SCIENTIFIC OPINION

EFSA Journal

ADOPTED: 10 November 2021

doi: 10.2903/j.efsa.2021.6970

Scientific Opinion of the Scientific Panel on Plant Protection Products and their Residues (PPR Panel) on testing and interpretation of comparative *in vitro* metabolism studies

EFSA Panel on Plant Protection Products and their Residues (EFSA PPR Panel), Antonio F Hernandez-Jerez, Paulien Adriaanse, Annette Aldrich, Philippe Berny, Tamara Coja, Sabine Duquesne, Andreas Focks, Marina Marinovich, Maurice Millet, Olavi Pelkonen, Silvia Pieper, Aaldrik Tiktak, Christopher J Topping, Anneli Widenfalk, Martin Wilks, Gerrit Wolterink, Ursula Gundert-Remy, Jochem Louisse, Serge Rudaz, Emanuela Testai, Alfonso Lostia, Jean-Lou Dorne and Juan Manuel Parra Morte

Abstract

EFSA asked the Panel on Plant Protection Products and their residues to deliver a Scientific Opinion on testing and interpretation of comparative *in vitro* metabolism studies for both new active substances and existing ones. The main aim of comparative *in vitro* metabolism studies of pesticide active substances is to evaluate whether all significant metabolites formed in the human *in vitro* test system, as a surrogate of the *in vivo* situation, are also present at comparable level in animal species tested in toxicological studies and, therefore, if their potential toxicity has been appropriately covered by animal studies. The studies may also help to decide which animal model, with regard to a particular compound, is the most relevant for humans. In the experimental strategy, primary hepatocytes in suspension or culture are recommended since hepatocytes are considered the most representative *in vitro* system for prediction of *in vivo* metabolites. The experimental design of $3 \times 3 \times 3$ (concentrations, time points, technical replicates, on pooled hepatocytes) will maximise the chance to identify unique (UHM) and disproportionate (DHM) human metabolites. When DHM and UHM are being assessed, test item-related radioactivity recovery and metabolite profile are the most important parameters. Subsequently, structural characterisation of the assigned metabolites is performed with appropriate analytical techniques. In toxicological assessment of metabolites, the uncertainty factor approach is the first alternative to testing option, followed by new approach methodologies (QSAR, read-across, *in vitro* methods), and only if these fail, *in vivo* animal toxicity studies may be performed. Knowledge of *in vitro* metabolites in human and animal hepatocytes would enable toxicological evaluation of all metabolites of concern, and, furthermore, add useful pieces of information for detection and evaluation of metabolites in different matrices (crops, livestock, environment), improve biomonitoring efforts via better toxicokinetic understanding, and ultimately, develop regulatory schemes employing physiologically based or physiology-mimicking *in silico* and/or *in vitro* test systems to anticipate the exposure of humans to potentially hazardous substances in plant protection products.

Keywords: xenobiotic, reactive metabolites, interspecies metabolism, suspension/plated hepatocytes, clearance, PBK, QSAR, in silico

Requestor: European Food Safety Authority

Question number: EFSA-Q-2020-00551

Correspondence: pesticides.ppr@efsa.europa.eu

Panel members: Antonio F Hernandez-Jerez, Paulien Adriaanse, Annette Aldrich, Philippe Berny, Tamara Coja, Sabine Duquesne, Andreas Focks, Marina Marinovich, Maurice Millet, Olavi Pelkonen, Silvia Pieper, Aaldrik Tiktak, Christopher J Topping, Anneli Widenfalk, Martin Wilks and Gerrit Wolterink.

Declarations of interest: The declarations of interest of all scientific experts active in EFSA's work are available at https://ess.efsa.europa.eu/doi/doiweb/doisearch.

Acknowledgments: This opinion is dedicated to the beautiful memory of Dr Alfonso Lostia (1982–2020). The Panel wishes to thank the following hearing experts for the views provided to this scientific output: Ugo Zanelli, Paul Whalley; Carsten Kneuer, Matthew Himmelstein; the observer Sandra Coecke (Directorate General - Joint Research Centre); the EFSA staff Martina Panzarea, Mathilde Colas.

Suggested citation: EFSA PPR Panel (EFSA Panel on Plant Protection Products and their Residues), Hernandez-Jerez AF, Adriaanse P, Aldrich A, Berny P, Coja T, Duquesne S, Focks A, Marinovich M, Millet M, Pelkonen O, Pieper S, Tiktak A, Topping CJ, Widenfalk A, Wilks M, Wolterink G, Gundert-Remy U, Louisse J, Rudaz S, Testai E, Lostia A, Dorne J-L and Parra Morte JM, 2021. Scientific Opinion of the Scientific Panel on Plant Protection Products and their Residues (PPR Panel) on testing and interpretation of comparative *in vitro* metabolism studies. EFSA Journal 2021;19(12):6970, 61 pp. https://doi.org/10.2903/j.efsa.2021.6970

ISSN: 1831-4732

The EFSA Journal is a publication of the European Food Safety Authority, a European agency funded by the European Union.

Table of contents

1. Introduction

1.1. Background and Terms of Reference as provided by EFSA

The requirement to perform *in vitro* comparative metabolism studies for pesticide active substances related to human health is set by Commission Regulation (EU) No 283/2013 setting out the data requirements for active substances, in accordance to Regulation (EC) No 1107/2009 of the European Parliament and of the Council concerning the placing of plant protection products on the market, which establishes in Section 5 "Toxicological and metabolism studies" that "Comparative *in vitro* metabolism studies shall be performed on animal species to be used in pivotal studies and on human material (microsomes or intact cell systems) in order to determine the relevance of the toxicological animal data and to guide in the interpretation of findings and in further definition of the testing strategy. An explanation shall be given or further tests shall be carried out where a metabolite is detected *in vitro* in human material and not in the tested animal species".

In addition, the same Regulation defines that "The relevance of generating toxicity data in animal models with dissimilar metabolic profiles to those found in humans shall be addressed, if such metabolic information is available, and taken into consideration for study design and risk assessment".

The scope of *in vitro* comparative metabolism studies is to assess potential interspecies differences in metabolism of a pesticide active substance. The aim is to ensure that all metabolites formed by the *in vitro* human material (e.g., human hepatocytes) are also present in, and thus covered by, one or more of the animal species used in toxicological studies. When a metabolite is present only in human material "an explanation shall be given or further tests shall be carried out" to assess its toxicological relevance. The same applies for metabolites in considerable excess in human material as compared to animals when deemed essential from toxicity point of view.

Currently no validated test methods are available to conduct these studies. Consequently, in the context of the toxicological dataset submitted for the approval or renewal of pesticide active substance, these studies are either conducted using different experimental layouts or in some cases they are not provided. Therefore, this poses some challenges for the national authorities and EFSA when evaluating the studies received by the applicants during the peer review process of pesticide active substance.

In vitro comparative metabolism studies are widely used for pharmaceuticals in various stages of development, from a new chemical entity characterization to a pre-clinical safety assessment and this knowledge could support the use of these studies in the pesticide context (FDA, 2020). However, it became clear in the context of the EU pesticide peer review that further guidance on the minimum requirements in conducting *in vitro* comparative metabolism studies on pesticide active substances and assessing related data would be needed (Outcome of the pesticides peer review meeting on general recurring issues in mammalian toxicology, EFSA, 2016).

In addition, following implementation of Regulation (EU) 283/2013, biomonitoring methods for human body fluids and tissues are required. In the absence of a minimal set of data on interspecies differences in metabolism, it could be very complex to address the question of using the metabolism data in rat as a surrogate of metabolism in human for biomonitoring methods.

In order to set a common scientific background among all the stakeholders involved in the peer review process (i.e., Applicants, European Commission, Member States, Pesticide Steering Network), the EFSA Pesticide Unit organised a workshop in 2018 on comparative *in vitro* metabolism studies in regulatory pesticide risk assessment. An event report on the outcome of the workshop was published in 2019 (EFSA, 2019a). The workshop indicated the need of guidance on how to perform these studies. The discussion held was used as foundation stone for developing the current project.

The project was therefore started with the aim to develop an EFSA Guidance illustrating the testing strategy that should be applied to assess interspecies comparative *in vitro* metabolism, the minimum requirements that should be met in the testing protocol for the selected assays and interpretation of the study results.

During the activity, the European Commission was informed about the status of the Guidance and commented that it would be beneficial to address not only the problem formulation related to the identification of human metabolites not properly assessed by toxicological studies using laboratory animal species, but also how to assess the identified human metabolites from the toxicological point of view. Therefore, to address the second problem formulation (evaluation of toxicological relevance of human metabolites not properly assessed by toxicological animal studies), the PREV Unit considered that further experience would be needed and that the contribution of the EFSA PPR Panel would allow

further development of the activity. This resulted in a Scientific Opinion of the PPR Panel as final output offering the opportunity to better engage with the European Commission by considering the possibility of transforming the Scientific Opinion into a Guidance document which can then be taken note by the Commission (as it has happened in the past).

The Scientific Opinion will provide scientific advice on how to perform comparative *in vitro* metabolism studies and how to interpret the results in the context of pesticides risk assessment. This will likely contribute to enhance harmonisation when applicants conduct these studies and to better support regulators when evaluating the data, thus increasing consistency during the Peer Review process.

Therefore, EFSA requested the PPR Panel to develop a Scientific Opinion to:

a) illustrate the testing strategy that should be applied to investigate interspecies comparative *in vitro* metabolism, the minimum requirements that should be included in the testing protocol for the selected assays and interpretation of the study results for the identification of human metabolites not properly assessed by toxicological studies using laboratory animal species.
b) assess the relevance of the toxicological profile of human metabolites not properly assessed by toxicological studies using laboratory animal species.

1.2. Interpretation of the terms of Reference

The PPR developed this scientific document as an Opinion and not as a Guidance as requested in the terms of reference. Based on the experience gained with the application of this opinion, the development of a Guidance may be considered in the future.

The opinion focusses on the regulatory context of pesticide active substances at the EU level to satisfy the data requirement set in Commission Regulation (EU) No 283/2013 related to conducting comparative *in vitro* metabolism studies. However, the Scientific Opinion is relevant for all areas of chemical risk assessment.

The following considerations from the legislation (Commission Regulation (EU(No 283/2013)) should be kept in mind when conducting comparative *in vitro* metabolism studies and interpreting the results obtained:

- *'The relevance of generating toxicity data in animal models with dissimilar metabolic profiles to those found in humans shall be addressed, if such metabolic information is available, and taken into consideration for study design and risk assessment' (Annex, Section 5 Introduction, L93/21)*.
- *'Comparative in vitro metabolism studies shall be performed on animal species to be used in pivotal toxicology and on human material (microsomes or intact cell systems) in order to determine the relevance of the toxicological animal data and to guide in the interpretation of findings and in further definition of the testing strategy' (Annex, Section 5.1.1. Absorption, distribution, metabolism and excretion after exposure by oral route, L93/22)*.
- *'An explanation shall be given, or further tests shall be carried out where a metabolite is detected in vitro in human material and not in the tested animal species' (Annex, Section 5.1.1. Absorption, distribution, metabolism and excretion after exposure by oral route, L93/22)*.

Taking into account the regulatory text, the first problem formulation is the identification of human metabolites not sufficiently assessed by toxicological studies using laboratory animal species. Therefore, this opinion is aimed to provide scientific and technical discussion on how to perform comparative *in vitro* metabolism studies, how to interpret the results obtained and which further assessment shall be carried out when a metabolite is found only in human material (i.e. unique human metabolite, UHM) or present at higher amounts in humans than animals (i.e. disproportionate human metabolite, DHM) (see Section 4 for more detailed definition).

When the legal text of the data requirement is implemented in the context of this opinion, the points above are interpreted as follows:

a) The opinion focuses on *in vitro* methodologies and study design related to liver metabolism with the aim to identify UHM and DHM because the liver is considered the organ with the highest metabolic capacity compared to other organs and tissues (see Section 3 for more detailed information).

b) The results of comparative *in vitro* metabolism studies will support the peer review process of the risk assessment of pesticides active substances in the interpretation of toxicological test results and their extrapolation to the assessment of hazards and risks for humans.
c) Data requirement mentions that further action is needed when '*a metabolite is detected in vitro in human material and not in the tested animal species*' as well as '*the relevance of generating toxicity data in animal models with dissimilar metabolic profiles to those found in humans shall be addressed, if such metabolic information is available, and taken into consideration for study design and risk assessment*' (see Section 6). The study design and experimental strategy outlined in this opinion intend to optimise the detection of potential UHM and allow the evaluation of DHM.

Although this opinion focuses on the use of *in vitro* test systems related to liver metabolism, depending on the existing information on the test item, extrahepatic tissue metabolism needs consideration as well. If organs other than the liver are studied, the same principles as indicated for the liver apply to identify UHM and DHM. It is anticipated that the study design described in this opinion can also be used for other aims related to the use of toxicokinetic (TK) information to support chemical risk assessment and recommendations will be given in the opinion.

2. Overview of the Structure and Key points of the Scientific Opinion

Comparative *in vitro* metabolism studies aim to assess potential interspecies differences in metabolism of a test item. This Scientific Opinion is tailored to the EU data requirement on pesticide active substances,[1] and therefore, the test item is used as synonym of the pesticide active substance. The comparison of metabolites formed *in vitro* should allow evaluation whether all metabolites formed in the human test system are also present at comparable level in pivotal animal species used in toxicological studies and therefore, if their relevance for human risk assessment has been appropriately addressed.

The experimental strategy outlined in this opinion is schematically presented in Figure 1 and aims to optimise the chances to detect all relevant metabolites, by minimising sources of variability and/or bias during the planning and execution of the studies. This variability is related to the specific and critical factors and conditions in the test system, test item, preliminary studies, experimental design of the main study, analytical matters and evaluation of results, i.e. all factors that can influence the amount of metabolite formed and therefore its analytical detection (see Section 3).

a) Test system and test item

Primary hepatocytes from all animal species used in the EU pesticide active substance regulatory data package on mammalian toxicology (i.e. rat, mouse, dog and rabbit) and from humans are the recommended test system. Primary hepatocytes are recommended since human hepatocytes are considered the best surrogate for prediction of human *in vivo* metabolism (Obach et al., 2016) (see Section 3.1). The recommendation of hepatocytes is consistent with the legal requirements for pesticides where use of '*microsomes or intact cell systems*' is cited. It reflects the latest and most appropriate state of knowledge and science.

Test item characterisation (e.g. radiochemical purity, solubility) together with determination of cytotoxicity in preliminary experiments is important element for concentration selection in the main experiment.

b) Preliminary experiments

Regarding hepatocytes, cytotoxicity and metabolic competence are the most important pieces of information required before performing the actual experiments. Determination of *in vitro* clearance is useful for deciding whether long-term incubation in cultured hepatocytes is required to assess slow-metabolising test items.

