Opinion Article

Towards an Animal-Free Human Health Assessment: Starting from the Current Regulatory Needs

Peter M. J. Bos[1], Liesbeth Geraets[1], Lianne de Wit-Bos[1], Minne Heringa[1,2] and Jacqueline van Engelen[1]

[1]National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands; [2]present address: Reckitt Benckiser, Hygiene Home, Amsterdam, The Netherlands

Abstract

Societal concern for animal welfare and scientific concerns about the predictive power of animal models for the human situation are driving forces for the development of animal-free approaches for the safety testing of chemicals. A paradigm shift towards an assessment of human health risks that is fully based on animal-free approaches is not foreseen within the next decades. To accelerate the use of animal-free innovations (*in vitro* experiments, *in silico* methods, etc.) in the EU, it has multiple advantages to simultaneously work towards a new risk assessment paradigm and to aim the development of animal-free methods at better meeting current regulatory needs. To achieve this, a multi-stakeholder collaboration is needed already in the development phase, where regulators can inform on the regulatory needs and the criteria for acceptance. As a first step, the present paper discusses what information is needed within the context of four areas of chemical safety assessment in the EU: 1) classification, labelling and packaging, 2) the derivation of health-based guidance values and product limits, 3) risk assessments of exposure situations of concern and 4) addressing specific topics of societal concern. Further agreements on the level of detail and uncertainty, robustness, predictive value, reproducibility and validation are a prerequisite to develop tools that can be trusted and that will be legally binding.

1 Introduction

In everyday life, humans come into contact with a variety of chemicals that have a wide range of uses and applications, including food additives, food contact materials, industrial chemicals, plant protection products (PPP), biocides and cosmetics. For many decades, the safety of human exposures has been systematically evaluated by studying potential intrinsic adverse health effects of chemicals in animal experiments and by extrapolating this information to human exposure scenarios and/or by determining risk management measures to protect humans from the intrinsic hazardous properties.

However, already in 1959, Russell and Burch drew attention to the ethical aspects of animal experiments and introduced the concept of the three Rs, replacement, reduction and refinement of animal experiments (Russell and Burch, 1959). Although initially slowly acknowledged, the societal concern for animal welfare has significantly grown in the past decades. At the same time, there is an increasing concern about the use of animal experiments for

scientific reasons, i.e., whether animals are appropriate models to predict adverse health effects in humans (Schiffelers et al., 2007). Multiple reasons underlie this concern. The biology and physiology of rats and mice, the animal species used most commonly for testing, may differ significantly from that of humans. In addition, human endpoints such as a general ill-feeling, headache, psycho-somatic and learning effects are not covered by the endpoints generally investigated in animal experiments. A third scientific reason also meets present societal concerns and more or less follows from the previous reason. It relates to the question: Can chemicals in our daily environment be responsible for human diseases such as Alzheimer's disease, Parkinson's disease, obesity or attention deficit hyperactivity disorder (ADHD), some of which are accompanied by high costs for human society?

The animal experiments as required and traditionally performed in accordance with the OECD Guidelines for the Testing of Chemicals or the EU Test Methods are, for various reasons, only designed to cover certain predetermined endpoints, and are neither able nor developed to cover all potential relevant human

diseases. These experiments are very resource-intensive in terms of costs and man-hours, which limits, for instance, the number of animals and thereby the statistical power to detect adverse effects. Further, the suitability of the animal species tested as a predictive model for potential human health effects is seldom verified beforehand.

In addition, development of innovative nanomaterials and the transition towards a circular economy (e.g., green chemistry or non-toxic environments) require adequate, cheap and rapid methods for toxicity screening. To enable industries to implement safe and sustainable-by-design approaches, new methodologies are needed to assess safety and sustainability throughout the entire life cycle but that do not require extensive data gathering and toxicity testing.

For these reasons, initiatives from both a scientific as well as a societal drive to reduce the number of animal experiments are emerging with the final aim to fully assess the safe use of chemicals using animal-free tests and approaches (NCad, 2016). This is reflected in the increasing research on animal-free approaches, the development of new strategies consisting of animal-free methods and tools, based mainly on *in silico* models and *in vitro* tests, and the development of strategic roadmaps for the transition towards an animal-free approach (Adeleye et al., 2015; Berggren et al., 2017; Blaauboer et al., 2016; Daneshian et al., 2016; Embry et al., 2014; FDA, 2017; Gocht et al., 2015; ICCVAM, 2018; Krewski et al., 2010).

It is often stated that the current legislative frameworks do not accept animal-free or *in vitro* data. It is true that for the assessment of risks of chemical substances for man and environment, animal studies are commonly performed. Alternatives to these studies may only be applied if the legal frameworks for the assessments explicitly offer a possibility for this. In a report for the Dutch Ministry of Economic Affairs, RIVM has analyzed ten European frameworks for assessment of chemical substances to see whether the possibility to use alternatives is hindered. The investigation showed that, in a strict sense, there are no legal barriers (Heringa et al., 2014). However, the absence of legal barriers in most frameworks does not mean the road is free to apply alternatives. This is mainly because of the lack of validated methods to predict complex endpoints such as repeated-dose toxicity, reproductive toxicity, etc., as validation and regulatory (e.g., OECD) acceptance of methods is required by all these regulations.

Another important issue is that currently developed methods may not be sufficiently suitable to assess the safety of chemicals, i.e., they do not always sufficiently address the regulatory questions to be answered in a safety assessment. In relation to the safety assessment of chemicals, a large number of different questions can be posed. For instance: Is this chemical an endocrine disruptor? Is this chemical harmful for children or newborns? What is the effect of a combination of pesticides present on fruit? At what dose will this chemical have an effect? The question: Does this chemical cause cancer? might imply a number of sub-ques-

tions: Does this chemical cause cancer in a specific exposure situation (via skin contact or via inhalation; route-specific toxicity and dose-response relationship)? Does this chemical cause cancer in combination with other chemicals (cocktail effect)? Does this chemical cause cancer in an individual (susceptibility)? Therefore, it is important to specify the information that is needed and be aware of the specific safety assessment question that needs to be answered.

In this present publication, we focus on the current regulatory needs, i.e., the current frameworks for which information on the effects of chemicals on human health needs to be generated to assure their safe production, application and use.

Depending on the regulatory framework and situation with the substance (e.g., already on the market or not), the questions above may or may not need to be answered. Anyhow, it is important to realize that answering the different regulatory questions requires different pieces of information and thus may require specific tests or testing strategies. We have distinguished four types of regulatory questions:

1) What are the hazardous properties of the chemical? E.g.: Is it a sensitizer? Does it cause cancer? (see Section 2.1)
2) What is a safe level of exposure, so safe use of the chemical can be assured? (see Section 2.2)
3) What effects can be expected when people have been exposed to a level higher than the safe level? (see Section 2.3)
4) Does a certain chemical cause a certain disease such as, e.g., Alzheimer's or Parkinson's disease, that is not covered directly by the hazardous properties of Question 1?

It will, therefore, be described what regulatory safety (or risk) assessment actually entails, discussed what the current information requirements are from the viewpoint of the regulatory context in the EU, and from there, what this will mean for the focus and development of animal-free innovations, rather than discussing the innovations themselves. Sometimes a qualitative outcome may suffice, while in other situations specific strategies combining multiple tools are needed to achieve a quantitative outcome. These aspects will be of importance if the current regulatory needs are to be met.

This paper, therefore, intends to stimulate the discussion on how to bridge the gap between the regulatory needs for safety or risk assessment and to provide input for the development of robust and fit-for-purpose animal-free innovations and thereby accelerate the use of animal-free tools in regulatory human safety assessments in the short term.

2 Current assessment of human health risks in a regulatory context in the EU: information needs

In the EU, regulatory bodies are involved in human health risk assessments at multiple levels. The safety evaluations may regard the whole supply chain (manufacture, distribution, storage,

---

Abbreviations

AOP, adverse outcome pathway; BLV, Biological Limit Value; BMD, benchmark dose; BMC, benchmark concentration; CMR, carcinogenicity, germ cell mutagenicity or reproductive toxicity; CLP, Classification, Labelling and Packaging; HBGV, Health-Based Guidance Value; PPP, plant protection products; QIVIVE, Quantitative *In vitro - In vivo* Extrapolation

use via recycling/circular use up to waste disposal) and may be performed *a priori*, i.e., before a chemical is introduced on the market, whether or not in a new application, or *a posteriori*, i.e., for a current exposure situation for which a concern has arisen. Depending on the use of a chemical, market authorization is evaluated within respective regulatory frameworks. Committees or panels of agencies such as ECHA and EFSA evaluate potential human health risks. At a Member State level, national authorities may request a health risk assessment of specific exposure situations of concern. The type and detail of information required for such an assessment will depend on the regulatory context of the evaluation.

For the purpose of the development of new methods and strategies to replace animal experiments in a regulatory context, the question to be answered is why animal experiments are being performed in the first place, i.e., what information is to be obtained from them and how will the information be used. This must be clearly understood so that animal-free innovations may be developed in such a way that the data they provide is acceptable in regulatory processes.

