kobets T, Williams GM. 2019. Review of the evidence for thresholds for DNA-reactive and epigenetic experimental chemical carcinogens. Chem Biol Interact. 301:88–111.

Laroche C, Aggarwal M, Bender H, Benndorf P, Birk B, Crozier J, Dal Negro G, De Gaetano F, Desaintes C, Gardner I, et al. 2018. Finding synergies for 3Rs – toxicokinetics and read-across: report from an EPAA partners' forum. Regul Toxicol Pharmacol. 99:5–21.

Madia F, Worth A, Whelan M, Corvi R. 2019. Carcinogenicity assessment: addressing the challenges of cancer and chemicals in the environment. Environ Int. 128:417–429.

Morgan MM, Johnson BP, Livingston MK, Schuler LA, Alarid ET, Sung KE, Beebe DJ. 2016. Personalized *in vitro* cancer models to predict therapeutic response: challenges and a framework for improvement. Pharmacol Ther. 165:79–92.

[NCI] National Cancer Institute Carcinogenesis. 1977. Technical report series No. 14. Bioassay of lindane for possible carcinogenicity. CAS No. 58-89-9; NCI-CG-TR-14. US Department of Health, Education and Welfare, Public Health Service, National Institutes of Health. [accessed 2019 Sep]. https://ntp.niehs.nih.gov/ntp/htdocs/lt_rpts/tr014.pdf.

[NIEHS] National Institute of Environmental Health Sciences. 2018. 2018–2023 strategic plan. Advancing environmental health sciences – improving health. NIH Publication No. 18-ES-7935. US Department of Health and Human Services. [accessed 2019 Sep]. https://www.niehs.nih.gov/about/strategicplan/strategicplan20182023_508.pdf.

[NRC] National Research Council of the National Academy of Sciences. 1983. Risk assessment in the Federal Government: Managing the process. Washington (DC): National Academy Press. [accessed 2020 Feb]. https://www.nap.edu/catalog/366/risk-assessment-in-the-federal-government-managing-the-process.

[NRC] National Research Council of the National Academy of Sciences. 1994. Science and judgement in risk assessment. Washington (DC): National Academy Press.

NRC. 2009. Science and decisions. Advancing risk assessment. Washington (DC): National Academies Press.

[NTP] National Toxicology Program. 2010. Technical report on the toxicology and carcinogenesis studies of beta-myrcene (CAS No. 123-35-3) in F344/N rats and B6C3F1 mice (gavage studies). NTP Tech Rep Ser. (557):1–163.

NTP. 2016. Report on carcinogens. 14th ed. Research Triangle Park (NC): US Department of Health and Human Services, Public Health Service. [accessed 2019 Sep]. https://ntp.niehs.nih.gov/go/roc14.

[OECD] Organisation for Economic Co-operation and Development. 2016. Series on testing and assessment No. 255. Guidance document on the reporting of defined approaches to be used within integrated approaches to testing and assessment. ENV/JM/MONO(2016)28. Paris (France): OECD; p. 27.

Roberts HR. 1982. Continuing the evolution of food safety. Regul Toxicol Pharmacol. 2:77–83.

Russell WMS, Burch RL. 1959. The principles of humane experimental technique. London: Methuen. Reprinted by UFAW; p. 238.

Sauve-Ciencewicki A, Davis KP, McDonald J, Ramanarayanan T, Raybould A, Wolf DC, Valenti T. 2019. A simple problem formulation framework to create the right solution to the right problem. Regul Toxicol Pharmacol. 101:187–193.

Slob W. 2014. Benchmark dose and the three Rs. Part II. Consequences for study design and animal use. Crit Rev Toxicol. 44:568–580.

Sonich-Mullin C, Fielder R, Wiltse J, Baetcke K, Dempsey J, Fenner-Crisp P, Grant D, Hartley M, Knaap A, Kroese D, et al. 2001. IPCS conceptual framework for evaluating a mode of action for chemical carcinogenesis. Regul Toxicol Pharmacol. 34:146–152.

Thorpe E, Walker AI. 1973. The toxicology of dieldrin (HEOD). II. Comparative long-term oral toxicity studies in mice with dieldrin, DDT, phenobarbitone, -BHC and -BHC. Food Cosmet Toxicol. 11: 433–442.