Regarding the test item, it is of importance to assure that it is not excessively bound to plasticware or other non-biological materials in the system (see Section 3.2).

c) Experimental design of the main study

The PPR Panel recommends to maximise the chances of identifying UHM and DHM by taking into account the test item concentration, incubation time and technical replicates in the experimental design (see Section 3.3). Metabolite formation from a test item is dependent on:

[1] Active substances refer to both new and existing one in the renewal process.

- the kinetic characteristics of metabolising enzymes affecting the selection of test item concentration,
- the length of incubation time affecting the sensitivity of detecting selective metabolites and
- the number of technical replicates affecting the robustness of the measured metabolite amount.

To limit the contribution of individual biological variability and having a response representative of the 'average' population, the use of pooled hepatocytes is recommended (see Appendix C for further details). As a recommendation, a standard hepatocyte concentration of 1 million hepatocytes/mL is suggested, but dependent on the velocity of conversion this can be changed and ideally optimised (Appendix A).

For the above-mentioned reasons, the design of 3 × 3 × 3 (concentrations, time points, technical replicates of pooled test systems) would allow a reasonable setting to identify UHM and DHM, which may be less likely with simpler designs such as 1 concentration and incubation time (Whalley et al., 2017).

The study design includes also considerations on the applicability domain of the test (Section 3.4) and selection of species (Section 3.5), which may result in modifications of the execution of the actual experiments.

d) Definition of unique human metabolite (UHM) and disproportionate human metabolite (DHM)

According to this Scientific Opinion, metabolites of concern are those that are unique to human, i.e. not detectable in incubations with test material from other test species, i.e. UHM, and those that are present in more than fourfold abundance at any sampling time point in human hepatocytes as compared with other species, i.e. DHM (Section 4). The assignment of a metabolite as disproportionate may be challenging especially at abundance ratios close to a factor of 4, considering biological and technical complexity and variability. However, the 3 × 3 × 3 experimental design (see above) provides ample experimental results and supports integrated analysis and decision-making.

e) Chemical Analysis

Regarding the analysis of the incubates, test item-related radioactivity recovery and metabolite profile are the most important parameters in the first step. On this basis, the assignment of DHM and UHM is made. In the second step, structural characterisation of the assigned metabolites is performed with appropriate mass spectrometry (MS) and nuclear magnetic resonance (NMR) techniques. The scheme of metabolic pathways is an important part of chemical analysis and for further consideration of potential consequences (Section 5).

f) Toxicological assessment of metabolites

For a DHM that has already been co-tested in *in vivo* animal toxicity study with the parent compound, a pragmatic approach to avoid additional animal testing is to apply an additional uncertainty factor (UF) to the default UF (100), resulting in an overall larger UF for deriving a health-based guidance value (HBGV) for the parent compound. The default UF of 4 allows for interspecies difference in toxicokinetics, consequently the metabolite formed in quantities more than fourfold in humans compared to animal species is not covered by animal testing and would require an additional UF, the value of which is dependent on the relative quantity of the DHM in human and animal hepatocytes.

Due to current and developing toxicological assessment schemes employing new approach methodologies (NAMs) and adherence to the 3Rs principle, the hazard identification (Section 6.2) would begin with the application of QSAR and read-across methodologies, and proceed to various *in vitro* toxicological testing methods and batteries. If these approaches do not provide sufficient level of necessary information, *in vivo* animal toxicity studies, e.g. a 90-day rat study, could be considered. This step-wise approach is in alignment with earlier EFSA guidance documents and scientific opinions.

In case a human metabolite is not covered by parent compound and even if a limited toxicological database exists, additional UF can be applied for setting health-based guidance value for this unique human metabolite (Section 6.3). The additional UF may be determined by expert judgement on a case-by-case basis. It is also possible to apply an additional uncertainty factor for extrapolation of a low observed adverse effect level (LOAEL) to a no observed adverse effect level (NOAEL), or the data from the critical study may be modelled to derive a benchmark dose (lower confidence limit) (BMDL), which can then be used as the reference point (also referred to as point of departure) for the derivation of the health-based guidance value.

g) Recommendations for the future

Obviously, the comparative *in vitro* metabolism study can provide a significant piece of information to assess the potential role of metabolites of concern in toxicological risk assessment of a pesticide active substance. Additionally, knowledge of *in vitro* metabolites in incubations with test material from humans and animal toxicology test species would add useful pieces of information for the toxicological evaluation of metabolites in different matrices (crops, animals, ecosystems, residues in environment). Improved understanding of toxicokinetics would support current biomonitoring efforts. Finally, comprehensive integrative regulatory schemes would benefit from physiologically based or physiology-mimicking virtual or *in vitro* test systems to anticipate and prevent the exposure of humans and ecosystems to potentially hazardous active substances in plant protection products (see Section 7).

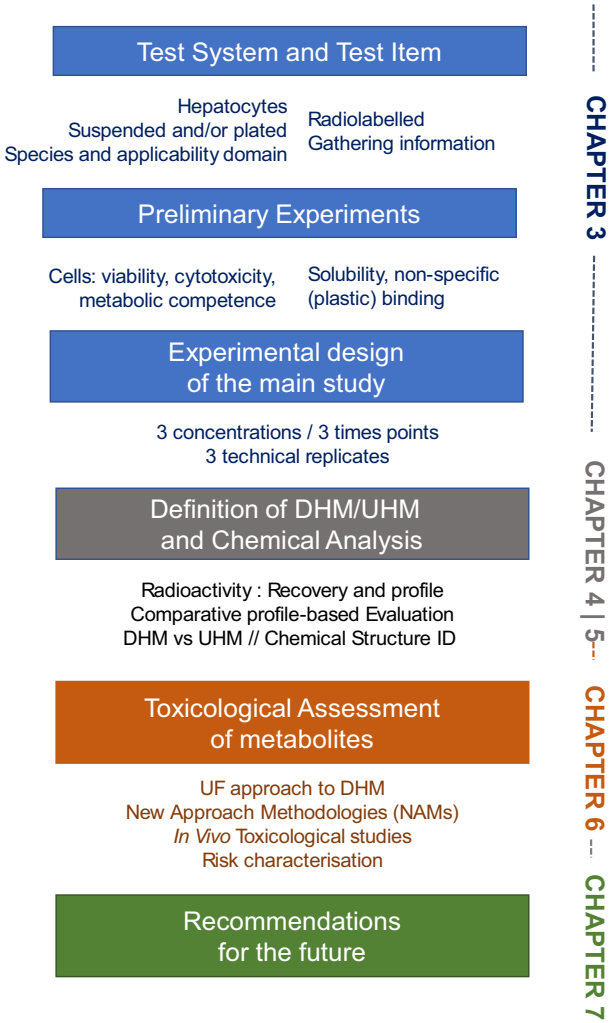

Figure 1: Overview of the structure and key points of the scientific opinion

3. Study design and experimental strategy

3.1. Test systems

There are three principal prerequisites for the selection of an adequate *in vitro* test system to compare metabolism of a test item between different species.

1) The system should mimic the most important *in vivo* metabolising organ (liver[2]) in terms of biotransformation. No single *in vitro* test system can mimic all the complexities of human (or

[2] If a scientifically justified suspicion about the involvement of an extrahepatic tissue in metabolism and metabolites of toxicological significance arises, approaches to address this possibility should be considered (see also Section 7.5).

animal) TK processes, but the liver is considered the most important organ for biotransformation of the majority of chemicals, including pesticide active substances, being responsible to a significant extent for total metabolism.

2) The test system should be compliant with Good *In vitro* Method Practices (GIVIMP) in-house validation standards (OECD, 2018a) or should have been used for metabolism studies for an extended time, so that researchers and regulators (not necessarily in the pesticide area) have sufficient confidence in its performance.
3) Adequate liver-derived test systems from humans and the species of pivotal toxicity tests should be available.

In the pharmaceutical area, liver microsomes are a suitable choice for comparing the metabolite profile among humans and animal species, although they do not contain the whole metabolic machinery. This is because the final assessment of human metabolites relies on *in vivo* human data (phase-1 trials; FDA, 2020). However, in the area of pesticides, *in vivo* human data are rarely available, and therefore in most cases, the assessment of human metabolites can only rely on *in vitro* data. Therefore, *in vitro* studies should be **robust enough** to detect both UHM and DHM and should give the complete picture on metabolic pathways by using the most appropriate experimental test system, particularly important for interpretation of the results.

For addressing the **first problem formulation** of this opinion, i.e. to identify DHM and UHM, isolated primary hepatocytes, either suspended (for short-term incubation) or plated (for longer term incubation), are considered the best choice. Hepatocytes, as the main hepatic cell type, mimic intact liver in terms of most important biotransformation enzymes. Decades of experience exist in using hepatocytes in biomedical research and drug development, particularly for elucidating *in vitro* metabolism, metabolite formation and clearance in drug development by the pharmaceutical industry. Cryopreserved pooled hepatocytes are commercially available. The human hepatocyte test system (plated hepatocytes) has been validated by EURL-ECVAM regarding Cytochromes P450 (CYP) induction (Bernasconi et al., 2019). Hepatocytes isolated from different animal species, including rat, mouse, dog and rabbit have been used in research and drug development for a long time (FDA, 2020). Pooled primary hepatocytes are considered the most suitable model since they average out possible outlier in enzymatic activities and they contain the physiologically relevant expression of phase I and phase II biotransformation enzymes, appropriate concentration of the cofactors, and expression of transporter proteins.

Metabolites formed *in vivo* can be well predicted by using hepatocytes in suspension. The predictability of metabolites formed by subsequent reactions gives more uncertainty for decision-making (Dalvie et al., 2009). This may be attributed to, e.g. the short incubation time that can be used with hepatocytes in suspension (maximum about 4 h, see Appendix C). For such cases, long-term incubation test systems can be used (e.g. 24 h, see Appendix C), for which plated hepatocytes are recommends in this opinion (Ballard et al., 2014; Kratochwil et al., 2017), and if available, the results of long-term incubations are the basis for the assessment.

For addressing **other follow-up problem formulations**, such as formation of reactive metabolites or isoform-specific metabolism, microsomes, S9 fractions, heterologous expression systems or other suitable approaches such as chemical trapping agents might be suitable choices depending on a specific problem formulation.

3.2. The test item

The PPR Panel recommends to gather existing knowledge from the open literature (retrieved by means of the literature search, EFSA, 2011) and/or additional databases (e.g. ADME data in US-EPA dashboard[3], ECVAM databases[4] or OpenFoodTox, see Appendix F) on the test item itself (when it is an existing one subject to renewal) or on structurally similar test items before the comparative *in vitro* metabolism study is conducted. Other information, e.g. on hepatic clearance, may also be helpful for selecting the test conditions, when a comparative metabolism study is designed.

In silico testing, such as QSAR analysis for metabolism prediction can facilitate the planning of the study by postulating the main metabolic pathways in humans and potential human metabolites,

[3] https://comptox.epa.gov/dashboard
[4] Worth et al. (2018a,b): https://data.europa.eu/euodp/en/data/dataset/jrc-eurl-ecvam-rodent-in-vitro-intr-clear-db and https://data.europa.eu/euodp/en/data/dataset/jrc-eurl-ecvam-rodent-in-vivo-biotrans-db
[5] https://www.efsa.europa.eu/en/data/chemical-hazards-data

including primary and more distal metabolites (also called secondary metabolites, i.e. formed by a second or more reactions), but this information cannot be used as alternative to an *in vitro* study, and should be considered as supplementary information only.

In order to design the study in a proper way, information on the physical–chemical properties and characterisation of the test item are required (GIVIMP, OECD, 2018a). More details are given in Appendix B.

As part of a toxicological dossier developed for a pesticide active substance, prior to approval for use within EU (Regulation EC (No) 1107/2009), an *in vivo* TK study, usually performed in the rat following the OECD Test Guideline 417, is required (OECD, 2010). The radiolabelled test item is administered to the animals and afterwards all relevant metabolites (present at ≥ 5% of the administered dose) are identified in body fluid and tissues (e.g. plasma, serum, blood, urine and faeces) to provide a metabolic scheme for the test item. The PPR recommends that the radiolabel material used in the *in vitro* metabolism study to be of similar composition than the one used in in the *in vivo* metabolism study. The identification refers to the exact structural determination of relevant metabolites. The knowledge obtained with the *in vivo* rat kinetic study provides important insight into the potential major metabolites that could be formed *in vitro*, as in general grossly identical metabolites are expected to be formed *in vitro* and *in vivo* (see also Section 6.2.1). This information, together with the possible derivation of TK parameters such as *in vivo* clearance and half-life, might help on the *a priori* selection of the most appropriate test system conditions (e.g. short-term vs. long-term incubation). In the rat *in vivo* TK study, a general overview about the metabolism is gained and a tentative identification of major metabolites done. However, more detailed TK parameters such as clearance, half-life or bioavailability of metabolites are not possible to be derived, because of experimental design and unspecificity of the analytical method used (i.e. quantifying total radioactivity, not differentiating between parent chemical and metabolites).

The *in vivo* TK study results on the test item may help to design and interpret the comparative *in vitro* metabolism study. As pointed out in OECD TG 417, the *in vivo* rat study provides useful information about the analytical method for parent and metabolites (OECD, 2010). The results would help to select an analytical method for detection and identification of the metabolites in the *in vitro* hepatocyte study.

For the active substances subject to a renewal process, the full package of toxicological and *in vivo* TK studies is already available and the comparative *in vitro* metabolism studies can be informed on that basis. For the new active substances, timing of conducting the *in vitro* metabolism study might vary depending on the characteristics of the test item and the policy of the laboratories or companies. If TK parameters could be derived from the *in vivo* study, they may contribute to constructing a PBK model for the test item and metabolites, which can be used to provide more insights into dose-dependent internal exposure to parent chemical and metabolites (see Section 7.6). Comparison of metabolism between *in vitro* rat hepatocytes and the *in vivo* TK rat study may be helpful to extrapolate the results of the *in vitro* human hepatocyte study to *in vivo* hepatic metabolism and contributes to the acceptability assessment of the comparative *in vitro* metabolism study. However, there is very little published data on this potential. Regarding human metabolites, available studies indicate that the metabolite profile obtained *in vitro* can reflect the *in vivo* metabolite pattern, especially concerning primary metabolites (Anderson et al., 2009; Dalvie et al., 2009). Generation of *in vitro* metabolism data in laboratory animals and comparison of these data with *in vivo* TK data for pesticide active substances provide the information how *in vitro* metabolism data reflect *in vivo* metabolite profiles.

In the area of pesticides, according to EU pesticide active substance data requirements, another important source of information on metabolite formation are pesticide residue metabolism studies in livestock, the availability of which, however, depends upon the intended pesticide use.

3.3. Preliminary experiments

Before the comparative metabolism study is to be performed, the PPR Panel recommends to conduct preliminary experiments to assess (1) the cytotoxic potency of the test item to the test system, (2) the metabolic competence of the test system and (3) the non-specific binding of the test item.