Animal data used *a priori* are meant to provide information on potential human health effects such that a safe production and use of a chemical can be assured. In this setting, animal data serve the purpose of a safety assessment before a chemical is introduced on the market. The amount of data required, and thus the type of animal experiments, depends on the applicable regulatory framework. An extensive toxicity profile is required for, e.g., PPP and biocidal products, but in other frameworks the data requirements are less extensive and may depend on the level of production or use, or on the expected level of exposure.

Animal data may also be used *a posteriori* to evaluate whether a current exposure scenario for which a concern has arisen might result in human health risks. *A posteriori* risk assessments of exposure scenarios of societal concern generally are within the compass of governmental responsibility. These are scenarios for situations arising from a concern for either high exposure (or a different route of exposure) to a chemical or a specific hazard endpoint, or both. A risk assessment then is performed to answer the question whether the exposure scenario introduces (or has introduced) a human health risk and, if so, what health risks can be expected, to what extent in terms of type and severity of effects, and what the size and nature of the affected population is. Finally, it will be evaluated what measures need to be taken to reduce exposure. Recent examples include acrylamide and the rubber infill in artificial grass pitches (ECHA, 2017a; RIVM, 2017; Svensson et al., 2003; Tareke et al., 2002; US EPA/ATSDR/CPSC, 2016). In practice, such a risk assessment needs to be performed only for a relatively small number of chemicals as compared to the tens of thousands of chemicals present in the environment. However, these are scenarios that often receive the highest media attention and create high societal concern. In these situations, there generally is no time to generate new data and to perform a risk assessment; one has to rely on information that is already available for the chemical in question.

Often a tiered approach is followed, with data requirements increasing with anticipated exposure. One reason for such an approach is the aspect of proportionality of data requirements in relation to the level of production and/or use. Data requirements may also be restricted because of the limitations of an animal experiment. For instance, not all relevant human endpoints can be studied in traditional toxicological animal experiments. It is therefore noted that the official data *requirements* in a framework may be less than the actual information *need* for a complete and adequate safety assessment from a scientific point of view.

In the next sections, it will be discussed what information presently generated using animal experiments is needed within the context of an *a priori* regulatory safety assessment, i.e., for Classification, Labelling and Packaging (CLP) and for the derivation of a Health Based Guidance Value (HBGV). The focus in this publication will be on the information that is needed within the legal frameworks in the EU, which is somewhat different from, for example, the regulatory situation in the United States (Busquet and Hartung, 2017; Krewski et al., 2010). Subsequently, the possibilities of animal-free tools for the purpose of an *a posteriori* risk assessment will be pointed out, as well as the need to investigate the possibilities to address topics of current societal concern.

2.1 *A priori* safety assessment: classification of chemicals (CLP)

Within the EU, the mandatory hazard classification of chemicals and mixtures being manufactured, imported, used or distributed within the EU is laid down in the Classification, Labelling and Packaging (CLP) Regulation (EC) 1272/2008 (EC, 2008). Currently, the harmonized classification and labelling of approximately 4000 chemicals is included in this regulation. The criteria of the international UN Globally Harmonised System of Classification and Labelling of Chemicals (GHS) are incorporated in the EU CLP Regulation. Classified chemicals are labelled with appropriate Hazard (H-) and Precautionary (P-) statements. H-statements relate to specific hazards and P-statements deal with actions to prevent exposure, actions to respond to exposure, with storage conditions and with disposal. These statements aim to warn of specific hazards and to limit or reduce exposure, often without knowledge and/or consideration of what a safe level of exposure actually might be.

By mandating the classification and labelling of chemical substances and mixtures, this regulation guarantees that employees and consumers in the EU have clear information about the hazards of chemicals (EC, 2008). Chemicals proposed to be classified for carcinogenicity, germ cell mutagenicity or reproductive toxicity (CMR) categories 1A, 1B or 2, or for respiratory sensitization, or chemicals that are active substances in the meaning of Regulation EC 1107/2009 (i.e., PPP) or Regulation (EU) No 528/2012 (biocidal products) shall normally be subject to harmonized classification and labelling.

Currently, hazard evaluation under CLP is based on available data that is generated within the various legal frameworks (such as PPPs, biocidal products, and REACH) and performed per endpoint. CLP therefore does not require additional testing (animal or animal-free) for the purpose of classification for human health hazards (ECHA, 2017b), and no conclusion is drawn for an endpoint for which no (sufficient) information has been generated in the various legal frameworks. Depending on the hazardous proper-

Tab. 1: Overview of *in vitro* tests for each toxicological endpoint that are currently recognized and accepted for classification purposes according to the CLP-Regulation

Endpoint	*In vitro* test
Acute toxicity	
Skin irritation	OECD TG 439 (reconstructed human epidermis test method)[a]
Skin corrosion	OECD TG 430 (transcutaneous electrical resistance test) OECD TG 431 (reconstructed human epidermis test) OECD TG 435 (*in vitro* membrane barrier test)
Eye irritation	-
Eye damage	OECD TG 437 (bovine corneal opacity and permeability test) OECD TG 438 (isolated chicken eye test) OECD TG 460 (fluorescein leakage test) OECD TG 491 (short time exposure test) OECD TG 492 (reconstructed human cornea-like epithelium test)
Skin sensitization	OECD TG 442C/D/E (tests for peptide/protein binding, keratinocyte response, and monocytic/dendritic cell response, respectively)[b]
Respiratory sensitization	-
Mutagenicity[c]	OECD TG 471 bacterial reverse mutation test OECD TG 473 *in vitro* mammalian chromosomal aberration test OECD TG 476, 490 *in vitro* mammalian cell gene mutation test OECD TG 487 *in vitro* mammalian cell micronucleus test
Repeated dose toxicity	-
Carcinogenicity	-
Reproductive toxicity: - Effects on fertility and sexual function - Effects on development	 -

[a] This test method cannot distinguish between corrosives and irritants when used alone. Thus, in the case of positive results, the potential corrosive properties should be excluded or confirmed based on data obtained from an *in vitro* skin corrosion test (ECHA, 2017b).
[b] These tests are not regarded as stand-alone tests, but may be used in a weight of evidence approach. These should not be used currently to sub-categorize skin sensitizers into subcategories 1A and 1B (ECHA, 2017b).
[c] Positive *in vitro* tests must be accompanied with a chemical structure similarity to known germ cell mutagens to be accepted for a classification, and then only a category 2 classification can be obtained.

ties of the chemical, a restriction or ban might apply for the whole supply chain, i.e., from manufacture, distribution, storage, and use up to waste disposal. The classification of chemicals therefore also triggers risk reduction measures in other legislations, as many European and national legislations are directed at the safe use of hazardous chemicals. Examples are legislations concerning consumer products (cosmetics, toys, cleaning products), the protection of workers, transport safety and external safety.

The regulatory need for this type of safety assessment is thus to obtain information on the *intrinsic* hazards of chemicals. While the majority of the tests as currently presented for evaluation of human health hazard assessment concerns *in vivo* tests, *in vitro* tests are accepted and used for a limited number of endpoints as stand-alone for hazard classification. Table 1 presents an overview of *in vitro* tests that are currently recognized and regulatorily accepted for classification purposes. In this context, it is noteworthy that *in vitro* tests for skin corrosion and irritation have been introduced at the level of UN GHS in 2019 (United Nations,

2019a) and that an *in vitro* test for eye irritation presently is under discussion (United Nations, 2019b). Furthermore, non-testing approaches such as grouping and QSARs are also accepted. As can be seen from Table 1, classification based on non-animal tests is now common practice for some endpoints, but not for the more complex, systemic endpoints.