Timp W, Feinberg AP. 2013. Cancer as a dysregulated epigenome allowing cellular growth advantage at the expense of the host. Nat Rev Cancer. 13:497–510.

United Nations. 2017. Globally harmonized system of classification and labelling of chemicals (GHS). 7th ed. New York: United Nations. [accessed 2019 Sep]. https://www.unece.org/fileadmin/DAM/trans/danger/publi/ghs/ghs_rev07/English/ST_SG_AC10_30_Rev7e.pdf.

[US EPA] United States Environmental Protection Agency. 1984. Risk assessment and management: framework for decision making. EPA 600/9-85-002. [accessed 2019 Sep]. https://nepis.epa.gov/Exe/ZyPDF.cgi/20008KTF.PDF?Dockey=20008KTF.PDF.

US EPA. 1986. Guidelines for carcinogen risk assessment, 51 FR 33992. [accessed 2019 Sep]. https://cfpub.epa.gov/ncea/risk/recordisplay.cfm?deid=54933.

US EPA. 1987. National Center for Environmental Assessment. Integrated Risk Information System (IRIS) Chemical assessment summary. Gamma-Hexachlorocyclohexane (gamma-HCH); CASRN 58-89-9. [first online 1987 Jan 31; accessed 2019 Sep]. https://cfpub.epa.gov/ncea/iris/iris_documents/documents/subst/0065_summary.pdf.

US EPA. 1988. Reviews of environmental contamination and toxicology. In: Ware GW, editor. Office of drinking water health advisories. Vol 104. NewYork (NY); Springer Verlag. [accessed 2019 Sep]. https://link.springer.com/content/pdf/bfm:978-1-4613-8785-5/1.pdf.

US EPA. 1991. National primary drinking water regulations - synthetic organic chemicals and inorganic chemicals; monitoring for unregulated contaminants; National Primary Drinking Water Regulations Implementation; National Secondary Drinking Water Regulations; Final Rule. Fed Regist. 56:3526.

US EPA. 1997. Office of Prevention, Pesticides and Toxic Substances. EPA ID No.: 009001. Lindane: request for a waiver from the requirement for a series 83-2b carcinogenicity study with mice; DP Barcode: D238392; Submission No.: S528522; Reregistration Case No.: 0315. Washington (DC): United States Environmental Protection Agency.

US EPA. 2000. Office of Prevention, Pesticides, and Toxic Substances (7501C). Pesticide fact sheet. Name of chemical: pymetrozine. Reason for issuance: conditional registration. [accessed 2020 Feb]. https://www3.epa.gov/pesticides/chem_search/reg_actions/registration/fs_PC-101103_01-Aug-00.pdf.

US EPA. 2001. National Center for Environmental Assessment. Integrated Risk Information System (IRIS) Chemical assessment summary. Chloroform; CASRN 67-66-3 first online 31 January 1987, last revised 19 October 2001. [accessed 2020 Feb]. https://cfpub.epa.gov/ncea/iris/iris_documents/documents/subst/0025_summary.pdf.

US EPA. 2005a. Risk Assessment Forum. Guidelines for carcinogen risk assessment. EPA/630/P-03/001F; Washington (DC): United States Environmental Protection Agency. [accessed 2019 Sep]. https://www.epa.gov/risk/guidelines-carcinogen-risk-assessment.

US EPA. 2005b. Reregistration Eligibility Decision (RED) for fluometuron. Office of Prevention, Pesticides, and Toxic Substances. [accessed 2020 Feb]. https://www3.epa.gov/pesticides/chem_search/reg_actions/reregistration/red_PC-035503_1-Sep-05.pdf.

US EPA. 2005c. Memorandum – Pirimicarb: report of the Cancer Assessment Review Committee; PC Code: 106101; TXR No.: 0053618. Office of Prevention, Pesticides, and Toxic Substances. [accessed 2020 Feb]. https://archive.epa.gov/pesticides/chemicalsearch/chemical/foia/web/pdf/106101/106101-2005-07-13a.pdf.

US EPA. 2014. Framework for human health risk assessment to inform decision making. EPA/100/R-14-001. [accessed 2019 Sep]. https://www.epa.gov/sites/production/files/2014-12/documents/hhra-framework-final-2014.pdf.