1) Cytotoxic potency of the test item to the test system

Biotransformation enzyme activity can be reduced in chemical-induced stress and cytotoxicity situations. Therefore, comparative metabolism studies are to be performed at test item concentrations

that do not affect hepatocyte function. Test item-induced toxic effects can be assessed by performing a cytotoxicity study as described in more detail in Appendix C.1.1, allowing to establish the highest test item concentration that does not affect hepatocyte function to be used in the comparative metabolism studies. As toxicity can be caused by a biotransformation product (i.e. metabolite), it is of importance that the cytotoxicity studies are performed with metabolically active hepatocytes. Studies need to be performed with hepatocytes of all selected species, since species-dependent differences in cytotoxic potencies of the test item may occur.

2) Metabolic competence of the test system

Adequate functioning of the test systems would be needed to obtain relevant data on the test item's metabolite formation. Such information can be obtained by assessing the time-dependent metabolite formation of well-known reference compounds, which can also be applied as positive control in the comparative metabolism study itself (more information is presented in Appendix B.5 and Appendix C.1.2). Comparative metabolism studies should only be performed with hepatocytes batches with sufficient metabolic activity.

3) Non-specific binding of the test item

For the adequate interpretation of the data obtained in the comparative metabolism study, insight into the availability of the test item to the test system (hepatocytes) is essential. Especially for very lipophilic test items, non-specific binding, e.g. to plastic and/or medium constituents, may largely reduce the availability of the test item to the hepatocytes. More information on this topic is presented in Appendix C.1.3.

Besides these preliminary studies, performance of a pilot clearance study can be helpful. Such a clearance study provides information on the velocity of the conversion of the test item, which can be used to estimate whether a short-term incubation would suffice to reach 20% conversion or whether a long-term incubation would be required. More information on such clearance studies is presented in Appendix C.2.

3.4. Experimental design of the main study

Selection of the experimental test conditions for conducting the comparative *in vitro* metabolism study and application of well-defined acceptance criteria is critical to obtain robust and reliable data. The choice of, amongst others, hepatocyte concentration and culture medium have a large impact on intrinsic clearance values obtained with *in vitro* hepatic studies as shown in an analysis of clearance data obtained with different experimental set-ups (Louisse et al., 2020). Because intrinsic clearance in hepatocytes is practically wholly determined by metabolism, it can be expected that also with *in vitro* metabolite formation studies, experimental set-up influences the study outcome. The experimental design described in this Scientific Opinion is considered optimal to maximise the chances of detecting potential metabolites, particularly those which are formed with a slow rate of formation and therefore prone to be not detected, since occurring at concentrations below the limit of detection of the analytical method. Since the aim of such studies is to compare the metabolic profiles across species and to possibly identify also DHM (in addition to UHM), it is important to use the same experimental set-up, when possible, across species. By assuring this, the identified *in vitro* species differences will not result from differences in experimental set-up but will reflect (differences in) biology. It is intended to minimise sources of experimental variability which would affect the comparison of metabolic profiles.

Although the design described below serves as an example regarded by the PPR Panel as a minimum option, it has to be kept in mind that pesticide active substances are a highly variable group of chemicals belonging to different chemical classes with different properties, and therefore, it is expected that a single design (or even a range of designs) does not fit all different test items.

At least three concentrations: Concentrations of the test item have a significant impact on the detection of relevant metabolites. In the main experiment 3, concentrations are recommended: high, medium and low. One high concentration (e.g. 100 μmol/L), which should not be cytotoxic (i.e. viability $\geq$ 80% compared to solvent control), is recommended to maximise the chance of detecting metabolites that are formed via metabolic reactions with low affinity. However, at such high substrate concentrations, certain metabolic conversions may be inhibited, indicating that incubation with a lower substrate concentration is required (e.g. 1 μmol/L) to allow formation (and subsequent detection) of metabolites that would be formed via these reactions. It is highly recommended to also include an

intermediate concentration (e.g. 10 μmol/L) as it would provide relevant insights into the kinetics of the reactions underlying the metabolite(s) formation, such as the concentrations at which metabolic reactions get saturated.

At least three time points: In the main experiment, at least three incubation time points are recommended. An incubation period of 4 h is considered the maximum incubation time applicable with primary hepatocytes in suspension cultures, allowing the maximum metabolic conversion and formation of more distal metabolites. It is highly recommended to include at least two sampling time points (e.g. 1 and 2 h) (see Appendix H) to assess whether changes in metabolite levels are time dependent (indicating adequate functioning of the *in vitro* test system during the whole test duration). Furthermore, information on time-dependent changes in metabolite levels provides more insight into the kinetics of formation of primary and more distal metabolites, facilitating a better evaluation of DHM in humans compared to animals.

Three technical replicates of pooled test systems: It is of importance to perform three simultaneous incubations of each pooled hepatocyte preparation to address the variability of the assay protocol.

Depending on the outcome of the test, additional replicates, time points or concentrations might be needed to substantiate whether e.g. certain reactions are saturated. In addition, depending on the chemical properties of the pesticide active substance, such as solubility and cytotoxicity, chemical-specific optimal concentrations can be selected.

Slow-metabolised test item: To ensure that the test item has been metabolised and that metabolites are formed, the PPR Panel recommends to observe a minimum of 20% decrease of the concentration of the test item (and its related conversion into metabolites) at the end of the 4-hour incubation period in hepatocytes from all species. The proposed value of 20% takes into account several factors such as variability and method of analysis (FDA, 2018). In case no or limited metabolism is observed (i.e. less than 20% decrease of the test item), the PPR Panel recommends to perform a longer incubation of, e.g. 24 h in plated hepatocyte incubations to avoid that potentially unique metabolites are not detected due to slow conversion *in vitro* (Ballard et al., 2014; Kratochwil et al., 2017). Based on available data (e.g. having data on structurally similar compounds or results from a pilot *in vitro* clearance study), a long-term incubation could be immediately selected. For test items which are not metabolised or metabolised at very low rate, or for which the half-life is known to be shorter than 24 h, the omission of the 24-h incubation can be reasonably justified.

Detailed information on a recommended experimental set-up is given in Appendices 2 and 3.

3.5. Applicability domain of the test

The described experimental set-up in Appendix C.3 will not be applicable for all test items. For example, for a volatile test item, a closed system would be required to prevent evaporation and for substances that have a poor solubility in water, dosing in medium without proteins may cause problems, requiring a tailored study design. There may be cases in which, due to the knowledge on the physico-chemical and toxicological properties of the test item, a comparative *in vitro* metabolism study does not appear scientifically necessary or testing for this purpose not technically practicable. Examples of such test items are some inorganic active substances (e.g. calcium carbonate) for which *in vitro* testing is not needed.

For test items that are complex mixtures the recommendation is to test the most representative component(s) within the mixture, the selection of which is done on the basis of different criteria, including but not limited to, the results of *in vivo* ADME study or toxicity information (e.g. in silico methods or other NAMs, if available), as described in the EFSA Guidance on mixture risk assessment (EFSA Scientific Committee, 2019). This issue is becoming more and more relevant considering the use of essential oils and biopesticides.

3.6. Selection of species

As a default approach, the PPR Panel recommends to perform comparative *in vitro* metabolism studies by using hepatocytes from human and from all animal species used in the EU pesticide active substance regulatory data package on mammalian toxicology (i.e. rat, mouse, dog and rabbit).

However, a tiered approach could be considered depending on the problem formulation. For identification of UHM and DHM, the comparative *in vitro* metabolism studies could be conducted by using human and rat hepatocytes in a first step. Depending on the outcome (e.g. metabolite present in humans and not in rats), it could be further investigated whether the human metabolite is formed in

all other species relevant for the assessment of the particular substance. This is based on the fact that the majority of *in vivo* toxicity studies in the EU pesticide active substance regulatory data package are performed in rats (from short-term to long-term toxicity as well as reproductive toxicity). If the rat is the most sensitive species, no further comparative metabolism studies in other species are necessary.

If the problem formulation is to address the human relevance of a reference point or a health-based guidance value derived from a study on a species other than the rat, the PPR Panel recommends to conduct further tests by employing hepatocytes from other species (i.e. mouse, dog, rabbit).

It may be anticipated that comprehensive comparative *in vitro* metabolism testing (i.e. to have all species, not only the rat) could provide information that is useful and cost-effective considering that the problem formulation in the EU pesticide assessment is rarely limited to the derivation of a health-based guidance value from a single species. The EU pesticide assessment also includes hazard identification by considering all the species, e.g. on developmental toxicity (i.e. rabbit), mutagenicity and carcinogenicity (i.e. mouse) as well as endocrine disruption (including dog). The Panel advises to consider a tailor-made and flexible approach depending on the problem formulation.

4. Definition of unique and disproportionate human metabolites

Differences in the metabolites profile between the species can be either qualitative or quantitative. The qualitative differences can originate from the lack of an enzyme in a particular species producing a metabolite of interest. Quantitative differences can be observed due to various reasons, like lower expression or different kinetics of the enzyme, a different (iso)enzyme catalysing the metabolite production or different rate of further metabolism of the metabolite (Weidolf and Wilson, 2016). Even if the metabolite profiles are qualitatively similar in different species, usually there are large differences in quantitative proportions of metabolites (e.g. Pelkonen et al., 2009; Weidolf and Wilson, 2016).

Interindividual differences of various metabolites produced in animal species used in the toxicological studies are usually regarded as of smaller importance vis-à-vis interspecies differences. In any case, historical data in metabolic enzyme activities and metabolite amounts, if available, would be useful to collect genetic shifts in strains over the time. In the area of pesticides, it is recommended that the same strain as the one predominantly used in the toxicological studies is used in the comparative *in vitro* metabolism studies to provide representativeness of the results.

Interindividual differences in humans (e.g. due to genetic polymorphisms in xenobiotic metabolising enzymes) are generally large (Zanger and Schwab, 2013). If the problem formulation is to identify UHM or DHM, the comparative *in vitro* metabolism study can be conducted by using pooled human hepatocytes from both genders as representative of an 'average' human, avoiding the large differences possibly related to the use of hepatocytes from a single donor.

If the problem formulation is to address the uncertainty from human variability in metabolism (currently addressed by a subfactor of 3.16 of the default UF value of 10 to account for TK differences across humans), the best strategy would be to identify the isoform-specific metabolism in simplified test systems (e.g. recombinant enzymes, microsomes) and confirm results by using hepatocytes from different single donors characterised for their relative (iso)enzyme content.

It is expected that comparison and analysis of metabolic pathways between species could be facilitated by using an electronic database like OECD MetaPath (Kolanczyk et al., 2012) if comparative *in vitro* metabolism is submitted electronically (EFSA, 2021).

4.1. Unique human metabolite(s)

A metabolite that is only detected in the human-derived test system and not in *in vitro* or *in vivo* laboratory species is defined as 'unique human metabolite; UHM'.

The PPR Panel recommends to consider the UHM to be present at least in one of the tested concentrations.

There might be a situation in which parent compounds or metabolites can act as inhibitors of metabolism at higher concentrations (e.g. Houston and Kenworthy, 2000; Wu, 2011; Leow and Chan, 2019). In addition, formation of more distal metabolites may lead to concentration-dependent changes in putative UHM. In both situations, the assignment of UHM (and also DHM, for that matter) may prove difficult and it has to be done on a case-by-case basis. Compounds may also act as inducers of biotransformation enzymes, but these effects are not likely to occur within 24 h (Bernasconi et al., 2019).

4.2. Disproportionate metabolite(s)

In the context of this Scientific Opinion, a disproportionate human metabolite (DHM) is defined as a metabolite that is present at a quantity higher than four times in human hepatocytes when compared to laboratory animal species hepatocytes at any sampling point.

The factor of 4 is derived from the TK interspecies subfactor of the default uncertainty value of 10 to account for animal to human differences in TKs as defined by IPCS (2005). The assessment of DHM in the pharmaceutical area relies on human data obtained in phase-1 clinical trials (FDA, 2020). In the area of pesticide active substances, *in vivo* human data are rarely available, and therefore, the identification of DHM as described in this opinion can only be based on *in vitro* data.

The analysis should consider any of the concentrations and the temporal trend for a given metabolite in order to define a DHM (see Appendix H). The PPR Panel recommends to define a DHM based on a direct comparison of appropriate metabolite amounts in human and animal metabolite profiles (see Section 5).

The study design outlined in this opinion is considered appropriate to allow an enhanced detection of DHM. However, as the translation to the *in vivo* situation involves many additional factors, the relevance of the DHM for the *in vivo* situation should still be evaluated, especially regarding the projected (relative) internal exposure of humans.

In case the highest concentration tested is different between species due to differential cytotoxicity, the comparison of the relevant metabolites across the species should be done by comparing the formation of these metabolite at the same non-cytotoxic concentrations.

5. Chemical analysis

Obtaining quantitative information about the number, identity and abundance of metabolites formed is one of the main goals of *in vitro* metabolite profiling (Spaggiari et al., 2014a,b). In the pharmaceutical field, it is an essential part of preclinical and clinical drug discovery and development (see e.g. Pelkonen et al., 2009). In the context of pesticide active substances and this Scientific Opinion, the principal goal is to compare the metabolite profiles of various species, especially rat, mouse, dog and rabbit, with the human profile, to identify UHM and DHM, which cannot be assessed in formal *in vivo* studies.

Primarily, concentrations of test item are selected on the basis of physico-chemical parameters such as LogP, LogD and solubility as well as cytotoxicity and assay requirements. However, from the beginning, it is of importance to consider potential prerequisites and requirements of the analytical system to become confident that metabolites and metabolic profiles measured in incubations with different species are measured in such a manner that the presence of UHM and quantitation of DHM can be confidently evaluated.

5.1. Suitability and selection of analytical tools

Because the rat *in vivo* TK studies is performed usually with the radiolabelled material to obtain definite knowledge on the TK characteristics of the studied test item, the same radiolabelled test item could be used in the *in vitro* metabolism studies, employing appropriate separation techniques to produce the metabolite profile (see Section 3.1).

High-pressure liquid chromatography (HPLC) coupled with mass spectrometry (MS) or high-resolution MS (HRMS) has become the prevailing analytical strategy for *in vitro* metabolism studies (Tolonen and Pelkonen, 2015; Spaggiari et al., 2014a,b). Combining the unique detectability and quantitative measurement of the radiolabelled test item with the superior separation and identification power of the modern LC-MS techniques (together if needed with NMR techniques) by the unlabelled compound could provide the optimal workflow for comparative *in vitro* metabolism studies.

Because all analytical techniques are limited by their capability to detect analytes at a certain threshold, tested concentrations should optimise the chance of keeping poorly soluble compounds in solution while reducing the chance for saturation of metabolic enzymes. Because modern online radiochromatography detectors typically have an unquenched ^{14}C counting efficiency of greater than 90% (Deakin, 2015), the PPR Panel preferred ^{14}C to ^{3}H except when atomic substitution proves difficult to implement or the structure of the compound of interest is altered. Hence, the concentrations of the ^{14}C radiolabelled compound may be increased if the specific activity is limited and sufficient solubility is achieved.