Classification on intrinsic hazardous properties is performed in a qualitative way for most endpoints (e.g., yes/no toxic to reproduction, yes/no carcinogenic), with the exception of the endpoints Specific Target Organ Toxicity Single Exposure (STOT-SE)/Repeated Exposure (STOT-RE), for which quantitative exposure boundaries apply. What thus is needed when moving towards classification based on animal-free tools, is a data-generating strategy that aims to assess whether a chemical has the *potential* to induce a specific health effect. Such a strategy will consist of a number of integrated tools that, in combination, provide information for classification. The basic steps are depicted in Figure 1.
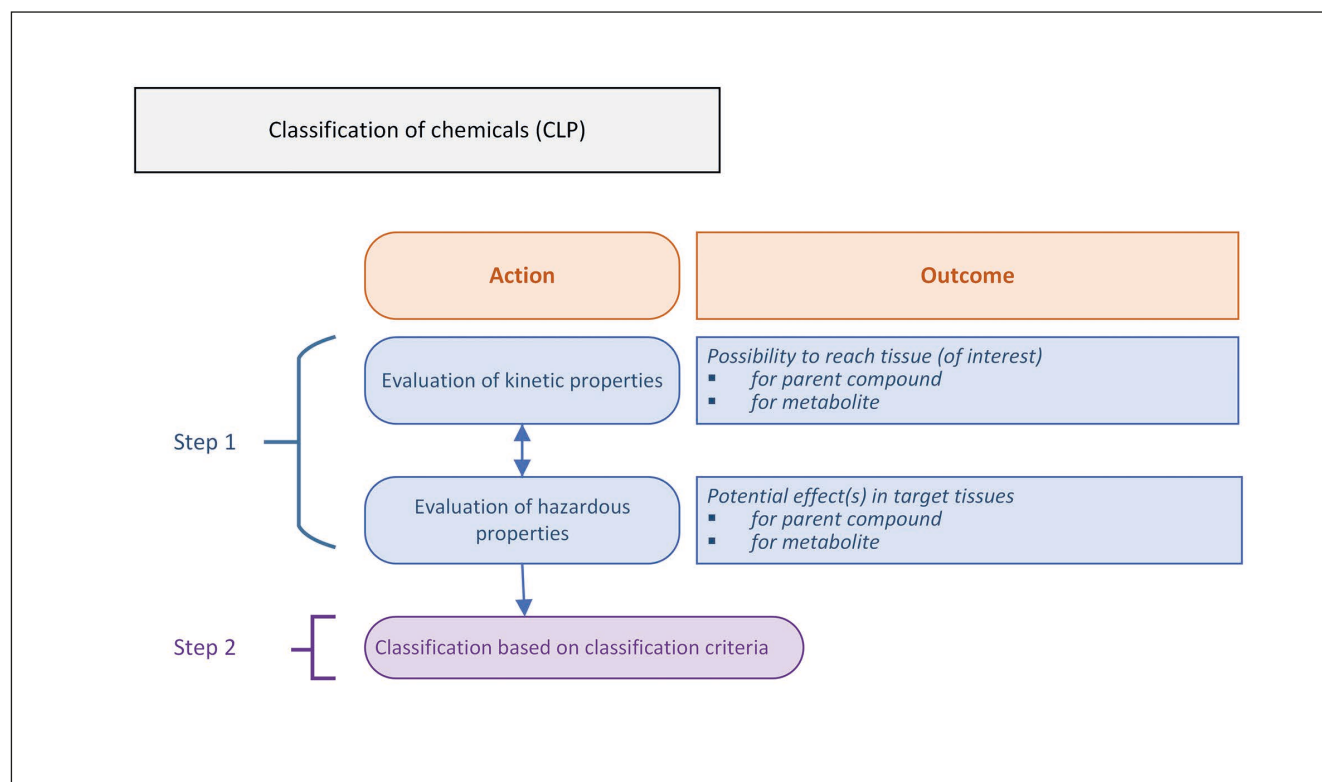
Fig. 1: Stepwise approach for generating data to meet the information needs for classification, labelling and packaging (CLP)

For local effects, information on kinetics generally plays no or only a limited role. In case of a systemic effect, information on kinetics gains importance and appropriate kinetic tools need to become part of the strategy. For the demonstration of a potential intrinsic effect of the chemical, it is important to know what the potential toxic agent will be (parent compound or metabolite) and whether the toxic agent reaches the target tissue(s) (absorption, distribution). For this purpose, kinetic and hazardous properties need to be generated in an integrated and coherent way. The actual exposure conditions in humans are not of importance. Quantitative extrapolation to a realistic human exposure scenario is not necessary for classification, which may call for less strict demands on the tools to be developed than when quantitative *in vitro - in vivo* extrapolation (QIVIVE) is required.

The present classification categories based on *in vivo* animal experiments are generally linked to a specific severity of an effect; boundaries have been defined for assignment into these categories. Obviously, when animal experiments are replaced by (a strategy of) animal-free innovations, these classification boundaries and criteria may need to be redefined, based on the endpoint observed in the new strategy or method. For instance, where skin corrosion criteria for the *in vivo* acute dermal irritation/corrosion test (OECD TG 404) are based on visible destruction of skin tissue in test animals, the criteria for *in vitro* methods are based on, e.g., the integrity of the skin barrier function (OECD TG 430 and 435) or on tissue viability (OECD TG 431) (United Nations, 2019a).

2.2 *A priori* safety assessment: derivation of guidance values and limit values

EU regulatory frameworks in which safety evaluations are performed for chemicals (or products) include REACH, biocidal products, PPPs, novel foods, feed additives, food improvement agents such as enzymes, additives and flavorings, food contact materials, veterinary medicinal products and cosmetics. The main aim of the safety evaluation in (most of) these frameworks is to derive an exposure limit, such as an HBGV or a product limit, which is used to control exposure in order to ensure the safe production, manufacturing and use of a chemical. Although the regulatory information requirements differ between the frameworks, they share the same purpose, i.e., establishing conditions for safe production and/or use. Currently, the establishment of an HBGV is (mostly) based on a reference point for the critical effect, being a no observed adverse effect level (NOAEL) or benchmark dose (BMD) or benchmark concentration (BMC) from an animal experiment, derived by applying assessment factors. In an *in vivo* experiment with intact animals, all physiological processes, including kinetics, and selected endpoints are implicitly addressed.

If guidance values or limit values are to be derived via an animal-free approach, then what information is needed for their derivation and how can this information be obtained? What elements should be included in a strategy for the derivation of HBGVs or product limits? Generally speaking, as a minimum prerequisite, information on the kinetics of a chemical and the critical effect(s) is needed. Kinetic information is needed to assess the absorption
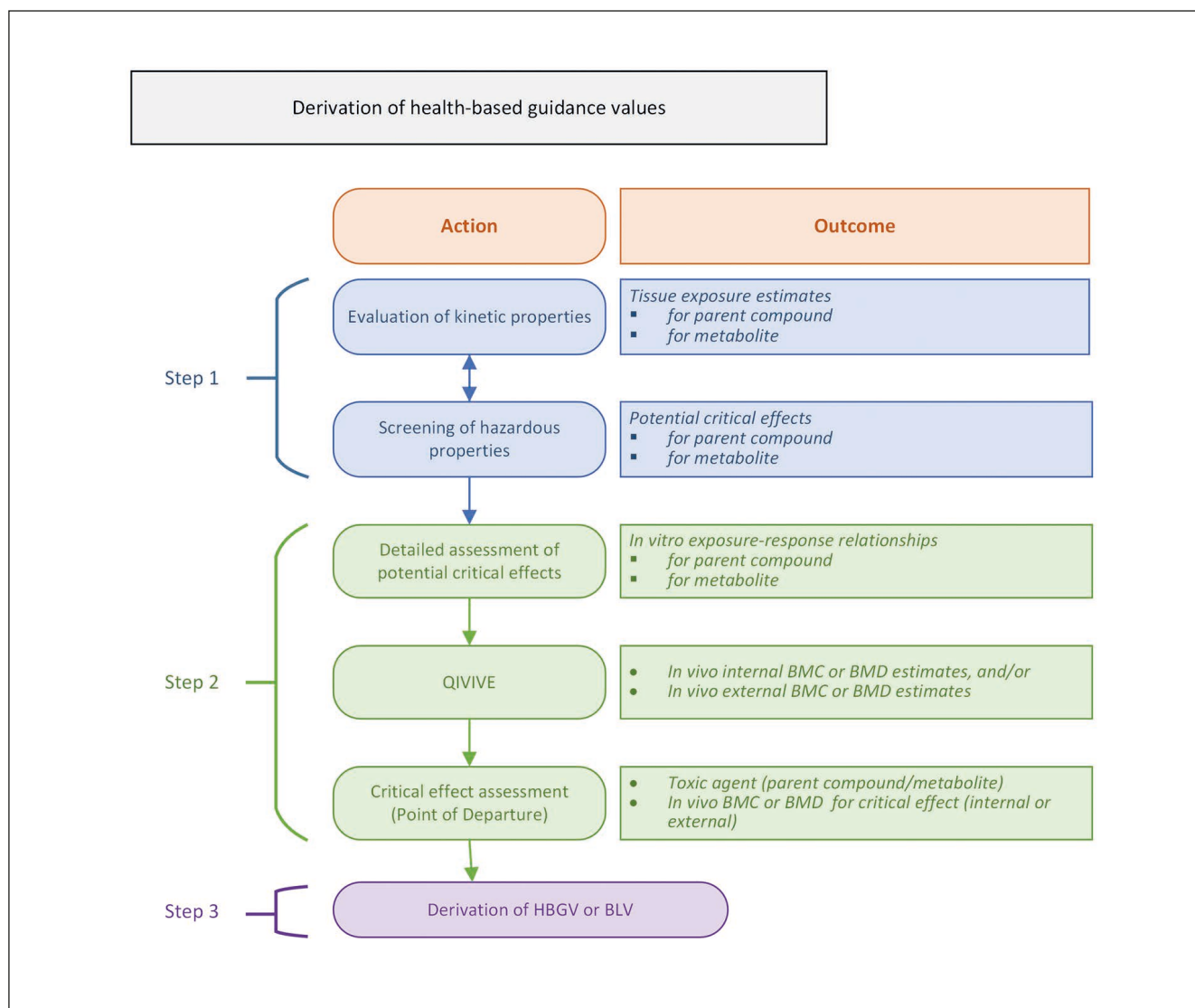
Fig. 2: Stepwise approach for generating data to meet the information needs for derivation of health-based guidance values (HGBV)

of the chemical, to identify the potential toxic agent (parent compound or metabolite), and to determine to what extent organs and tissues are exposed to the toxic agent. In contrast to the information needed for CLP, the kinetic information and the occurrence of the critical effect (exposure-response) need to be determined quantitatively.