US EPA. 2017. Isopyrazam; pesticide tolerance EPA-HQ-OPP-2016-0143; FRL-9960-76. Federal Register 82(100):24071. [accessed 2019 Sep]. https://www.govinfo.gov/content/pkg/FR-2017-05-25/pdf/2017-10765.pdf.

US EPA. 2018. Office of Pesticides Programs. Chemicals evaluated for carcinogenic potential. Annual cancer report 2018. [accessed 2019 Sep]. http://npic.orst.edu/chemicals_evaluated.pdf.

[US FDA] United States Food and Drug Administration. 2018. Memorandum. FAP A4810: review of the carcinogenicity potential of myrcene as a flavoring substance in food (FAP 5A4810). [accessed 2019 Sep]. https://www.regulations.gov/document?D=FDA-2015-F-4317-0073.

US Government. 1996. Food Quality Protection Act of 1996; an act to amend the federal insecticide, fungicide and rodenticide act and the federal food, drug, and cosmetic act. 110 Stat. 1489; Public law 104-170; 7 USC 136 note. [accessed 2019 Sep]. https://www.govinfo.gov/content/pkg/PLAW-104publ170/pdf/PLAW-104publ170.pdf.

US Government. 2016. Toxic Substances Control Act (TSCA) as amended by the Frank R. Lautenberg chemical safety act of the 21st century. United states code on FDSys. 15 USC chapter 53 toxic substances control; from Title 15: commerce and trade. [accessed 2020 Feb]. https://www.epa.gov/sites/production/files/2016-06/documents/bills-114hr2576eah.pdf.

Van der Laan JW, Buitenhuis WH, Wagenaar L, Soffers AE, van Someren EP, Krul CA, Woutersen RA. 2016. Prediction of the carcinogenic potential of human pharmaceuticals using repeated dose toxicity data and their pharmacological properties. Front Med (Lausanne). 3:45.

Waalkes MP, Anver MR, Diwan BA. 1999. Chronic toxic and carcinogenic effects of oral cadmium in the Noble (NBL/Cr) rat: induction of neoplastic and proliferative lesions of the adrenal, kidney, prostate, and testes. J Toxicol Environ Health Part A. 58:199–214.

[WHO IPCS] World Health Organisation – International Programme on Chemical Safety. 2004. IPCS risk assessment terminology. Part 1: IPCS/OECD key generic terms used in chemical hazard/risk assessment. Part 2. IPCS glossary of key exposure assessment terminology. IPCS harmonisation project document No. 1. Geneva: WHO. [accessed 2020 Feb]. http://www.inchem.org/documents/harmproj/harmproj/harmproj1.pdf.

WHO IPCS. 2010. WHO human health risk assessment toolkit: chemical hazards. IPCS harmonisation project document no. 8. Geneva: WHO. [accessed 2020 Feb]. https://apps.who.int/iris/bitstream/handle/10665/44458/9789241548076_eng.pdf;jsessionid=12D6B92111A2F3FF2B46D3B923FC84493?sequence=1.

Wilkinson CF, Killeen JC. 1996. A mechanistic interpretation of the oncogenicity of chlorothalonil in rodents and an assessment of human relevance. Regul Toxicol Pharmacol. 24:69–84.

Witchi H. 2005. A/J mouse as a model for lung tumorigenesis caused by tobacco smoke: strengths and weaknesses. Exp Lung Res. 31:3–18.

Wolf DC, Cohen SM, Boobis AR, Dellarco VL, Fenner-Crisp PA, Moretto A, Pastoor TP, Schoeny RS, Seed JG, Doe JE. 2019. Chemical

carcinogenicity revisited 1: a unified theory of carcinogenicity based on contemporary knowledge. Regul Toxicol Pharmacol. 103: 86–92.

Woutersen RA, Soffers AE, Kroese ED, Krul CA, van der Laan JW, van Benthem J, Luijten M. 2016. Prediction of carcinogenic potential of chemicals using repeated-dose (13-week) toxicity data. Regul Toxicol Pharmacol. 81:242–249.

Woutersen M, Beekman M, Pronk MEJ, Muller A, de Knecht JA, Hakkert BC. 2019. Does REACH provide sufficient information to regulate mutagenic and carcinogenic substances? Hum Ecolog Risk Asses Int J. 25:1996–2016.