The radiolabelled test item using ^{14}C is preferred for following the fate of the total mass of the substrate added. It can be replaced if it is demonstrated that (a) mass balance and metabolite identification can be adequately evaluated using the unlabelled test substance and (b) the analytical specificity and sensitivity of the method used with non-radioactive test substance is equal to, or greater than, those obtained with the radiolabelled test substance. The basic scenario of the study, suspended or plated hepatocytes in a suitable culture medium, simplifies the analytical performance, e.g. matrix effect is lower than in *in vivo* studies.

5.2. Production of comparative metabolite profiles

5.2.1. Use of radioactive techniques

Radiolabelled material is essential for studies of metabolite profiles and routes of metabolism because all significant moieties or degradation products can be tracked with appropriate labelling. This is especially of significance in estimation of the bound fraction, not extractable and recoverable of parent compound and metabolites, suggesting the production of reactive metabolites (see Appendix E). The quantitative aspect of radioactivity is also useful in establishing metabolite profiles when absolute structure of metabolites is unknown. As previously mentioned, the use of stable isotopes together with the radiolabelled isotope is encouraged to aid in identification of metabolites by using various spectroscopic methods (e.g. radioactive detector (RAM), MS, including fragmentation approaches (MS/MS, MS^n), HRMS or NMR) (see Figure 2).

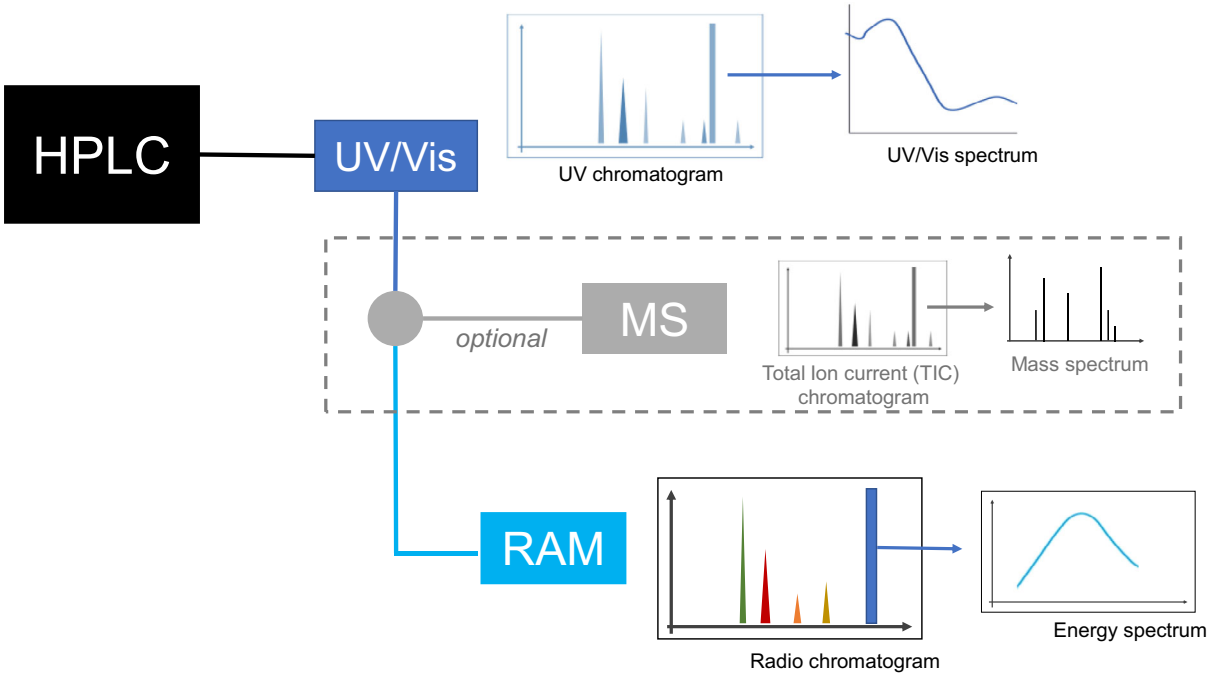

Figure 2: Schematic set-up for online HPLC-UV and RAM with optional online MS detection for structure elucidation

Non-radioactive substances could be used for comparative *in vitro* metabolism studies. Although metabolite quantitation and recovery are more difficult to measure with non-labelled compounds, comparative assay conditions should ascertain, to the extent it is possible, that individual metabolite could be compared across species and time points.

5.2.2. Metabolite profiles of comparative *in vitro* metabolism

As previously mentioned, radiolabelled test material is essential to distinguish labelled metabolites from exogenous matrix. Because the quantitative aspect of radioactivity is invaluable in establishing patterns when structures of metabolites are unknown (and their UV spectra and relative MS ionisation responses are not fully established), relative comparison is an acceptable approach. Figure 3 shows a

hypothetical example of the analytical workflow (also including the existing results of the in vivo rat TK study) with ultimate (radiolabelled) metabolite profiles of human and four experimental animal species.

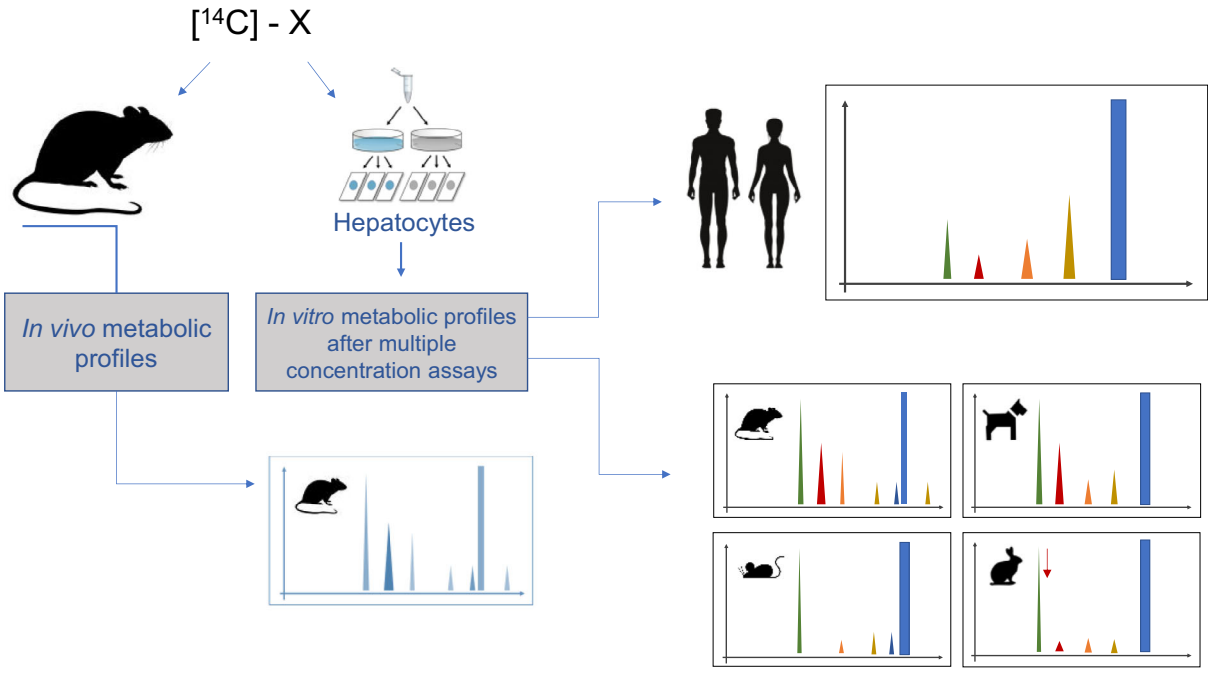

Figure 3: Metabolite profiles of comparative metabolism with radiolabelled compounds from *in vivo* rat and comparative *in vitro* studies

With the radiolabel as a marker, it is much easier to distinguish the precursor compound and its labelled metabolites from endogenous matrix compounds to obtain metabolic profiles. The detected radio-labelled metabolites should elute at the same retention time as observed in the LC-MS chromatogram (see Figure 2). An additional monitoring at low UV wavelength (e.g. 205–210 nm) (nonspecific wavelength where most organic molecules show some UV absorbance) could provide another information of the level of endogenous material co-eluting with the radioactive components of interest. This important information, related to matrix effect, could be useful for further developing the chromatographic separation method when metabolite identification is needed.

As an example, formation of metabolites of [^{14}C]- X with both qualitative and quantitative differences among the species examined is shown in Figure 4. The most notable difference is the total absence of the metabolites marked in black in any species tested, while this metabolite fraction appears present in humans (black arrow), representing an UHM.

Recoveries for all samples following reconstitution should be higher than 90% which allows a valid comparison between species. If a metabolite is detected *in vitro* in humans but not in rats but it is identified in the rat ADME study *in vivo*, this metabolite is not to be considered unique to humans.

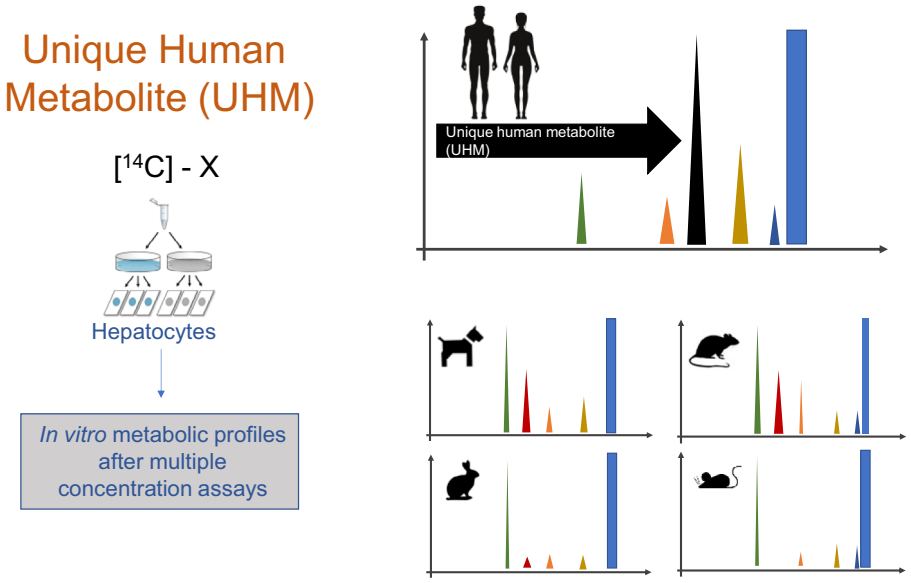

Figure 4: Metabolite profiles of a compound in incubations with hepatocytes from human, rat, mouse, dog and rabbit demonstrating qualitative differences between species

The formation of metabolites of [^{14}C]- X with relative quantitative differences between the species examined is presented in Figure 5; the most notable difference is the difference observed related to the amount of signal recovered from the metabolite marked in orange between human (30%) and any other species tested (Max 5%). The PPR Panel recommends to compare the amount of radioactivity between metabolites taken into account that the total radioactivity recovered is sufficient. This metabolite could be considered as DHM because the relative quantification (peak area: 30%/5% in this example) exceeds a factor of 4, taking into account the analytic variability.

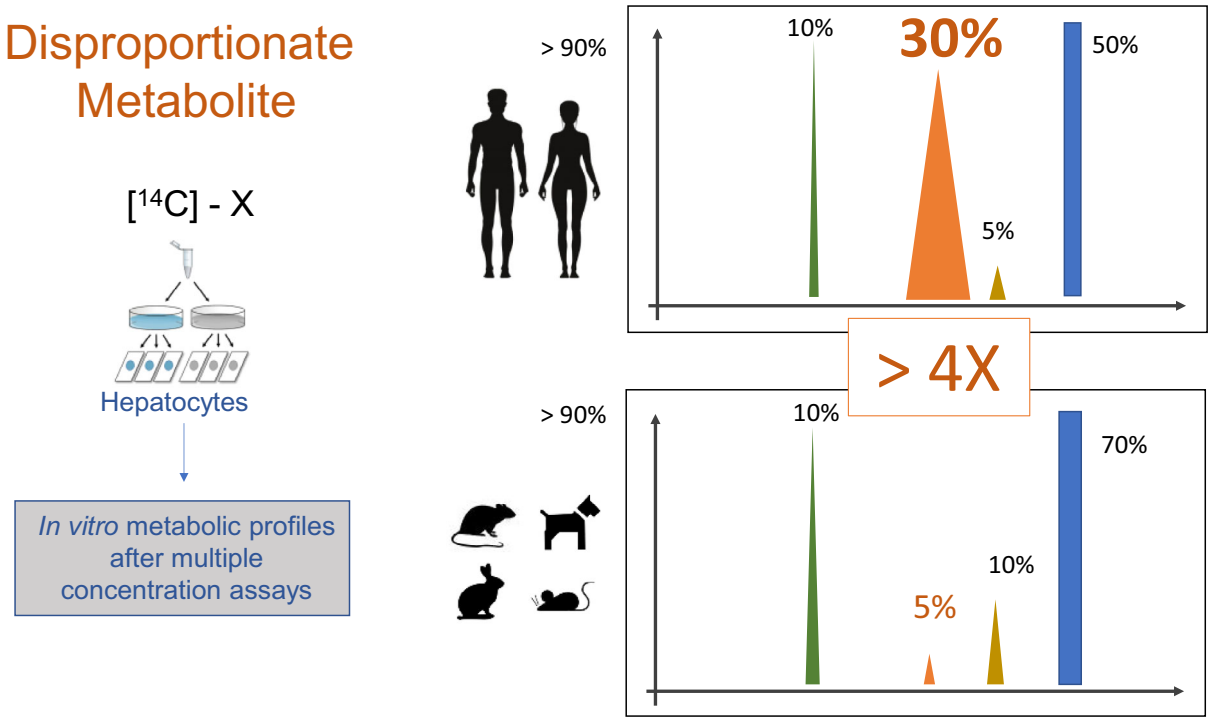

Figure 5: Metabolite profiles of a compound in incubations with hepatocytes from human, rat, mouse, dog and rabbit demonstrating quantitative differences between species

If one UHM DHM is detected in any of the tested situations, the PPR Panel recommends the identification of this metabolite, e.g. by use of MS and/or NMR, in line with metabolite identification described in the OECD 417 test guideline (OECD, 2010).

5.3. Metabolite identification

The PPR Panel recommends identifying all UHM and/or DHM metabolites detected at ≥ 5% of the administered dose, considering a total recovery of administered test substance (radioactivity) higher than 90% as well as describing a metabolic scheme for the test item.

Generally, LC-HRMS techniques offer the most effective technology to screen and/or identify metabolites, in both *in vitro* and *in vivo* conditions. The high mass accuracy of modern instrumentation enables reliable and accurate identification of the molecule, i.e. actual biotransformation with respect to the parent compound, as well as makes the fragment ion mass spectra data more informative in comparison to other types for MS/MS-instruments operating at unit resolution, i.e. triple quadrupole or ion trap MS-based technology. Furthermore, LC-HRMS technique makes possible to obtain accurate analyte concentrations for pharmacokinetic studies and clearance estimation.