The assessment of an HBGV or limit value via an animal-free approach may, for instance, follow a stepwise strategy as depicted in Figure 2. The first step can be considered a screening step in which, based on kinetic information and the screening of hazardous properties, potential critical effect(s) can be identified for the most toxic agent (parent compound or metabolite). To achieve this, a screening approach is needed in which information on kinetics (i.e., potential for a possible toxic agent to reach specific tissues) and hazardous endpoints can be generated in an integrat-

ed way. This is in fact already to some extent in line with the current approach used for genotoxicity testing, where three complementary *in vitro* assays are required, at least in a first tier, and each is performed in the absence and presence of metabolizing liver enzymes (S9). The screening step for derivation of an HBGV may, for instance, consist of a battery of tools through which information can be generated for multiple potential endpoints. This step needs to be designed such that potential hazards will not be missed. By including knowledge about physiological and biological processes that may lead to adverse effects, such as for example the adverse outcome pathway (AOP) network, one or more effects may be identified that can be considered a potential critical endpoint for HBGV derivation.

In a second step, these identified critical effects may be the subject of a more detailed and quantitative evaluation by ade-

quate *in vitro* assays and/or *in silico* models to establish exposure-response curves from which internal BMCs or BMDs can be derived. Also, the generation of hazard information needs to be conducted in combination with that of kinetic information. Kinetic information can be used to optimize the exposure conditions in the test designs to bring them in line with realistic human exposure scenarios and to determine an adequate dose metric.

For the derivation of an HBGV, the internal dose or concentration that represents a threshold for a potential critical effect needs to be extrapolated to an external exposure level by QIVIVE. It may also be extrapolated to an *in vivo* internal blood concentration or other relevant biomarker if guidance values, for instance a biological limit value (BLV), are designed for human biomonitoring. The determination of the most critical effect to serve as point of departure for an HBGV should be based on a comparison of the *in vivo* BMCs or BMDs (external or internal) with each other. In a third and final step, an HBGV or BLV is derived from the most relevant and critical BMC or BMD.

It is acknowledged that in principle such an approach requires that adequate animal-free tools are available for multiple relevant endpoints, but the advantage is that a chemical only needs to be tested in detail for the identified critical endpoints. Obviously, CMR properties need to be included in the screening step for critical endpoints. However, if CMR properties are a reason for non-authorization, a qualitative assessment may suffice.

The current practice is to address uncertainties in the derivation of guidance values, *inter alia* the extrapolation from animals to humans, by applying assessment factors. Similarly, uncertainties resulting from the use of an animal-free approach (e.g., the use of cell or organ systems instead of an intact organism or the extrapolation by QIVIVE) can be taken into account by applying assessment factors. Correction for a high source of uncertainty due to interspecies differences does not apply here if information obtained with human cells or tissues can be used.

Over the past few years, several stepwise or tiered approaches have been suggested to evaluate potential health risks or to perform a safety assessment for chemicals within specific regulatory frameworks (e.g., Adeleye et al., 2015; Berggren et al., 2017; Blaauboer et al., 2016; Dent et al., 2018; Desprez et al., 2018; Embry et al., 2014; Pastoor et al., 2014). These strategies serve different purposes and do not (always) focus on the derivation of HBGVs but are mentioned here to illustrate how animal-free strategies and methods can be combined in a tiered approach to generate answers to specific regulatory questions.

It is important to be aware that frameworks that do not require data depend on the data obtained within frameworks that do require data. An example is exposure values used in emergency response planning or land-use planning. Another example is the derivation of occupational exposure levels (OELs), although it is acknowledged that within REACH, derived no-effect levels (DNELs) for the worker are also determined. Therefore, it must be kept in mind that the information generated for the derivation of HBGVs for chemicals within a specific regulatory framework should be sufficient to derive occupational exposure levels and perform emergency response and land-use planning as well.

2.3 *A posteriori* risk assessment of exposure situations of concern

As mentioned before, *a posteriori* risk assessments for the evaluation of human health risks in a historic, current or emerging exposure scenario of substances for which a concern has arisen are performed for a relatively small number of chemicals. The exposure can be the main trigger to perform such a risk assessment, e.g., in case of a fire or leakage at an industrial site or in case a contaminant is detected in food, such as the recent acrylamide case (Tareke et al., 2002; van Bruggen et al., 2010). Another example may be unrecognized high emissions of a chemical during a long time period. A trigger for a full risk assessment may also come from public concern about the hazardous properties of a chemical compound, e.g., the use of rubber infill in artificial football pitches (RIVM, 2017).

In case the need for a full risk assessment is caused by concern about exposure, knowledge about the critical effect of a chemical and/or a safe level no longer suffices. People who are exposed want to know what health risks they can expect from that exposure. Therefore, quantitative information on the exposure-response relationship is needed, not only for the critical effect but also for other effects that might occur at higher exposure levels. Such an evaluation requires detailed knowledge on the toxicity profile of a chemical, preferably in relation to the specific characteristics of the exposure scenario in question, e.g., route of exposure, frequency, duration, and concentration/dose. When a public concern arises about specific hazardous properties of a chemical, in principal, a complete insight into the toxicity profile is not necessarily needed. Quantitative information about the exposure-response relationship for the hazardous endpoint(s) of interest suffices. Since exposure has already occurred (in contrast to most *a priori* risk assessments), available epidemiological data can provide an important additional source of information.

Whether the high level of exposure or the public concern triggers the risk assessment, there usually is only limited time to come up with an answer and one has to rely on the available data. The steps that need to be taken are reflected in Step 1 in Figure 3. The characteristics of the exposure scenario of interest are described in detail to determine and search for the appropriate kinetic and hazard information that fits best to address potential health risks resulting from that scenario. If the available information is insufficient for risk assessment, information gaps are identified. Generation of additional information to fill these gaps (Step 2) can only be performed if this is possible within a short time frame (and if sufficient budget is available), depending on the urgency to address the concern. Animal experiments are then not the appropriate way forward; especially as animal studies with repeated exposures are time-consuming.

Thus, animal-free innovations are needed to generate data about potential effects that might occur in the exposure scenario of concern such that potential effects in relation to the exposure scenario of interest can be evaluated and appropriate exposure-effect relationships can be assessed. In these situations, *in silico* tools (including PBPK-models) may also gain importance, even more than *in vitro* assays, as the latter will also require time to perform, to analyze and to interpret, and even more so, when
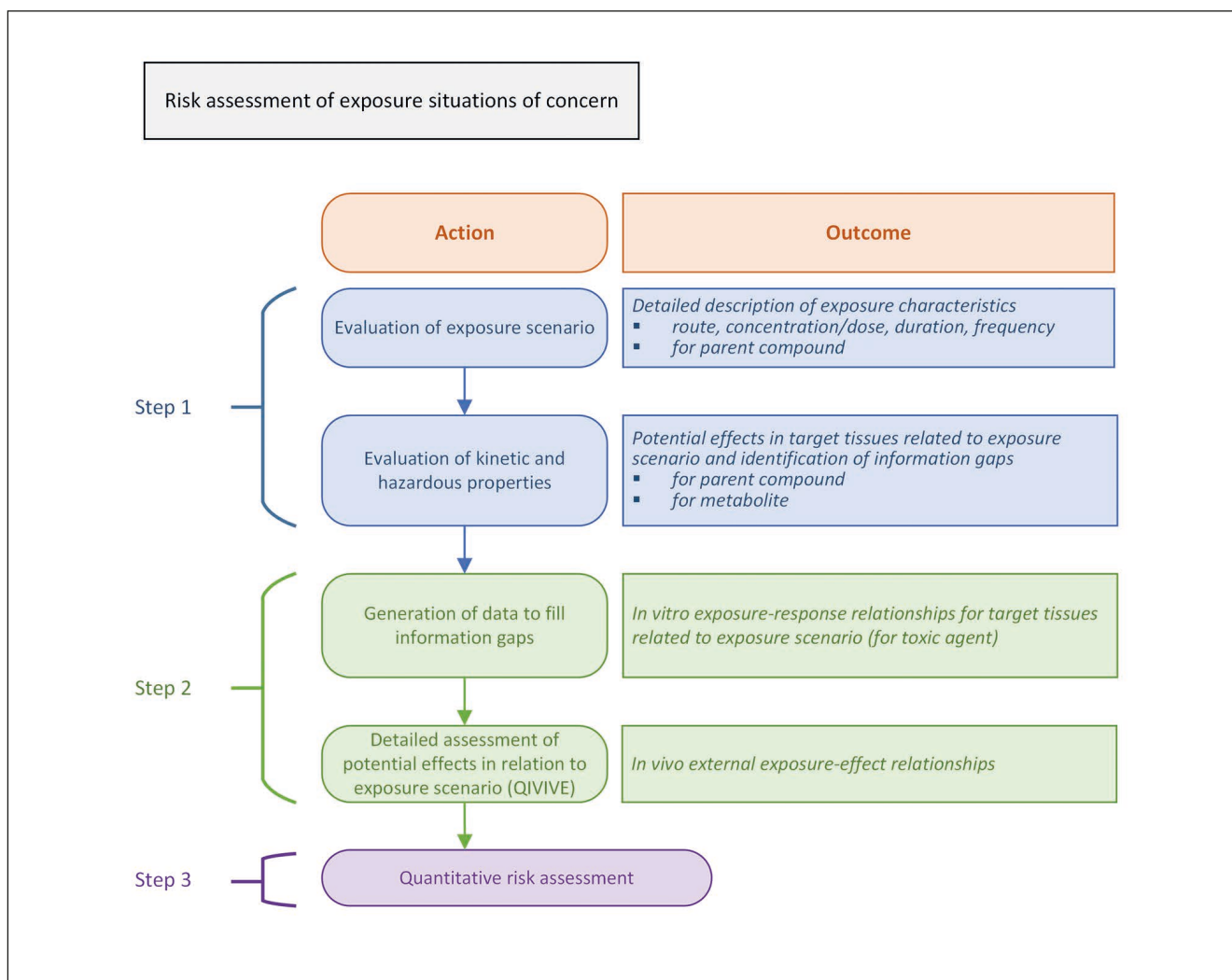
Fig. 3: Stepwise approach for generating data to meet the information needs for risk assessments of exposure situations of concern