A typical metabolite identification approach involves the following steps:

- predict possible metabolites using *in silico* methods (see Section 3.1) or based on researcher's experience;
- apply survey scans of the test and control samples;
- compare ion chromatograms of the test and control samples to determine the *m/z* values of expected and unexpected metabolites;
- conduct product ion scan and/or MS/MS; MSn on the *m/z* of suspected metabolites and propose metabolite structures;

Overall, the above strategies for metabolite identification typically are well suited for identification of metabolites but may be not sufficient for complete structure determination and further quantification of the identified metabolites.

- (Optional) confirm metabolite structures using NMR, synthesis of authentic standards or other chemical techniques combined with MS if needed (Anari and Baillie, 2005).

A prior knowledge of the likely routes of biotransformation of the analyte greatly facilitates the detection and structural characterisation of its metabolites. Together with accurate mass data with low mass error from the detected compounds and because common changes in molecular weights caused by metabolic reactions are well known, the combination of MS/MS and HRMS experiments can enable the differentiation in the biotransformation with different molecular formula but presenting the same nominal mass change. In addition to HRMS, the strength of time-of-flight mass spectrometry (TOF-MS) instruments for metabolite screening is their very high detection sensitivity when a wide mass range of data is acquired. Thus, prediction values of expected metabolites allow for the construction of user-defined MS acquisition protocols that frequently reveal the presence of minor metabolites, even in the presence of a vast excess of coeluting endogenous constituents.

When the LC–MS methods do not give clear structural information for the formed metabolite, or the structure obtained by LC–MS methods must be elucidated in a more detailed manner, the synthesis of possible metabolites and consequent comparison of the LC–MS data for the detected metabolite and synthesised compounds are mandatory. The identification of a metabolite via the MS information only, including additional structural information that could be obtained via fragmentation analysis, high-resolution measurements and/or advanced MS acquisition mode such as ion mobility (IMS)-based approaches is generally considered as sufficient in the present context, but in some cases have to be then completed by NMR for a full structural elucidation.

For formal naming of novel metabolites identified for the first time, the PPR Panel recommends to be consistent with IUPAC nomenclature. The PPR Panel recommends common naming, together with novel structures and/or release of an electronic code for the structure, i.e. the International Chemical Identifier (INCHI) code that is recommended by IUPAC and NIST. The INCHI code and software to generate this code for chemical drawings is freely available (http://inchi.info/software_en.html).

6. Toxicological assessment of Human Metabolites

The toxicological assessment of human metabolites (DHM and/or UHM) should start from evidence from existing studies, if available. Metabolites found as residues in crops and/or livestock or groundwater metabolites may be the same as DHM and/or UHM: hence already known and tested to assess their relevance. If such information is not available, the first step should be based on alternative approaches to animal testing such as NAMS, including read-across, *in silico* tools and/or *in vitro* methods. These alternative approaches are evolving very quickly, and therefore, this section describes different options to assess the available evidence.

For *in vitro* or *in vivo* testing of a metabolite, sufficient material is required, which may need to be synthesised. The PPR Panel acknowledges that testing of certain metabolites may be challenging, e.g. in case a metabolite is unstable. Therefore, the PPR Panel recommends a tailored testing strategy depending on the nature of the metabolite and by considering whether synthesis and testing of a metabolite is possible.

The PPR Panel recommends to assemble, weight and integrate all available evidence in a weight of evidence approach (EFSA Scientific Committee, 2017a,b) in order to reach conclusion on the biological relevance of human metabolites with regard to their toxicity (see Appendix I).

Although the different sources of information are reported separately, the PPR Panel recommends to follow an integrated approach to testing and assessment (IATA) (OECD, 2020a) by integrating all available data derived from multiple methods and sources, starting from *in silico* to *in vitro* and then to *in vivo* data. Pragmatic approaches to avoid animal testing are also described.

6.1. Uncertainty factor approach to a disproportionate human metabolite

While UHM has not been identified in tested laboratory species and is therefore *a priori* a compound of non-assessed toxicity, DHM has been co-tested in toxicity-tested laboratory species however occurring in *in vitro* systems at lower levels than in human hepatocytes.

The **UF approach** can be an option to address safety concerns around DHM of unknown toxicological concern, by conservatively assuming that DHM is the toxic moiety of the test item.

This pragmatic approach to avoid additional animal testing for a DHM could be envisaged by applying an additional UF to the default UF, resulting in a UF larger than 100 for deriving a HBGV for the parent compound. This approach can be useful in cases where human exposure is exhausting a small proportion of the HBGV. In such a case, the overall UF is calculated depending on the relative quantity of the DHM in humans compared to animal species.

The toxicity of a metabolite formed in up to fourfold quantities in humans compared to animal species is assumed to be covered by testing when applying an uncertainty factor of 100 to the reference point for the parent compound. Hence, it can generally be assumed that the quantity of **metabolite** produced in the animal species can be divided by an **UF of 25** (i.e. = 2.5 (interspecies differences in toxicodynamics (TD)) and human variability in TK (3.16) and TD (3.16), while the UF accounting for interspecies difference in TK (4) can be excluded (for explanation of the subfactors, see IPCS 2005) to arrive at the amount of metabolite in humans at the HBGV for the parent compound derived from the reference point divided by an UF of 100.

Two scenarios can be considered as examples for the approach (Figure 6).

1) In the first scenario, the metabolite in humans is formed up to four times higher levels than in rats. For example, taking a quantity of metabolite in rat as 5% and a reference point of 100 mg/kg body weight (bw) per day, a quantity of 5 mg metabolite/kg bw per day has been tested. Applying an UF of 100 results in an HBGV of 1 mg/kg bw per day for the parent compound, whereas the safe dose for the metabolite is 0.2 mg/kg bw per day using the UF of 25. At the HBGV of 1 mg/kg bw per day for the parent compound, the amount of the metabolite formed in humans is tested as safe up to 0.2 mg/kg bw per day. This represents 20% of the HBGV of 1 mg/kg bw per day and it is fourfold higher than the 5% in rats. Hence, for a fourfold higher quantity in humans compared to the animal species, the toxicity of the metabolite has been sufficiently covered at the level of exposure at the HBGV when applying the UF of 100 for the parent compound.
2) In the second scenario, the metabolite is formed in higher than fourfold quantities in humans compared to the animal species. For example, if the quantity of the

disproportionate metabolite in rat is 5% and the quantity in humans is 40% (eightfold higher), the toxicity of the disproportionate metabolite is not sufficiently covered when using an UF of 100. In this case, assuming a reference point of 100 mg/kg bw per day for the parent compound, the HBGV would be 1 mg/kg bw per day for the parent, equivalent to a dose of 0.2 mg/kg bw per day for the disproportionate metabolite (see example 1). When exposing humans to the parent compound at a level of 1 mg/kg bw per day, i.e. equal to the HBGV, the metabolite level produced in humans would be 0.4 mg/kg per day. Since only exposure up to 0.2 mg/kg bw per day to the metabolite (5%) can be considered as safe, the exposure in humans would be a concern. This concern can be accounted for by applying an additional UF to the HBGV. This additional UF can be derived by calculating the ratio of quantities human/animal metabolite *in vitro* and divide this by 4 (see Figure 6). In this example, an additional UF of 2 ([40/5]/4) would be calculated resulting in an overall UF of 200. Accordingly, the HBGV for the parent compound would be 0.5 mg/kg bw per day. Lowering the HBGV as described would adequately protect humans to the toxicity of the parent compound and the disproportionate metabolite.

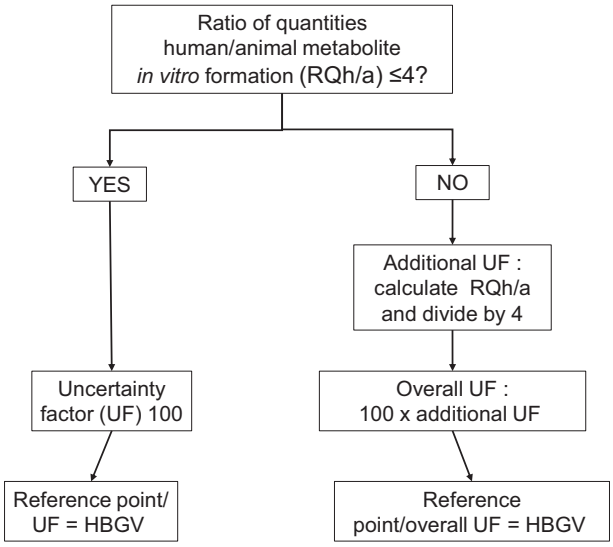

Figure 6: Avoiding additional animal testing of a DHM by using an additional uncertainty factor for the derivation of HBGV. Examples are given in Appendix J. (RQh/a: ratio of quantities of human/ animal metabolite formed *in vitro*).

6.2. Hazard identification

In line with other EFSA scientific documents (e.g. EFSA PPR Panel, 2016), the PPR Panel recommends, as a first step, to assess the potential genotoxicity of a DHM or UHM. In case a human metabolite is present only in one laboratory species, assessed only in very specific or limited studies (e.g. developmental toxicity study in rabbits or subchronic studies in dogs), further toxicological considerations are appropriate, since other endpoints might not be covered (e.g. carcinogenicity).

6.2.1. New Approach Methodologies (NAMs)/Non-animal or Alternative Testing Methods

One of the options to assess the potential toxicity of UHM and DHM would be animal testing, however, such studies are undesirable in view of the worldwide effort to reduce animal testing for ethical reasons. In addition, such studies are time consuming and expensive. Therefore, alternative methods to animal testing, also referred to as NAMs provide options to generate information on the potential toxicity of UHMs and DHMs. Although NAMs are widely used for data-poor chemicals, there is currently no EFSA harmonised guidance document on the use of NAMs in chemical risk assessment (see Section 7.9).

QSAR and Read-across methods

A first step to identify the potential hazard of the metabolite under investigation would be the use of non-testing methods such as QSAR and read-across. These methods allow to infer toxic properties (e.g. genotoxicity or general toxicity) of UHM or DHM based on data for structurally similar substances. The application of QSAR and read across for genotoxicity and general toxicity, and some worked examples, have been previously described in detail in the PPR Panel Guidance on the establishment of the residue definition for dietary risk assessment (EFSA PPR Panel, 2016). Below a brief description of both approaches is presented.

QSAR models are statistically based or knowledge-based models which link chemical structure to biological activity for a number of compounds and then predict such biological activity for data-poor chemicals within the same applicability domain. The increasing availability of data and modern machine learning techniques will further facilitate the development and increase the predictivity of QSARs (e.g. Pawar et al., 2019; Muratov et al., 2020; Keyvanpour and Shirzad, 2021; Shah et al., 2021). The *in silico* part of assessment generally starts with the QSAR analysis, e.g. using the OECD Toolbox evaluation,[6] of the parent and related metabolites to reveal highest similarities (e.g. Dice index, analysis of functional groups).

While a simple similarity measure does not always reflect changes in reactivity and itself may not be enough to fully justify a read-across prediction (Schultz et al., 2015, 2019), further analyses at organic functional group levels are performed. In these analyses, potential significance of various functional groups in terms of TK and TD is considered.

QSARs are predominantly used for predicting genotoxic properties of substances with a known chemical structure. Currently, their use in predicting general toxicity is limited. For predicting genotoxicity, QSAR models are generally considered to give reliable results when the substance to be predicted is within the applicability domain (Benigni et al., 2019, 2020). When using QSARs, conclusions on the genotoxicity are generally drawn on the outcome of two or more independent QSAR models for each genotoxicity endpoint. All relevant genotoxicity endpoints, i.e. gene mutations, and structural and numerical chromosomal aberrations, have to be considered. The conclusion on the genotoxic potential of the metabolite(s) is justified by providing all necessary information, e.g. which models were used, the applicability domain and reliability of the models.

Read-across is a technique used for predicting endpoint information for one substance by using data on the same endpoint from one or more other substances which have been shown to be structurally related. In a read-across approach, experimental data on genotoxicity or other endpoints for one or more chemicals (source chemical(s)) can be used to make a prediction for the same endpoint for one or more different chemicals (target chemical(s)), e.g. UHM and DHM. The primary assumption is that the source and target chemical(s) provoke similar effects related to the assessed endpoints, usually based on structural similarity, and therefore exhibit similar biological activity (OECD, 2007b). Guidance on grouping and read-across or how these principles can be implemented have been published by ECHA (2008, 2013), OECD (2014) and EFSA (2019b).

The read-across approach is a stepwise process, starting with defining the endpoint that is evaluated, and establishing a working hypothesis and justification for read-across. The structural similarities of the UHM or DHM and the source chemicals (for instance the parent or other chemicals) need to be assessed and the dissimilarities must be considered. The PPR Panel recommends to discuss the relevance of the similarities and dissimilarities identified for making use of read-across to evaluate the endpoint considered in the analysis. Based on the evaluation, it can be concluded whether the UHM or DHM is similar to the source chemicals and is expected to exert the same toxicity as the source chemicals or whether the UHM or DHM should be considered dissimilar and no conclusion on its toxic effect can be drawn.

For the QSAR as well as read-across approaches, it is essential that all steps in the process are well documented so that the adequacy and scientific validity of the approaches can be assessed.

The PPR Panel recommends to draw a final conclusion on the similarity between toxic potential of the UHM or DHM and the structurally related chemicals upon the outcome from the QSAR and read-across approaches and other available information, e.g. for *in vitro* tests, using a weight of evidence approach (EFSA Scientific Committee, 2017a).

[6] https://www.oecd.org/chemicalsafety/oecd-qsar-toolbox.htm#Guidance_Documents_and_Training_Materials_for_Using_the_Toolbox

In vitro methods. Genotoxicity testing of UHM or DHM (in case genotoxicity is not adequately covered by available data or QSAR predictions) should follow the outline of the EFSA genotoxicity testing strategies (EFSA, 2011; EFSA PPR Panel, 2016, 2017c). It is important to consider the structure of the metabolite, because this could affect the behaviour of the compound, e.g. with regard to the uptake into the cell. A possibility that a metabolite is further metabolised, needs consideration, also regarding the adequacy of the standard S9-mix. Relevant *in vitro* genotoxicity tests have been included in the OECD guideline programme.

Alternative *in vitro* test methods and strategies to measure TK processes (intestinal absorption and metabolism, hepatic transport and metabolism, etc.) (Olesti et al., 2021) and TD processes at different levels of biological organisation (molecular, biochemical, cellular, etc.) are currently being developed, e.g. EUToxRisk (https://www.eu-toxrisk.eu/), DNT IATA (EFSA PPR Panel, 2021), EPA list (US EPA, 2021) and OECD guidelines. As an excellent example, skin sensitisation can indicate the progress in development and application of non-animal test methods (OECD, 2020b). Concerning TK processes, only measurement of skin absorption using *in vitro* methods is included in OECD guidelines (OECD TG 428; OECD, 2004) and *in vitro* CYP induction in human hepatocytes has been validated (Bernasconi et al., 2019). Regarding TD processes, most OECD-approved guidelines are related to endocrine disruptor screening (e.g. OECD TG 455, 456, 458, 493; OECD, 2016, 2011, 2020c, 2015, respectively)). Many cell-based test panels for target organ toxicities have been developed and many of them are available commercially. Examples of such panels include microphysiological test systems (Marx et al., 2020), tissue chip technologies (Rusyn and Roth, 2021), imaging-based (cell painting) phenotypic profiling (Hussain et al., 2020; Nyffeler et al., 2020) and *in vitro* cell stress panels (Hatherell et al., 2020). A recent example of the use of such *in vitro* information for toxicological evaluation is coumarin (Baltazar et al., 2020). Recently, the USA consumer product safety commission (CPSC) drafted a guidance to clarify to stakeholders the informational requirements of CPSC to evaluate NAMs and IATAs, e.g. related to concordance and reproducibility data, false-positive and false-negative rates, applicability domain, test endpoint and validation studies (CPSC, 2021). The recently adopted 11th Revision of the SCCS Note for Guidance for assessing the safety of cosmetic ingredients (SCCS, 2021) is also a valuable document to be consulted, considering that animal testing in banned in Europe for these kinds of products and alternative approaches are described. This or similar guidance from regulatory agencies will facilitate the use of data obtained by NAMs and IATAs to provide information on the toxicity of UHMs and DHMs.

Non-guideline *in vitro* test methods are increasingly being used for providing information as part of the weight of scientific evidence in characterising mechanism or mode of action or a hazard in order to better understand the toxicity observed with parent and metabolite(s) and used in risk-based decision-making. Furthermore, the results of these *in vitro* methods are important as building blocks for molecular initiating and key events and their relationships in developing adverse outcome pathways (AOPs) which would be useful in the frame of IATA schemes, provided they are scientifically justifiable and suitable for risk assessment purposes. However, the PPR Panel recommends to consider the use *in vitro* methods, as well as experimental evidence obtained with the approaches mentioned above, on a case-by-case basis and by assessing the relevance and reliability of the alternative approaches (OECD, 2018a). When assessing the internal validity of the study, the criteria set by National Toxicology Programme-Office of Heath Assessment and Translation (NTP-OHAT) is recommended to be applied (NTP, 2019).

6.2.2. *In vivo* methods

If the above *in silico* and *in vitro* assessments do not allow for the convincing conclusion as to the absence or insignificant probability of a human hazard, toxicological relevance of DHM and UHM can be based on *in vivo* testing (EFSA PPR Panel, 2016) as a last resort. The testing strategy should take into account the toxicological profile of the parent compound and the possibility to explore specific hazards.

For toxicity assessment of DHM and UHM, for which safety concerns cannot be excluded by other means and methods, *in vivo* tests may be the last option. Because of the unknown toxicity of the human metabolites of concern and the need to set health-based guidance values, a 90-day rat study (OECD TG 409; OECD, 1998) on the metabolite can be an option. This is unless an alternative study would better reflect the most reasonable comparison by using lower number of animals. The PPR Panel recommends to assure comparability of testing conditions by using the same strain of laboratory animals and the same experimental conditions used for the parent.

6.3. Risk characterisation

Analogously to cosmetic ingredients (SCCS, 2021), an approximate risk assessment to human metabolites could be based on internal doses, if considered justified.

For this assessment of the internal dose, the PPR Panel highly recommends building a generic PBK model to estimate the internal human exposure to the parent compound and on this basis also to the metabolites of concern, following the OECD Guidance document on the characterisation, validation and reporting of PBK models for regulatory purposes (OECD Guidance Document No331, 2021).

In the absence of data and in judiciously selected cases, the use of an adjusted internal threshold of toxicological concern (TTC) (Partosch et al., 2014) might be a suitable approach in case of a very low exposure (below 1 μg/kg bw per day). Work is ongoing to develop robust internal TTC thresholds, especially in the area of cosmetics; meanwhile the SCCS has proposed an interim conservative internal TTC of 1 μmol/L plasma concentration, which is supported by the published experience on pharmaceuticals, a literature review of non-drug chemical/receptor interactions and analysis of ToxCast™ data (SCCS, 2021). This internal TTC value applies only to non-genotoxic substances.

If a human metabolite is considered to be covered by toxicological evaluation of a parent compound, also its risk assessment is covered.

When the toxicity of a human metabolite is not covered by the parent compound data, even if only a limited toxicological database exists, available data may still be useful for risk assessment. In such cases an additional UF might apply to setting of health-based guidance value for a human metabolite (EFSA Scientific Committee, 2012). Additional UFs, usually in the range of 3–10, may be applied to account for limited or missing data, e.g. to extrapolate from an LOAEL to an NOAEL or from a short-term to a long-term study. The value of the UFs must be determined by expert judgement on a case-by-case basis. As an alternative approach to the application of an additional uncertainty factor for extrapolation of an LOAEL to an NOAEL, the data from the critical study may be modelled to derive a BMDL as a potential reference point to be used for the derivation of the health-based guidance value.

7. Recommendations for the future

7.1. Relevance of comparative metabolism studies to other areas

The PPR recommends to:

- **reflect upon the metabolites or degradation products formed in groundwater or as residues in plants and/or livestock studies**. Some DHM and UHM might be identified in residues and livestock or groundwater. Information obtained from testing the metabolites of pesticide active substance could help in assessing their relevance in this area.
- **consider the formation of reactive metabolites**. The comparative *in vitro* metabolism studies may also suggest the formation of reactive (see Appendix E), potentially toxic metabolites, which requires tentative identification. It is of importance to use this information in studies concerning residues and their potential toxicity (crops, food-producing animals, food processing, etc.) for targeted search and tentative identification and for checking whether there could be additional sources of human exposure.
- **explore the use of *in vitro* metabolism studies in other areas such as residues**. For example, with the aim to replace *in vivo* livestock metabolism studies and reduce animal testing (Montesissa et al., 1996). It is noted that OECD guideline TG already exists for *in vitro* metabolism using fish S9 and fish hepatocytes (OECD TG 319A and B, OECD, 2018c,d).

7.2. Human relevance of toxicity effects, within a weight of evidence approach

The PPR recommends to:

- **Consider contribution of comparative *in vitro* metabolism studies for the assessment of the human relevance of toxic effects observed in animals**. The species-specific formation of metabolites can also be useful as additional line of evidence for assessing the human relevance of toxicity effects observed in animal species. The use of hepatocytes from all animal species used in toxicological studies (i.e. rat, mouse, dog and rabbit) would be strongly recommended to address human relevance of adverse effects observed in animal species, within a weight of evidence approach (EFSA Scientific Committee,

2017a). It is clear that a species-specific effect e.g. in rats can be only tentatively attributed to the concurrent formation of a 'unique' metabolite and further mechanistic studies on the plausibility of the hypothesis are needed. However, this can flag a potential issue about the relevance of deriving the reference values from toxicological studies from that species. Information from comparative *in vitro* metabolism can be added to dose-temporal concordance table within the ICPS frameworks for assessment of mode of actions and human relevance.

- **For a pesticide active substance re-evaluation, comparative *in vitro* metabolism studies should be considered under the renewal process as a tool to identify the relevant species for setting health-based guidance values**. The species-specific formation of metabolites may be also important in the discussion on which animal species the setting of reference values should be based. In case rat and human have a similar metabolic profile, while the dog is characterised by a peculiar metabolic pattern with the lowest NOAEL observed in the 1-year dog study, then it should be proven that the metabolites formed only in dogs are those responsible for the observed effects and the dog is therefore not the relevant species for human risk assessment. This recommendation is in line with the EU Regulation 283/2013.
- **Consider comparative *in vitro* metabolism studies as a tool to improve the design of toxicological data set and reduce animal testing (3Rs)**. For new pesticide active substances safety assessment, specific studies should be performed on defined species and this is currently not subject to changes. However, even if it is not the primary goal of this Scientific Opinion, it is worth mentioning that in other areas (e.g. drug development) comparative *in vitro* metabolism studies provide useful information to select the most relevant species in toxicological studies for new developed test item. In this case, the kinetic information is collected before the other toxicity studies are conducted and the most appropriate species is selected, with a substantial reduction in the number of animals tested (EMA, 2009; FDA, 2020). A first attempt to follow this line is given in Commission Regulation (EU) No 283/2013, where it is stated: *5.5 Long term toxicity and carcinogenicity – If comparative metabolism data indicate that either rat or mouse is an inappropriate model for human cancer risk assessment, an alternative species shall be considered*. This consideration could be subject to discussion when it comes to the update of the data requirements for pesticides in the future, as this is in line with the animal welfare policy in the EU. EU legislation on the protection of animals used for scientific purposes identifies guiding principles for the ethical use of animals in experimental procedures (Directive 2010/63/EU), based on the 3Rs principles (replace, reduce and refine). As part of the DG-Sante, EURL ECVAM has a specific mandate specified in EU legislation and includes a number of duties to advance the 3Rs principles of animal testing. In this context and based on the internationally established principles of the 3Rs, EFSA is committed to promote use of data derived from alternative approaches or NAMs, where possible (EFSA, 2014). The commitment also applies to EMA, having established a specific WG on the Application of the 3Rs in Regulatory Testing of Medicinal Products (EMA, 2017a,b) and to ECHA (2017).

7.3. Human body fluid and tissues

The PPR recommends to:

- **Consider the possibility to use mandatory field studies for efficacy and residues as an occasion to collect urine samples in operators**. As explained in the terms of reference, following implementation of Regulation (EU) 283/2013, analytical methods are required for the substances that are included in the residue definition for human body fluids and tissues. In the development of plant protection products, the applicant will perform residues field studies in which the product is handled by **operators** who are potentially exposed to the active substance in the formulated product. It would be possible to obtain biological specimens from the exposed people, e.g. fractionated collection of urine over 24 h. Provided that the specific metabolites monitored in urine are identical in *in vitro* studies in human test system, the specimens could be examined for the presence and amounts of the parent compound and respective specific metabolites, thus providing information on the metabolites occurring *in vivo* in humans. Although the amount of active substance and metabolites is expected to be low, the structure of potential human metabolites is already known from *in vitro* metabolism studies. Confirming the *in vitro* human metabolites *in vivo*

would contribute to the confidence in the conducted human risk assessment and to propose proper residue definition for human body fluids and tissues. The PPR noted that all ethical considerations should be taken when conducting these studies and analyses.

- Consider **follow-up biomonitoring studies** on residents and workers. It could provide additional relevant information on the metabolism of active substances in humans. Further information on human biomonitoring following exposure to pesticides can be found in an EFSA supporting publication (RPA, HSL and IEH, 2017).

7.4. Additional parameters to be measured in the *in vitro* metabolism study

The PPR recommends to:

- **Consider the possibility to evaluate in the same experiment the *in vitro* metabolic clearance, Km and Vmax**. This is possible by quantifying the concentration of parent compound in the *in vitro* test systems across experimental concentrations and over time points. Although this information is not strictly needed for the identification of potential UHM or DHM, it may be useful for future applications in TK modelling of active substances of plant protection products. The PPR Panel noted that for adequate determination of Km and Vmax, experimental conditions will need to be optimised, such as determination of the optimal incubation time point and inclusion of more than three concentrations.
- Pay attention to the **metabolite profile of both phase I and II metabolites**, which may suggest the formation of potential reactive intermediates and metabolites, such as glutathione or sulfate-conjugates and their downstream products.

7.5. Non-hepatic metabolism

The PPR recommends to:

- **Address non-hepatic metabolism when information from the *in vivo* rat TK study can indicate that non-hepatic metabolism plays a major role**. When *in vivo* rat metabolites formed are predominantly products of esterases (e.g. carboxylesterases (CEs)) or other biotransformation enzymes (e.g. monoamino oxidases (MAO)), and when the conversion rate is high (i.e. practically all the parent is expected to be initially hydrolysed on the basis of the rat *in vivo* TK data), one should consider using both a liver-based tool and a suitable extrahepatic tool relevant for the exposure route (i.e. an enterocyte test system for oral exposure). If hydrolysis is very rapid in the enterocyte test system, systemic exposure is expected to be mainly to the principal product(s) of hydrolysis. An enterocyte testing system (or analogous) is needed to evaluate the extent of first-pass metabolism and the extent of exposure of the liver to the predominant hydrolysis product. Consequently, the PPR Panel recommends to use the principal hydrolysed metabolite as a substrate for the liver-based test system to evaluate its further phase I and II metabolic fates. Interspecies differences in liver metabolism of pesticides play a critical role to depict their comparative metabolism, kinetics and associated consequences on toxicity (detoxification/bioactivation). In addition, intestinal metabolism plays an important role in the first pass metabolism and depicting interspecies differences related to such processes can support the development of quantitative *in vitro–in vivo* extrapolation and PBK models. These can be further integrated in the risk assessment process to take into account the species-specific ADME, differences in internal dose at the target organ and toxicity in a quantitative manner.
- **Development of *in vitro* systems to investigate such species differences in intestinal metabolism**. The test system should be able to express intestinal phase I/phase II enzymes and transporters in order to identify isoform-specific metabolism (human enzymes and corresponding orthologous in experimental animals) and the relative contribution compared to the liver. The performance of case studies will be necessary to validate these test systems using pesticide active substances with known metabolism. This will provide, in parallel to hepatic metabolism, a model to characterise inter-species differences in the first pass metabolism.

- **Development of a database** providing quantitative information on the expression and activities of phase I, phase II enzymes and transporters in organs with biotransformation capacity, including the liver, intestine and kidney at the isoform level in test species and humans.
- **Development of quantitative *in vitro–in vivo* extrapolation (QIVIVE) and PBK models** for each species integrating liver, intestinal, kidney metabolism, etc., transport and excretion while testing and validating them with pesticide active substances with known and unknown ADME properties.

7.6. Study design of the *in vivo* rat ADME study

The PPR recommends to:

- **improve the design of the *in vivo* rat ADME study (OECD 417; OECD, 2010)**. The design of the *in vivo* rat ADME study does not always allow proper estimation of TK parameters for their application in the construction of a PBK model, if only total radioactivity was measured. Critical factors for estimating TK parameters are dose (multiple), concentration of a parent and principal metabolites (plasma, excreta, tissues) and time (multiple points). TK parameters to be estimated are absorption, bioavailability (intestinal, post-hepatic), protein binding, volume of distribution, metabolite formation (fractional), clearances, half-life and excretion pattern in biological fluids.
- In addition to the measurement of total radioactivity, **specific methods to measure parent and metabolites separately should be applied**, not only at the final time point but also at intermediate time points.