in silico tools can be developed that are sufficiently flexible to mimic the exposure conditions under evaluation, thereby making it possible to apply available information directly to the exposure scenario for which the risk assessment is performed. Development of such flexible tools will be of great advantage for risk assessment since the exposure conditions of an emerging situation (e.g., route of exposure, frequency, duration and concentration) often cannot be foreseen. Finally, a quantitative risk assessment can be performed for the exposed populations in Step 3.

Thus, to serve the purpose of risk assessments of past, occurring or emerging exposures, it would be favorable if tools were developed that can aid a risk assessment by quickly generating new information for a chemical or by supporting the extrapolation of available data to the specific exposure scenario under evaluation. Incorporation of PBPK-modelling already has proven its value in several risk assessments (Andersen, 2003; Bos et al., 2006; Clewell et al., 2005; Mielke and Gundert-Remy, 2012). However, the development of these types of models requires a large amount of good-quality data from *in vitro* and/or *in vivo* tests.

2.4 Addressing specific topics of societal concern

A large potential benefit of animal-free approaches is that specific topics can be addressed that cannot be or cannot be fully tackled by animal experiments. Especially topics that are presently the subject of societal discussions may benefit, such as whether exposure to chemicals may be related to the development of contemporary diseases such as diabetes, cardiovascular diseases, behavioral effects such as attention deficit hyperactivity disorder (ADHD) or autism, Alzheimer's disease, etc. (Bellanger et al., 2015; Gwinn et al., 2017). Animal models currently used for regulatory purposes do not provide information on these endpoints.

In addition, also questions with regard to effects of combined exposures to (mixtures of) multiple chemicals are not sufficiently addressed by animal studies (Gwinn et al., 2017). The current
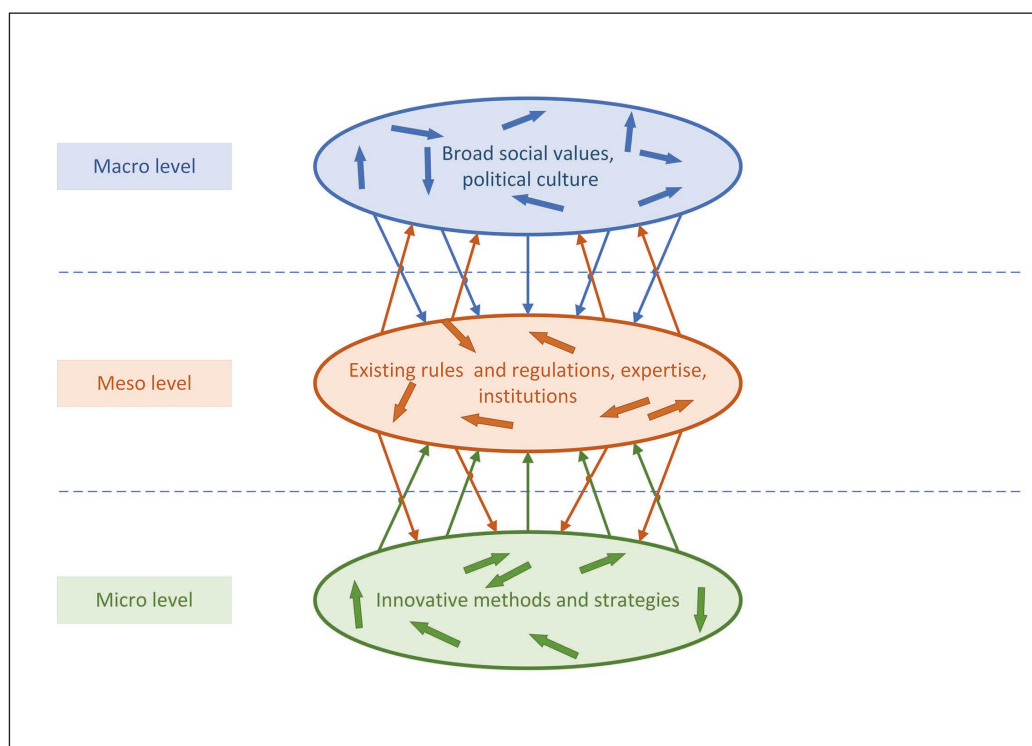
Fig. 4: Multilevel perspective on the transition towards a health safety or risk assessment based on animal-free innovations

safety or risk assessment paradigm based predominantly on animal experiments is oriented towards the health effects of single chemicals, whereas in practice the human body is exposed to a cocktail of chemicals.

These topics favor a system in which human-relevant cell systems for new endpoints and human-relevant technologies are used. Lessons can be learned from advances in human biomedical research and disease models. In addition, combinations of *in vitro* and *in silico* tools are needed in which the mixture effect of many different combinations of chemicals can be studied. This information need will not be easily met. However, from a societal and also from a health-economic point of view, it is worthwhile to explore the possibilities of new methods and methodologies to include this information in human safety and risk assessment (Pistollato et al., 2016).

3 Implementation of animal-free innovations in the EU: next steps

Societal concerns for reasons of animal welfare as well as scientific data gaps and concerns are driving forces for the development of animal-free innovations. Use of animal-free approaches potentially can improve the safety and risk assessment for certain scenarios where animal studies currently are not requested or conducted for various reasons. In addition, animal-free approaches, including appropriate *in silico* tools, may provide the flexibility to better address specific human exposure scenarios in practice. Tools based on human physiology and biology and

the use of human tissues and cells may improve the relevance of the outcome for the assessment of human health risks (Burden et al., 2015). The ideal picture would be that developments of animal-free innovations will finally result in an approach that makes it possible to determine a complete profile of toxic effects and of human diseases for every chemical coming to or already on the market, with acceptable costs in terms of resources and time.

While the final timelines and costs are still unclear, developments based on, e.g., adverse outcome pathways (AOPs), ontology mapping, virtual organs, and body-on-a-chip systems may ultimately lead to approaches that provide such toxicity profiles. An assessment of human health risks based on such approaches will require a shift towards a new safety assessment paradigm, one that is based on human biology and physiology rather than the present evaluation of individual endpoints in animal experiments (Marx et al., 2016; Staal et al., 2017). Interesting and challenging initiatives and proposals are already being presented with "the virtual human" as the ultimate remote prospect (Piersma et al., 2019). New information on chemical toxicity could be derived by artificial intelligence mining of existing data sets, combined with information from the new technologies mentioned above (Hartung, 2019).