7.7. PBK modelling

The PPR recommends to:

- **build a generic PBK model to estimate the internal exposure to the parent compound and to the metabolites of concern in humans and test species**. As described in this Scientific Opinion, the data generated from comparative *in vitro* metabolism studies are used to assess interspecies differences in metabolite formation in a qualitative and quantitative manner. A direct translation of such *in vitro* data to the *in vivo* situation is not straightforward, and more insight and better understanding is needed here. For an optimal translation of animal toxicity data to the human situation regarding interspecies differences in kinetics, quantitative insight into interspecies differences in time-dependent internal concentrations would be needed, as an ultimate aim, for both the parent chemical and its metabolites at defined exposure scenarios. Such time-dependent internal exposures do not solely depend on biotransformation in the liver, but also on other kinetic processes, which can differ extensively between species, including uptake and metabolism from the gastrointestinal tract, plasma protein binding, plasma: tissue partitioning, excretion rate and for certain chemicals extrahepatic metabolism and transporter protein-mediated kinetics. These chemical- and species-specific kinetic processes can be integrated in PBK models, which can be used to predict/simulate internal concentrations at defined exposure scenarios. For the development of such PBK models, *in vitro*, *in silico* and *in vivo* kinetic data can be used to obtain chemical-specific input parameters. In that regard, the *in vitro* biotransformation data produced in the context of the comparative *in vitro* metabolism studies described in this Scientific Opinion can provide important chemical-specific input data for PBK model development in both test species and humans.

In the context of this Scientific Opinion, such PBK-TD models can be used to (1) quantitatively assess species differences in internal exposure to DHM and (2) for quantitative *in vitro* to *in vivo* extrapolation of NAM-based toxicity data (mainly from *in vitro* studies) obtained for UHM and/or DHM, estimating external exposure to the pesticide active substance.

In 2021, the OECD released a 'Guidance document on the characterization, validation and reporting of PBK models for regulatory purposes' OECD Guidance Document No. 331), with 'the goal of increasing confidence in the use of these models parameterized with data from *in vitro* and *in silico* methods'. Application of high-quality and robust input data using standardised approaches, e.g. as proposed in this opinion, is considered critical for increased confidence in such PBK models based on *in vitro*- and *in silico*-derived *chemical-specific kinetic parameters.*

7.8. Use of pathway-related uncertainty factors *and chemical-specific adjustment factors*

The PPR recommends to:

- **use pathway-related uncertainty factors and chemical-specific adjustment factors instead of default factor by investigating isoform-specific metabolism**. The WHO proposed the use of chemical-specific adjustment factors (CSAF) when chemical-specific data are available or pathway-related UFs for the human TK dimension, when the metabolic pathway is known in humans (Renwick, 1993; IPCS, 2005; Dorne and Renwick, 2005; EFSA Scientific Committee, 2012; Bhat et al., 2017; Dorne, 2010). Such CSAFs have been applied in the context of EFSA's risk assessment for cadmium through a meta-analysis of TD biomarkers combined with PBK modelling (EFSA CONTAM, 2009). A further example is the phosphates as food additives (E 338- 341, E343, E450 – 452) where mechanistic information and a probabilistic approach was used for deriving a CSAF reducing it from the default of 100 to 4 (Smeraldi et al., 2020). In addition, human variability in TK processes using the pharmaceutical database has been investigated for specific isoforms of phase I (e.g. CYP3A4, CYP2C9, CYP2C19, CYP2D6; PON1) and phase II enzymes (e.g. GSTs) by means of meta-analysis of parameters reflecting acute (Cmax) and chronic exposure (AUC) (Testai et al., 2021). The resulting variability distributions were combined with isoform-specific *in vitro* metabolism data for food relevant chemicals (e.g. pesticide active substances, contaminants) to predict their blood concentrations in humans by means of PBK and QIVIVE modelling. Finally, PBK-TD modelling has also been developed for specific pesticide active substances and endpoints such as acetyl-cholinesterase inhibition (Kasteel et al., 2020; Testai et al., 2021). Such developments will allow the integration of quantitative mode of action and AOP information in chemical risk assessment. Since the AOP framework does not include the TK dimension, the Aggregated Exposure Pathway (AEP) concept has been proposed to integrate such dimension and EFSA's Scientific Committee has recommended to further develop knowledge on MoA, AOP and AEPs (EFSA Scientific Committee, 2019). In this context, it is foreseen that the identification of metabolic pathways for pesticide active substances and other chemicals in test species and humans using *in vitro* assays combined with chemical-specific data and PBK modelling can allow to explore integration of AEPs and AOPs and refine default UFs to science-based UFs. However, according to the current EU legislation on pesticide active substances, the default UF should not be lowered below a 100-fold.

7.9. Overarching guidance on TK and NAMs for hazard characterisation

The PPR recommends to:

- **develop an EFSA harmonised guidance document on the use of TKs and metabolism data in chemical risk assessment**. The PPR panel recommends the development of a guidance document on the use of TKs and metabolism data in chemical risk assessment by EFSA's Scientific Committee. This will provide harmonised methodologies for regulated products and contaminants using tiered approaches depending on the regulatory context and data availability. These include the use of basic ADME data, the development of PBK and QIVIVE models and the integration of quantitative metrics on interspecies differences and human variability in TK processes including subgroups of the human population (children, inter-ethnic differences, individuals harbouring the same genetic polymorphisms) (OECD, 2021; Testai et al., 2021). Finally, such a guidance document will provide a basis to derive reference points and health-based guidance values on an internal dose basis for risk characterisation through illustrative case studies.
- **To develop an EFSA harmonised guidance document on the use of NAMs in chemical risk assessment**. The PPR panel recommends further integration of NAMs in chemical risk assessment (e.g. high-throughput techniques, *in vitro*, OMICs and in silico methods such as PBK and QIVIVE models) as currently investigated worldwide (OECD, US EPA, EFSA). Relevant examples include large open-source platforms such as the US-EPA CompTox Chemicals Dashboard, ECHA's REACH database and EFSA's OpenFoodTox and new QSAR models predicting TK and TD properties of chemicals such as the VEGA hub and the OPERA suite from US-EPA (Dawson et al., 2021; Williams et al., 2021). In practice, case studies could explore how to integrate NAMs for a range of regulatory contexts (pesticide active substances, food additives, contaminants) using the existing OECD Harmonised Template (OHT) 201 for

structuring mechanistic data and the OHT 58 for kinetic data to implement these tools in a structured and transparent manner.

7.10 Guidance on comparative *in vitro* metabolism

The PPR Panel recommends performing a critical assessment of the studies submitted following the experimental design as introduced in this opinion to gain experience and to be used in the drafting of a future guidance on comparative *in vitro* metabolism.

References

de Albuquerque NCP, Carrão DB, Habenschus MD and de Oliveira ARM, 2018. Metabolism studies of chiral pesticides: a critical review. Journal of Pharmaceutical and Biomedical Analysis, 147, 89–109. https://doi.org/10.1016/j.jpba.2017.08.011

Anari MR and Baillie TA, 2005. Bridging cheminformatic metabolite prediction and tandem mass spectrometry. Drug Discovery Today, 10, 711–717.

Anderson S, Luffer-Atlas D and Knadler MP, 2009. Predicting circulating human metabolites: how good are we? Chemical Research in Toxicology, 22, 243–256.

Arlotto MP, Trant JM and Estabrook RW, 1991. Measurement of steroid hydroxylation reactions by high-performance liquid chromatography as indicator of P450 identity and function. Methods in Enzymology, 206, 454–462. https://doi.org/10.1016/0076-6879(91)06114-I

Baillie TA and Rettie AE, 2011. Role of biotransformation in drug-induced toxicity: influence of intra- and inter-species differences in drug metabolism. Drug Metabolism and Pharmacokinetics, 26, 15–29. https://doi.org/10.2133/dmpk.dmpk-10-rv-089

Ballard TE, Orozco CC and Obach RS, 2014. Generation of major human excretory and circulating drug metabolites using a hepatocyte relay method. Drug Metabolism and Disposition, 42, 899–902. https://doi.org/10.1124/dmd.114.057026

Baltazar MT, Cable S, Carmichael PL, Cubberley R, Cull T, Delagrange M, Dent M, Hatherell S, Houghton J, Kukic P, Li H, Lee M, Malcomber S, Middleton AM, Moxon TE, Nathanail AV, Nicol B, Pendlington R, Reynolds G, Reynolds J, White A and Westmoreland C, 2020. A next-generation risk assessment case study for coumarin in cosmetic products. Toxicological Sciences, 176, 236–252. https://doi.org/10.1093/toxsci/kfaa048

Barr JT, Lade JM, Tran TB and Dahal UP, 2019. Fraction unbound for liver microsome and hepatocyte incubations for all major species can be approximated using a single-species surrogate. Drug Metabolism and Disposition, 47, 419–423.

Bell CC, Lauschke VM, Vorrink SU, Palmgren H, Duffin R, Andersson TB and Ingelman-Sundberg M, 2017. Transcriptional, functional, and mechanistic comparisons of stem cell-derived hepatocytes, HepaRG cells, and three-dimensional human hepatocyte spheroids as predictive *in vitro* systems for drug-induced Liver injury. Drug Metabolism and Disposition, 45, 419–429. https://doi.org/10.1124/dmd.116.074369

Benigni R, Battistelli CL, Bossa C, Giuliani A, Fioravanzo E, Bassan A, Fuart GM, Rathman J, Yang C, Mostrag-Szlichtyng A, Ol S and Tcheremenskaia O, 2019. Evaluation of the applicability of existing (Q)SAR models for predicting the genotoxicity of pesticides and similarity analysis related with genotoxicity of pesticides for facilitating of grouping and read across. EFSA Supporting Publication 2019;16(3):EN-1598, 220 pp. https://doi.org/10.2903/sp.efsa.2019.EN-1598

Benigni R, Serafimova R, Parra Morte JM, Battistelli CL, Bossa C, Giuliani A, Fioravanzo E, Bassan A, Gatnik MF, Rathman J, Yang C, Mostrag-Szlichtyng A, Sacher O and Tcheremenskaia O, 2020. Evaluation of the applicability of existing (Q) SAR models for predicting the genotoxicity of pesticides and similarity analysis related with genotoxicity of pesticides for facilitating of grouping and read across: an EFSA funded project. Regulatory Toxicology and Pharmacology, 114, 104658. https://doi.org/10.1016/j.yrtph.2020.104658

Bernasconi C, Pelkonen O, Andersson TB, Strickland J, Wilk-Zasadna I, Asturiol D, Cole T, Liska R, Worth A, Müller-Vieira U, Richert L, Chesne C and Coecke S, 2019. Validation of *in vitro* methods for human cytochrome P450 enzyme induction: outcome of a multi-laboratory study. Toxicology in Vitro, 60, 212–228. https://doi.org/10.1016/j.tiv.2019.05.019

Bhat V, Meek MEB, Valcke M, English C, Boobis A and Brown R, 2017. Evolution of chemical-specific adjustment factors (CSAF) based on recent international experience; increasing utility and facilitating regulatory acceptance. Critical Reviews in Toxicology, 47, 729–749.

Buratti FM, Scardala S, Funari E and Testai E, 2011. Human glutathione transferases catalyzing the conjugation of the hepatoxin Microcystin-LR. Chemical Research in Toxicology, 24, 926–933. https://doi.org/10.1021/tx2000976

Buratti FM, Scardala S, Funari E and Testai E, 2013. The conjugation of Microcystin-RR by human recombinant GSTs and hepatic cytosol. Toxicology Letters, 219, 231–238. https://doi.org/10.1016/j.toxlet.2013.03.015

Buratti FM and Testai E, 2015. Species- and congener-differences in microcystin-LR and –RR GSH conjugation in human, rat and mouse hepatic cytosol. Toxicology Letters, 232, 133–140. https://doi.org/10.1016/j.toxlet.2014.10.020

Burton RD, Hieronymus T, Chamem T, Heim D, Anderson S, Zhu X and Hutzler JM, 2018. Assessment of the biotransformation of low-turnover drugs in the HμREL human hepatocyte coculture model. Drug Metabolism Dispose, 46, 1617–1625. https://doi.org/10.1124/dmd.118.082867

Chan TS, Yu H, Moore A, Khetani SR and Tweedie D, 2013. Meeting the challenge of predicting hepatic clearance of compounds slowly metabolized by cytochrome P450 using a novel hepatocyte model, HepatoPac. Drug Metabolism and Disposition, 41, 2024–2032. https://doi.org/10.1124/dmd.113.053397

CPSC (Consumer Product Safety Commission), 2021. Title: notice of availability: proposed guidance on alternative test methods and integrated testing approaches. Available online: https://www.regulations.gov/document/CPSC-2021-0006-0002

Dalvie D, Obach RS, Kang P, Prakash C, Loi CM, Hurst S, Nedderman A, Goulet L, Smith E, Bu HZ and Smith DA, 2009. Assessment of three human *in vitro* systems in the generation of major human excretory and circulating metabolites. Chemical Research in Toxicology, 22, 357–368. https://doi.org/10.1021/tx8004357

Dawson DE, Ingle BL, Phillips KA, Nichols JW, Wambaugh JF and Tornero-Velez R, 2021. Designing QSARs for parameters of high-throughput toxicokinetic models using open-source descriptors. Environmental Science & Technology, 55, 6505–6517. https://doi.org/10.1021/acs.est.0c06117

De Graaf IA, Olinga P, de Jager MH, Merema MT, de Kanter R, van de Kerkhof EG and Groothuis GM 2010. Preparation and incubation of precision-cut liver and intestinal slices for application in drug metabolism and toxicity studies. NatProtoc, 5 pp. 1540–1551. https://doi.org/10.1038/nprot.2010.111.

Deakin T, 2015. Radisotopic characterization as an analytical tool: current status, limitations and future challenges. Bioanalysis, 7, 541–555. https://doi.org/10.4155/bio.15.13. PMID: 25826136.

Dorne JLCM, 2010. Metabolism, variability and risk assessment. Toxicology, 268, 156–164.

Dorne JL and Renwick AG, 2005. The refinement of uncertainty/safety factors in risk assessment by the incorporation of data on toxicokinetic variability in humans. Toxicological Sciences, 86, 20–26. Epub 2005 Mar 30 PMID: 15800035. https://doi.org/10.1093/toxsci/kfi160

Dorne JLCM, Richardson J, Livaniou A, Carnesecchi E, Ceriani L, Baldin R, Kovarich S, Pavan M, Saouter E, Biganzoli F, Pasinato L, Zare Jeddi M, Robinson TP, Kass GEN, Liem AKD, Toropov AA, Toropova AP, Yang C, Tarkhov A, Georgiadis N, Di Nicola MR, Mostrag A, Verhagen H, Roncaglioni A, Benfenati E and Bassan A, 2021. EFSA's OpenFoodTox: an open source toxicological database on chemicals in food and feed and its future developments. Environment International, 146. Epub 2020 Dec 8 PMID: 33395940. https://doi.org/10.1016/j.envint.2020.106293.