The shift to a new safety assessment based on human biology and physiology will require a regulatory acceptance process that is rather complex. Schiffelers et al. (2014) have applied the multilevel perspective on technology transitions (Geels, 2002) to describe the influences needed to induce regulatory acceptance at the micro, meso and macro levels (Fig. 4). Innovations take place at the micro level, where new methods are developed, test-

ed and validated. The meso level entails the current situation or regime, where institutions with existing rules and regulations find their place and influence the acceptance of innovations. Societal features, such as the social attitude towards animal welfare and also the political culture, are present at the macro level. Developments at the macro level, e.g., concern for animal welfare and about the predictive power of animal models, and at the micro level, e.g., increasing availability and applicability of animal-free innovations, are important drivers to stimulate the necessary changes at the meso level for regulatory acceptance. However, discussions are needed at and between these three levels to assure a broad acceptance of safety and risk assessment based on animal-free innovations. Discussion topics include the uncertainty about the value of new animal-free innovations (micro level), the lack of harmonization of regulatory requirements and acceptance criteria (meso level), and the high levels of risk aversion and the concern for animal welfare (macro level) (Schiffelers et al., 2014). In optimizing the process, critical drivers are commitment, communication, cooperation and co-ordination.

However, despite the huge efforts undertaken, it is not realistic to expect that such an approach will be possible for all chemicals entering the market, at least not in the foreseeable future (Sewell et al., 2017). It will take decades to achieve a paradigm shift towards a safety assessment fully based on animal-free approaches.

However, it may turn out to be disadvantageous for the development of animal-free innovations if significant steps in the replacement of animal experiments by animal-free innovations cannot be taken in the meantime, i.e., if the implementation of these innovations is postponed until it is possible to generate a complete, quantitative profile of human health risks. The momentum towards a transition may be lost if the practical advantages of animal-free approaches for an adequate assessment of human health risks do not become clear in the short term. At the moment, much effort is put into the development of new methods (micro level). At the same time, application of animal-free innovations in the current safety and risk assessments may provide the proof-of-principle for new developments. Therefore, next to the developments towards a new safety or risk assessment paradigm, implementation of animal-free approaches in the current human safety assessments should be stimulated.

There is desire both of the developers of new tools and technologies (micro level) and of society urging the reduction of animal testing (macro level) to speed up the process of human risk assessment based on human biology and physiology. This is mainly the role of regulatory implementation and acceptance (meso level), for which co-ordination and collaboration at all three levels (micro, meso and macro) is key.

3.1 Need for co-ordination and collaboration
For an innovation to be successfully taken up in regulatory safety testing, it is important to start from a foreseen need and to determine the kind of information and the level of detail required for that need rather than to first develop and finalize a sophisticated tool before seeking where it can be best applied. The needs for safety assessments (e.g., CLP, derivation of HGBVs and product

limits) and risk assessments of exposure situations of concern are discussed above.

Depending on the regulatory context, the needs clearly vary, and different criteria are applicable to the level of detail and of uncertainty of the data to be generated. Qualitative information may suffice for one purpose, while a quantitative evaluation such as QIVIVE may be required for other purposes. Especially when based on animal-free innovations, fulfilment of a regulatory need will often require information that is obtained in different scientific fields and thus requires a multi-disciplinary approach. At least for systemic endpoints, new tools for hazard evaluation and kinetics need to be developed in coherence. For instance, if hazard evaluation appears to be optimally generated by a combination of different techniques, such as omics, computational tools and/or *in vitro* assays, the question is how these are to be combined and how and where information on kinetics needs to be included. The tools to be developed should be complementary and compatible.

To stimulate the introduction of animal-free innovations for the assessment of human health risks, it is therefore of utmost importance to intensify the dialogue between the different stakeholders involved. What is then needed, is the development of a safety assessment strategy describing the information needs to answer a specific regulatory question, the steps to be taken, and the criteria that should be met. To achieve such an integrated assessment strategy that is fit-for-purpose, a multi-disciplinary collaboration, including regulators, scientists, and representatives from industry, is preferred already at the developmental stage of animal-free innovations (Benfenati et al., 2016; Burden et al., 2015; Meek and Lipscomb, 2015; Sewell et al., 2017). In this, there is an important role in store for regulatory officers, risk assessors and risk managers to clearly formulate the regulatory questions to be answered and to define the criteria involved (Bell et al., 2018; Punt et al., 2018). Scientists can put forward the most recent developments and their possibilities. It can be decided in consultation how the regulatory needs can best be addressed such that compatible tools can be developed and criteria be defined that should be met for the respective regulatory goals.

Working in collaboration and combining regulatory needs and technical possibilities will provide the best assurance for implementation and application of animal-free innovations within the field of regulatory safety testing within the foreseeable future. Involvement of regulatory officers already in the development phase of animal-free innovations will further contribute to the creation of broad support for acceptance of these innovations and to strategies that will be fit for the purpose of safety and risk assessment and that are trusted, accepted and legally binding.

3.2 Acceptance and implementation of animal-free innovations
Whether implementation of animal-free innovations for *a priori* safety assessments in regulatory frameworks or for *a posteriori* risk assessment of chemicals is successful, will depend on several factors: First of all, there should be no legal boundaries in the European frameworks for the use of animal-free innovations. It has been determined that this is in principle indeed the case (Heringa et al., 2014). An important hurdle for regulatory

acceptance rather appears to be a lack of reproducible and robust methods for the more complex endpoints (Sewell et al., 2017). For instance, to sufficiently address health effects from repeated exposures, a testing strategy consisting of a combination of advanced methods is required to identify critical effects and to determine the toxicity profile of a chemical. Unless such a strategy encompasses a sufficient number of relevant endpoints, an animal experiment will still be performed instead. Secondly, a fundamental change in mind-set is necessary for a successful shift from an assessment where animal testing is still seen as the gold standard towards a safety assessment that is completely based on animal-free innovations (Schiffelers et al., 2007; Sewell et al., 2017). Gathering data from animal experiments for several endpoints has been the way forward for many decades, and the experience over the past decades with methods based on animal models has created a certain degree of trust in animal data and has provided at least a certain feeling of safety. It has also created a basis for making legal decisions, and new methods that are to be implemented should cover at least the same legal basis and acceptance as the ones they are replacing. This has created some reserve towards implementation of animal-free approaches in safety and risk assessment. We know what we have and to blaze a new and uncertain trail may lead to a road of trial and error, which is considered not to be the best way forward when human health is concerned.

It is acknowledged that there are uncertainties involved in using animal experiments for human safety or risk assessment; These uncertainties are generally accounted for by applying assessment factors. Similarly, the use of animal-free innovations will also be accompanied with uncertainties, but these will be other uncertainties with which we have little or no experience. Starting the introduction of animal-free innovations in a stepwise approach provides the opportunity to gain practical experience, and thus trust in the predictive power of robust animal-free innovations. At the same time, it can be ensured that they meet the requirements of the regulatory frameworks and the needs of regulators and risk assessors. Gradual implementation of practical, straightforward, and fit-for-purpose animal-free approaches also provides possibilities for interim adjustments and improvements where considered necessary.

Additional issues of importance that need to be addressed during the development of animal-free innovations include determination of an acceptable level of uncertainty, robustness, reproducibility and transparency. To create trust, the issues regarding uncertainty, robustness, reproducibility, predictive power and transparency need to be agreed upon and accepted by the respective stakeholders. Combining building blocks of complex animal-free innovations, such as omics, organs-on-a-chip, computational tools, ontology mapping, etc., into an approach that meets a specific regulatory need, will require a multidisciplinary approach (Benfenati et al., 2016). Such building blocks may be difficult to understand for someone who is not an expert in the multiple scientific fields involved but nevertheless has to evaluate these building blocks individually and in conjunction and has to judge their merits. It is therefore of importance that end-users of these tools, such as regulatory officers, risk assessors and risk

managers are acquainted with the new technologies and in addition are involved already during their development. This further supports that new developments, in combination with the mentioned requirements they should meet, need to be established via a dialogue of multiple stakeholders (ECHA, 2018; Punt et al., 2018).

Another important issue for the implementation of animal-free innovations is that the validation process needs to be reconsidered (Burgdorf et al., 2019; NAS, 2017; Piersma et al., 2018; Tralau et al., 2015). Animal experiments are considered to be the gold standard for validation, but this is in contradiction with one of the scientific reasons to replace animal experiments: They may not always be the appropriate model for human safety and risk assessment. In addition, when new topics such as new endpoints or combined or aggregated exposures are introduced, the amount of presently available animal data may be too limited to support their validation. Therefore, development of animal-free innovations should go hand in hand with research on how to validate these innovations and to develop appropriate methods for validation.

A possibility to meet the reserve and concerns towards introduction of animal-free innovations for safety evaluation is the introduction of a trusted assessment environment, or a safe harbor, where an open dialogue between industry and regulatory authorities on the safety assessment of a chemical or a product, based on data obtained with animal-free methods, can be encouraged.

Another possibility is the use of an appropriate measure of post-marketing control of exposure as a safety net. This can be achieved by the introduction of BLVs and the set-up of health surveillance systems. New animal-free approaches will require more detailed biological and kinetic information. This will provide the opportunity to derive adequate BLVs that will be better exposure estimates than external HBGVs, especially in case of exposure via multiple routes and/or from multiple sources. BLVs describe the actual internal exposure. Setting up surveillance systems to regularly measure these BLVs and perform basic health assessments will help to control the exposure. Furthermore, the information thus obtained can in its turn be used to evaluate the adequacy of the safety evaluation. This may help to fill in data gaps and to fine-tune the safety assessment.

4 Summary and conclusions

Development of animal-free innovations that are directly designed to meet the present regulatory needs has multiple advantages (Burden et al., 2015). This may contribute to a reduction of animal use in the short term. Instead of an uncontrolled proliferation of new animal-free innovations, the development of animal-free innovations will be focused on what is really needed from a regulatory point of view, leading to fit-for-purpose tools. This will advance the possibilities for their use within regulatory safety assessments. Agreements among the respective stakeholders on the level of detail and uncertainty, robustness, predictive value, reproducibility and method for validation are prerequisites to arrive at tools that can be trusted and that will be legally bind-

ing (Bell et al., 2018; Punt et al., 2018). To achieve this, clear agreements are necessary on the criteria that the tools must meet as well as on the interpretation of the outcome of the tools and the resulting conclusions, decisions and consequences.

To identify the regulatory information needs, different areas of safety or risk assessment were discussed that presently rely on data from animal experiments, including CLP and derivation of HBGVs and product limits. In the EU, the majority of chemicals are currently regulated by minimizing exposure via CLP, by restricting the manufacture, distribution, use and storage of a chemical and/or the derivation of HBGVs. As third and fourth areas, risk assessment of human exposure scenarios of concern and addressing societal issues were identified, respectively.

In principle, *CLP* is based on the intrinsic hazard of a chemical combined with agreements on a categorization based on the test results. Simply put, a mutual agreement on, for example, a combination of individual (qualitative) *in vitro* and/or *in silico* hazard and kinetic assays with information about human physiology and biology and the establishment of new or adapted classification criteria may be sufficient for the classification of chemicals under an animal-free evaluation system.

For the derivation of *HBGVs and product limits*, quantitative information is required about the kinetics of a compound and its critical effect (including the dose-response relationship to derive an appropriate point of departure). An approach then may be the development of a screening strategy to first identify the most toxic agent(s) and the most probable critical endpoints. The screening step may subsequently be followed by a more detailed quantitative evaluation of the selected endpoints to estimate a quantitative relationship between an appropriate target dose metric and the respective health effects. This relationship can be translated to an external or internal estimate of safe human exposure by QIVIVE. In a third step, an HBGV or BLV may then be derived from this safe exposure estimate. Each of the three steps will require different tools (*in vitro* kinetics, *in vivo* kinetics, *in silico*) with different levels of quantification and certainty.

The consequence of animal-free approaches in safety assessment may be that limited data become available for an *a posteriori quantitative risk assessment* in case that, for instance, an HBGV is exceeded. However, it should be realized that present HBGVs are not always based on a complete toxicity database either. A complete toxicity profile is seldom available and often data gaps are identified. Risk assessments of emerging exposure situations may be subject to constraints of time. It would then be favorable if animal-free innovations were developed that could quickly generate data, especially if these innovations were sufficiently flexible to address a variety of exposure scenarios.

The fourth area regards the inclusion of specific topics in the assessment of *human health risks that until now cannot or cannot fully be addressed by animal experiments*. Animal-free innovations may be developed to evaluate whether chemicals may cause a human disease, thereby improving the safety assessment process. The assessment of human health risks from combined exposures to chemicals often is limited to chemicals with a similar mode of action, such as the group of organophosphates and the inhibition of cholinesterase. Also, exposure to well-known mixtures such as mineral oil or white spirit has been the subject of toxicity studies. But exposure to combinations of chemicals also includes other areas of interaction, such as induction or inhibition of metabolizing enzymes, which may affect the activation or deactivation of other chemicals. Animal-free innovations enabling faster data generation (at reasonable costs) could allow assessments of combined exposures.

A multi-stakeholder dialogue is needed to develop animal-free tools that meet the present regulatory needs. Therefore, close involvement of regulators in the development of animal-free innovations is a prerequisite to assure that the regulatory needs are clearly defined and that criteria for a robust assessment of human health following chemical exposure will be met (Benfenati et al., 2016; ECHA, 2018; Meek and Lipscomb, 2015; Punt et al., 2018). This dialogue should also address additional issues, including how to create trust and confidence in assessments based on animal-free innovations, agreement upon how to deal with uncertainties, and the level of robustness, reproducibility and transparency needed. Experience and thus trust can be gained with animal-free tools and models aimed at the present regulatory needs, which can be brought forward in the development of more complex innovations.

Therefore, a gradual implementation may aid the phasing-out of animal testing and pave the road for more complex and sophisticated innovations. This may accelerate the developments of animal-free innovations as well as acceptance of these innovations in human safety or risk assessment in a regulatory context.

References

Adeleye, Y., Andersen, M., Clewell, R. et al. (2015). Implementing toxicity testing in the 21st century (TT21C): Making safety decisions using toxicity pathways, and progress in a prototype risk assessment. *Toxicology 332*, 102-111. doi:10.1016/j.tox.2014.02.007

Andersen, M. E. (2003). Toxicokinetic modeling and its applications in chemical risk assessment. *Toxicol Lett 138*, 9-27. doi:10.1016/S0378-4274(02)00375-2

Bell, S. M., Chang, X., Wambaugh, J. F. et al. (2018). In vitro to in vivo extrapolation for high throughput prioritization and decision making. *Toxicol In Vitro 47*, 213-227. doi:10.1016/j.tiv.2017.11.016

Bellanger, M., Demeneix, B., Grandjean, P. et al. (2015). Neurobehavioral deficits, diseases, and associated costs of exposure to endocrine-disrupting chemicals in the European Union. *J Clin Endocrinol Metab 100*, 1256-1266. doi:10.1210/jc.2014-4323

Benfenati, E., Berggren, E., Fritsche, E. et al. (2016). Novel chemical hazard characterisation approaches. *EFSA J 14*, e00506-n/a. doi:10.2903/j.efsa.2016.s0506

Berggren, E., White, A., Ouedraogo, G. et al. (2017). Ab initio chemical safety assessment: A workflow based on exposure considerations and non-animal methods. *Comput Toxicol 4*, 31-44. doi:10.1016/j.comtox.2017.10.001

Blaauboer, B. J., Boobis, A. R., Bradford, B. et al. (2016). Considering new methodologies in strategies for safety assessment

of foods and food ingredients. *Food Chem Toxicol 91*, 19-35. doi:10.1016/j.fct.2016.02.019

Bos, P. M., Zeilmaker, M. J. and van Eijkeren, J. C. (2006). Application of physiologically based pharmacokinetic modeling in setting acute exposure guideline levels for methylene chloride. *Toxicol Sci 91*, 576-585. doi:10.1093/toxsci/kfj176

Burden, N., Mahony, C., Muller, B. P. et al. (2015). Aligning the 3Rs with new paradigms in the safety assessment of chemicals. *Toxicology 330*, 62-66. doi:10.1016/j.tox.2015.01.014

Burgdorf, T., Piersma, A. H., Landsiedel, R. et al. (2019). Workshop on the validation and regulatory acceptance of innovative 3R approaches in regulatory toxicology – Evolution versus revolution. *Toxicol In Vitro 59*, 1-11. doi:10.1016/j.tiv.2019.03.039

Busquet, F. and Hartung, T. (2017). The need for strategic development of safety sciences. *ALTEX 34*, 3-21. doi:10.14573/altex.1701031

Clewell, H. J., Gentry, P. R., Kester, J. E. et al. (2005). Evaluation of physiologically based pharmacokinetic models in risk assessment: An example with perchloroethylene. *Crit Rev Toxicol 35*, 413-433. doi:10.1080/10408440590931994

Daneshian, M., Kamp, H., Hengstler, J. et al. (2016). Highlight report: Launch of a large integrated European in vitro toxicology project: EU-ToxRisk. *Arch Toxicol 90*, 1021-1024. doi:10.1007/s00204-016-1698-7

Dent, M., Amaral, R. T., Da Silva, P. A. et al. (2018). Principles underpinning the use of new methodologies in the risk assessment of cosmetic ingredients. *Comput Toxicol 7*, 20-26. doi:10.1016/j.comtox.2018.06.001

Desprez, B., Dent, M., Keller, D. et al. (2018). A strategy for systemic toxicity assessment based on non-animal approaches: The cosmetics Europe long range science strategy programme. *Toxicol In Vitro 50*, 137-146. doi:10.1016/j.tiv.2018.02.017

EC (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 L 353/1*. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32008R1272&from=EN

ECHA (2017a). Annex XV report. An evaluation of the possible health risks of recycled rubber granules used as infill in synthetic turf sports fields. European Chemicals Agency (ECHA). Version number: 1.01. Date: 28 February 2017. https://echa.europa.eu/documents/10162/13563/annex-xv_report_rubber_granules_en.pdf/dbcb4ee6-1c65-af35-7a18-f6ac1ac29fe4

ECHA (2017b). 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. European Chemicals Agency (ECHA). Version 5.0, July 2017. https://echa.europa.eu/documents/10162/23036412/clp_en.pdf/58b5dc6d-ac2a-4910-9702-e9e1f5051cc5

ECHA (2018). More progress needed to replace animal tests. European Chemical Agency (ECHA), Newsletter February 2018, Issue 1. https://newsletter.echa.europa.eu/home/-/newsletter/entry/more-progress-needed-to-replace-animal-tests

Embry, M. R., Bachman, A. N., Bell, D. R. et al. (2014). Risk assessment in the 21st century: Roadmap and matrix. *Crit Rev Toxicol 44, Suppl 3*, 6-16. doi:10.3109/10408444.2014.931924

FDA (2017). FDA'S Predictive Toxicology Roadmap. U.S. Food and Drug Administration. https://www.fda.gov/downloads/ScienceResearch/SpecialTopics/RegulatoryScience/UCM587831.pdf

Geels, F. W. (2002). Technological transitions as evolutionary reconfiguration processes: A multi-level perspective and a case-study. *Res Policy 31*, 1257-1274. doi:10.1016/S0048-7333(02)00062-8

Gocht, T., Berggren, E., Ahr, H. J. et al. (2015). The SEURAT-1 approach towards animal free human safety assessment. *ALTEX 32*, 9-24. doi:10.14573/altex.1408041

Gwinn, M. R., Axelrad, D. A., Bahadori, T. et al. (2017). Chemical risk assessment: Traditional vs public health perspectives. *Am J Public Health 107*, 1032-1039. doi:10.2105/AJPH.2017.303771

Hartung, T. (2019). Opinion: AI beats animal testing at finding toxic chemicals. *The Scientist*. https://www.the-scientist.com/critic-at-large/opinion--ai-beats-animal-testing-at-finding-toxic-chemicals-65795

Heringa, M. B., de Wit, L., Bos, P. M. J. et al. (2014). Do current EU regulations for the safety assessment of chemical substances pose legal barriers for the use of alternatives to animal testing? RIVM Letter report 2014-0148. National Institute for Public Health and the Environment. Bilthoven, The Netherlands. https://www.rivm.nl/bibliotheek/rapporten/2014-0148.pdf

ICCVAM (2018). A Strategic Roadmap for Establishing New Approaches to Evaluate the Safety of Chemicals and Medical Products in the United States. Interagency Coordinating Committee on the Validation of Alternative Methods. doi:10.22427/NTP-ICCVAM-ROADMAP2018

Krewski, D., Acosta, D., Jr., Andersen, M. et al. (2010). Toxicity testing in the 21st century: A vision and a strategy. *J Toxicol Environ Health B Crit Rev 13*, 51-138. doi:10.1080/10937404.2010.483176

Marx, U., Andersson, T. B., Bahinski, A. et al. (2016). Biology-inspired microphysiological system approaches to solve the prediction dilemma of substance testing. *ALTEX 33*, 272-321. doi:10.14573/altex.1603161

Meek, M. E. and Lipscomb, J. C. (2015). Gaining acceptance for the use of in vitro toxicity assays and QIVIVE in regulatory risk assessment. *Toxicology 332*, 112-123. doi:10.1016/j.tox.2015.01.010

Mielke, H. and Gundert-Remy, U. (2012). Physiologically based toxicokinetic modelling as a tool to support risk assessment: Three case studies. *J Toxicol 2012*, 359471. doi:10.1155/2012/359471

NAS (2017). Using 21st century science to improve risk-related evaluations. Committee on Incorporating 21st Century Science into Risk-Based Evaluations. National Academy of Sciences. http://dels.nas.edu/Report/Using-21st-Century-Science-Improve/24635

NCad (2016). Transitie naar proefdiervrij onderzoek over mogelijkheden voor het uitfaseren van dierproeven en het stimuleren van proefdiervrije innovatie. Advies van het Nationaal Comité advies dierproevenbeleid (NCad) in opdracht van de staatssecretaris van Economische Zaken (in Dutch with English summary). https://www.ncadierproevenbeleid.nl/documenten/rapport/2016/12/15/ncad-advies-transitie-naar-proefdiervrij-onderzoek

Pastoor, T. P., Bachman, A. N., Bell, D. R. et al. (2014). A 21st century roadmap for human health risk assessment. *Crit Rev Toxicol 44, Suppl 3*, 1-5. doi:10.3109/10408444.2014.931923

Piersma, A. H., van Benthem, J., Ezendam, J. et al. (2018). Validation redefined. *Toxicol In Vitro 46*, 163-165. doi:10.1016/j.tiv.2017.10.013

Piersma, A. H., van Benthem, J., Ezendam, J. et al. (2019). The virtual human in chemical safety assessment. *Curr Opin Toxicol 15*, 26-32. doi:10.1016/j.cotox.2019.03.009

Pistollato, F., Ohayon, E. L., Lam, A. et al. (2016). Alzheimer disease research in the 21st century: Past and current failures, new perspectives and funding priorities. *Oncotarget 7*, 38999-39016. doi:10.18632/oncotarget.9175

Punt, A., Bouwmeester, H., Schiffelers, M.-J. W. A. et al. (2018). Expert opinions on the acceptance of alternative methods in food safety evaluations: Formulating recommendations to increase acceptance of non-animal methods for kinetics. *Regul Toxicol Pharmacol 92*, 145-151. doi:10.1016/j.yrtph.2017.11.015

RIVM (2017). Evaluation of health risks of playing sports on synthetic turf pitches with rubber granulate. Scientific background document. De Groot, G. M. and Oomen, A. G. (eds.), RIVM Report 2017-017. http://www.rivm.nl/bibliotheek/rapporten/2017-0017.pdf

Russell, W. M. S. and Burch, R. L. (1959). *The Principles of Humane Experimental Technique*. London, UK: Methuen. ISBN 0900767782

Schiffelers, M. J., Blaauboer, B. J., Fentener van Vlissingen, J. M. et al. (2007). Factors stimulating or obstructing the implementation of the 3Rs in the regulatory process. *ALTEX 24*, 271-278. doi:10.14573/altex.2007.4.271

Schiffelers, M. J., Blaauboer, B. J., Bakker, W. E. et al. (2014). Regulatory acceptance and use of 3R models for pharmaceuticals and chemicals: Expert opinions on the state of affairs and the way forward. *Regul Toxicol Pharmacol 69*, 41-48. doi:10.1016/j.yrtph.2014.02.007

Sewell, F., Doe, J., Gellatly, N. et al. (2017). Steps towards the international regulatory acceptance of non-animal methodology in safety assessment. *Regul Toxicol Pharmacol 89*, 50-56. doi:10.1016/j.yrtph.2017.07.001

Staal, Y., Pennings, J., Hessel, E. et al. (2017). Advanced toxicological risk assessment by implementation of ontologies operationalized in computational models. *Appl In Vitro Toxicol 3*, 325-332. doi:10.1089/aivt.2017.0019

Svensson, K., Abramsson, L., Becker, W. et al. (2003). Dietary intake of acrylamide in Sweden. *Food Chem Toxicol 41*, 1581-1586. doi:10.1016/S0278-6915(03)00188-1

Tareke, E., Rydberg, P., Karlsson, P. et al. (2002). Analysis of acrylamide, a carcinogen formed in heated foodstuffs. *J Agric Food Chem 50*, 4998-5006. doi:10.1021/jf020302f

Tralau, T., Oelgeschlager, M., Gurtler, R. et al. (2015). Regulatory toxicology in the twenty-first century: Challenges, perspectives and possible solutions. *Arch Toxicol 89*, 823-850. doi:10.1007/s00204-015-1510-0

United Nations (2019a). Globally Harmonized System of Classification and Labeling of Chemicals (GHS). Eighth revised edition. ST/SG/AC.10/30/Rev.8. https://www.unece.org/fileadmin/DAM/trans/danger/publi/ghs/ghs_rev08/ST-SG-AC10-30-Rev8e.pdf

United Nations (2019b). Use of non-animal testing methods for classification of health hazards: Status report. Sub-Committee of Experts on the Globally Harmonized System of Classification and Labelling of Chemicals. UN/SCEGHS/38/INF.20. http://www.unece.org/fileadmin/DAM/trans/doc/2019/dgac10c4/UN-SCEGHS-38-INF20e.pdf

US EPA/ATSDR/CPSC (2016). Federal Research Action Plan on Recycled Tire Crumb Used on Playing Fields and Playgrounds https://www.epa.gov/chemical-research/federal-research-action-plan-recycled-tire-crumb-used-playing-fields

van Bruggen, M., Bodar, C. W. M., Bos, P. M. J. et al. (2010). Risk assessment of ethylene oxide emission in Zoetermeer. RIVM Report 609300017. National Institute of Public Health and the Environment, Bilthoven, The Netherlands. https://www.rivm.nl/bibliotheek/rapporten/609300017.pdf

Conflict of interest

The authors declare that they have no conflicts of interest.

Acknowledgements