ECHA (European Chemicals Agency), 2008. Guidance on Information Requirements and Chemical Safety Assessment, Chapter R.6: QSARs and Grouping of Chemicals. 134 pp. European Chemicals Agency, Helsinki, Finland. Available online: https://echa.europa.eu/documents/10162/13632/information_requirements_r6_en.pdf

ECHA (European Chemicals Agency), 2013. Grouping of Substances and Read-across Approach — An illustrative example [ECHA-13-R-02-EN]. European Chemicals Agency, Helsinki, Finland. 11 pp. Available online: http://echa.europa.eu/documents/10162/13628/read_across_introductory_note_en.pdf

ECHA (European Chemicals Agency), 2017. Non-animal approaches Current status of regulatory applicability under the REACH, CLP and Biocidal Products regulations. 163 pp. European Chemicals Agency, Helsinki, Finland. Available online: https://echa.europa.eu/documents/10162/22931011/non_animal_approcches_en.pdf

EFSA (European Food Safety Authority), 2011. Submission of scientific peer-reviewed open literature for the approval of pesticide active substances under Regulation (EC) No 1107/2009 (OJ L 309, 24.11.2009, p. 1–50). EFSA Journal 2011;9(2):2092, 49 pp. https://doi.org/10.2903/j.efsa.2011.2092

EFSA (European Food Safety Authority), 2014. Modern methodologies and tools for human hazard assessment of chemicals. EFSA Journal 2014;12(4):3638, 87 pp. https://doi.org/10.2903/j.efsa.2014.3638

EFSA (European Food Safety Authority), 2016. Outcome of the pesticides peer review meeting on general recurring issues in mammalian toxicology. EFSA Supporting Publications 2016;13(8):1074, 22 pp. https://doi.org/10.2903/sp.efsa.2016.EN-1074

EFSA (European Food Safety Authority), 2016b. Guidance on the establishment of the residue definition for dietary risk assessment. EFSA Supporting Publications 2014;14(12):4549, 45 pp. https://doi.org/10.2903/j.efsa.2016.4549.

EFSA (European Food Safety Authority), 2019a. EFSA Workshop on *in vitro* comparative metabolism studies in regulatory pesticide risk assessment. EFSA Supporting Publications 2019;16(4):1618, 44 pp. https://doi.org/10.2903/sp.efsa.2019.EN-1618

EFSA (European Food Safety Authority), Bura L, Friel A, Magrans JO, Parra-Morte JM and Szentes C, 2019b. Guidance of EFSA on risk assessments for active substances of plant protection products that have stereoisomers as components or impurities and for transformation products of active substances that may have stereoisomers. EFSA Journal 2019;17(8):5804, 33 pp. https://doi.org/10.2903/j.efsa.2019.5804

EFSA (European Food Safety Authority), 2021. Training manual on MetaPath Data Evaluation Record (DER) Composer (Version 1). Zenodo, https://doi.org/10.5281/zenodo.4751079

EFSA CONTAM Panel (EFSA Panel on Contaminants in the Food Chain), 2009. Scientific Opinion on a request from the European Commission on cadmium in food. EFSA Journal, 2009;980, 139 pp. https://doi.org/10.2903/j.efsa.2009.980

EFSA PPR Panel (EFSA Panel on Plant Protection Products and their Residues), 2016. Guidance on the establishment of the residue definition for dietary risk assessment. EFSA Journal 2016;14(12):4549, 129 pp. https://doi.org/10.2903/j.efsa.2016.4549

EFSA PPR Panel (EFSA Panel on Plant Protection Products and their Residues), 2021. Draft Scientific Opinion on Development of Integrated Approaches to Testing and Assessment (IATA) on developmental neurotoxicity (DNT) risk assessment.

EFSA Scientific Committee, 2011. Scientific Opinion on genotoxicity testing strategies applicable to food and feed safety assessment. EFSA Journal 2011;9(9):2379, 69 pp. https://doi.org/10.2903/j.efsa.2011.2379

EFSA Scientific Committee, 2012. Guidance on selected default values to be used by the EFSA Scientific Committee, Scientific Panels and Units in the absence of actual measured data. EFSA Journal 2012;10(3):2579, 32 pp. https://doi.org/10.2903/j.efsa.2012.2579

EFSA Scientific Committee, Hardy A, Benford D, Halldorsson T, Jeger MJ, Knutsen HK, More S, Naegeli H, Noteborn H, Ockleford C, Ricci A, Rychen G, Schlatter JR, Silano V, Solecki R, Turck D, Benfenati E, Chaudhry QM, Craig P, Frampton G, Greiner M, Hart A, Hogstrand C, Lambre C, Luttik R, Makowski D, Siani A, Wahlstroem H, Aguilera J, Dorne J-L, Fernandez Dumont A, Hempen M, Valtueña Martínez S, Martino L, Smeraldi C, Terron A, Georgiadis N and Younes M, 2017a. Scientific Opinion on the guidance on the use of the weight of evidence approach in scientific assessments. EFSA Journal 2017;15(8):4971, 69 pp. https://doi.org/10.2903/j.efsa.2017.4971

EFSA Scientific Committee, Hardy A, Benford D, Halldorsson T, Jeger MJ, Knutsen HK, More S, Younes M, Naegeli H, Noteborn H, Ockleford C, Ricci A, Rychen G, Schlatter JR, Silano V, Solecki R, Turck D, Bresson J-L, Griffin J, Hougaard Benekou S, van Loveren H, Luttik R, Messean A, Penninks A, Ru G, Stegeman JA, van der Werf W, Westendorf J, Woutersen RA, Barizzone F, Bottex B, Lanzoni A, Georgiadis N and Alexander J, 2017b. Guidance on biological relevance. EFSA Journal 2017;15(8):4970, 73 pp. https://doi.org/10.2903/j.efsa.2017.4970

EFSA Scientific Committee, Hardy A, Benford D, Halldorsson T, Jeger M, Knutsen HK, More S, Naegeli H, Noteborn H, Ockleford C, Ricci A, Rychen G, Silano V, Solecki R, Turck D, Younes M, Aquilina G, Crebelli R, Gürtler R, Hirsch-Ernst KI, Mosesso P, Nielsen E, van Benthem J, Carfì M, Georgiadis N, Maurici D, Parra Morte J and Schlatter J, 2017c. Scientific Opinion on the clarification of some aspects related to genotoxicity assessment. EFSA Journal 2017;15(12):5113, 25 pp. https://doi.org/10.2903/j.efsa.2017.5113

EFSA Scientific Committee, More SJ, Bampidis V, Benford D, Bennekou SH, Bragard C, Halldorsson TI, Hernández-Jerez AF, Koutsoumanis K, Naegeli H, Schlatter JR, Silano V, Nielsen SS, Schrenk D, Turck D, Younes M, Benfenati E, Castle L, Cedergreen N, Hardy A, Laskowski R, Leblanc JC, Kortenkamp A, Ragas A, Posthuma L, Svendsen C, Solecki R, Testai E, Dujardin B, Kass GEN, Manini P, Jeddi MZ, Dorne JLCM and Hogstrand C, 2019. Guidance on harmonised methodologies for human health, animal health and ecological risk assessment of combined exposure to multiple chemicals. EFSA Journal 2019;17(3):5634, 77 pp. https://doi.org/10.2903/j.efsa.2019.5634

EMA (European Medicinal Agency), 2009. ICH guideline M3(R2) on non-clinical safety studies for the conduct of human clinical trials and marketing authorisation for pharmaceuticals. Available online: https://www.ema.europa.eu/en/documents/scientific-guideline/ich-guideline-m3r2-non-clinical-safety-studies-conduct-human-clinical-trials-marketing-authorisation_en.pdf

EMA (European Medicinal Agency), 2017a. Joint CVMP/CHMP Working group on the Application of the 3Rs in Regulatory Testing of Medical Products. Available online: https://www.ema.europa.eu/en/committees/working-parties-other-groups/chmp/expert-group-3rs

EMA (European Medicinal Agency), 2017b. Non-animal approaches - Current status of regulatory applicability under the REACH, CLP and Biocidal Products regulations Report. Available online: https://echa.europa.eu/documents/10162/22931011/non_animal_approcches_en.pdf/87ebb68f-2038-f597-fc33-f4003e9e7d7d

EUToxRisk. Available online: https://www.eu-toxrisk.eu/

FDA (Food and Drug Administration), 2018. Bioanalytical method validation. Guidance for industry. US Department of Health and Human Services. Food and Drug Administration, Center for Drug Evaluation and Research, Center for Biologics Evaluation and Research, Center for Veterinary Medicine. Available online: https://www.fda.gov/files/drugs/published/Bioanalytical-Method-Validation-Guidance-for-Industry.pdf

FDA (Food and Drug Administration), 2020. Guidance for industry: safety testing of drug metabolites. Available online: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/safety-testing-drug-metabolites

Foti RS, Wienkers LC and Wahlstrom JL, 2010. Application of cytochrome P450 drug interaction screening in drug discovery. Combinatorial Chemistry and High Throughput Screening, 13, 145–158. https://doi.org/10.2174/138620710790596718

Glatt H, 2000. Sulfotransferases in the bioactivation of xenobiotics. Chemico-Biological Interactions, 129, 141–170. https://doi.org/10.1016/s0009-2797(00)00202-7

Godoy P, Hewitt NJ, Albrecht U, and Hengstler JG, 2013. Recent advances in 2D and 3D in vitro systems using primary hepatocytes, alternative hepatocyte sources and non-parenchymal liver cells and their use in investigating mechanisms of hepatotoxicity, cell signaling and ADME. Archives Toxicology, 87, 1315–1530. https://doi.org/10.1007/s00204-013-1078-5. Epub 2013 Aug 23. PMID: 23974980; PMCID: PMC3753504.

Gough A, Soto-Gutierrez A, Vernetti L, Ebrahimkhani MR, Stern AM and Taylor DL, 2021. Human biomimetic liver microphysiology systems in drug development and precision medicine. Nature Reviews Gastroenterology and Hepatology, 18, 252–268. https://doi.org/10.1038/s41575-020-00386-1

Gouliarmou V, Lostia AM, Coecke S, Bernasconi C, Bessems J, Dorne JL, Ferguson S, Testai E, Remy UG, Brian Houston J, Monshouwer M, Nong A, Pelkonen O, Morath S, Wetmore BA, Worth A, Zanelli U, Zorzoli MC and Whelan M, 2018. Establishing a systematic framework to characterise *in vitro* methods for human hepatic metabolic clearance. Toxicology in Vitro, 53, 233–244. https://doi.org/10.1016/j.tiv.2018.08.004

Gouliarmou V, Pelkonen O and Coecke S, 2015. Differentiation-promoting medium additives for hepatocyte cultivation and cryopreservation. Methods in Molecular Biology, 1250, 143–159. https://doi.org/10.1007/978-1-4939-2074-7_10

Guguen-Guillouzo C and Guillouzo A, 2010. General review on *in vitro* hepatocyte models and their applications. Methods in Molecular Biology, 640, 1–40. https://doi.org/10.1007/978-1-60761-688-7_1

Hatherell S, Baltazar MT, Reynolds J, Carmichael PL, Dent M, Li H, Ryder S, White A, Walker P and Middleton AM, 2020. Identifying and characterizing stress pathways of concern for consumer safety in next-generation risk assessment. Toxicological Sciences, 176, 11–33. https://doi.org/10.1093/toxsci/kfaa054

Honkakoski P and Negishi M, 1997. The structure, function, and regulation of cytochrome P450 2A enzymes. Drug Metabolism Reviews, 29, 977–996. https://doi.org/10.3109/03602539709002240

Houston JB and Kenworthy KE, 2000. *In vitro-in vivo* scaling of CYP kinetic data not consistent with the classical Michaelis-Menten model. Drug Metabolism and Disposition, 28, 246–254. PMID: 10681367.

Hsu F, Chen YC and Broccatelli F, 2021. Evaluation of tissue binding in 3 tissues across 5 species, and prediction of volume of distribution from plasma protein and tissue binding. Drug Metabolism and Disposition.

Hussain F, Basu S, Heng JJH, Loo LH and Zink D, 2020. Predicting direct hepatocyte toxicity in humans by combining high-throughput imaging of HepaRG cells and machine learning-based phenotypic profiling. Archives of Toxicology, 94, 2749–2767. https://doi.org/10.1007/s00204-020-02778-3

International Programme on Chemical Safety & IPCS Workshop on Incorporating Uncertainty and Variability into Risk Assessment (2000 : Berlin, Germany), 2005. Chemical-specific adjustment factors for interspecies differences and human variability: guidance document for use of data in dose/concentration-response assessment. World Health Organization. Available online: https://apps.who.int/iris/handle/10665/43294

Kasteel EE, Nijmeijer SM, Darney K, Lautz LS, Dorne JLC, Kramer NI and Westerink RH, 2020. Acetylcholinesterase inhibition in electric eel and human donor blood: an in vitro approach to investigate interspecies differences and human variability in toxicodynamics. Archives of Toxicology, 94, 4055–4065.

Keyvanpour MR and Shirzad MB, 2021. An analysis of QSAR research based on machine learning concepts. Current Drug Discovery Technologies, 18, 17–30. https://doi.org/10.2174/1570163817666200316104404

Kolanczyk RC, Schmieder P, Jones WJ, Mekenyan OG, Chapkanov A, Temelkov S, Kotov S, Velikova M, Kamenska V, Vasilev K and Veith GD, 2012. MetaPath: an electronic knowledge base for collating, exchanging and analyzing case studies of xenobiotic metabolism. Regulatory Toxicology and Pharmacology. 63, 84–96. https://doi.org/10.1016/j.yrtph.2012.02.013. Epub 2012 Mar 7. PMID: 22414578.

Kramer NI, Di Consiglio E, Blaauboer BJ and Testai E, 2015. Biokinetics in repeated-dosing in vitro drug toxicity studies. Toxicology in Vitro, 30, 217–224. https://doi.org/10.1016/j.tiv.2015.09.005

Kratochwil NA, Meille C, Fowler S, Klammers F, Ekiciler A, Molitor B, Simon S, Walter I, McGinnis C, Walther J, Leonard B, Triyatni M, Javanbakht H, Funk C, Schuler F, Lavé T and Parrott NJ, 2017. Metabolic profiling of human long-term liver models and hepatic clearance. American Association of Pharmaceutical Scientists Journal, 19, 534–550. https://doi.org/10.1208/s12248-016-0019-7

Lassila T, Rousu T, Mattila S, Chesné C, Pelkonen O, Turpeinen M and Tolonen A, 2015. Formation of GSH-trapped reactive metabolites in human liver microsomes, S9 fraction, HepaRG-cells, and human hepatocytes. Journal of Pharmaceutical and Biomedical Analysis, 10, 345–351. https://doi.org/10.1016/j.jpba.2015.07.020

Lauschke VM, Hendriks DF, Bell CC, Andersson TB and Ingelman-Sundberg M, 2016. Novel 3D culture systems for studies of human liver function and assessments of the hepatotoxicity of drugs and drug candidates. Chemical Research in Toxicology, 29, 1936–1955. https://doi.org/10.1021/acs.chemrestox.6b00150

Leow JWH and Chan ECY, 2019. Atypical Michaelis-Menten kinetics in cytochrome P450 enzymes: a focus on substrate inhibition. Biochemical Pharmacology, 169, 113615. https://doi.org/10.1016/j.bcp.2019.08.017

Linkov I, Loney D, Cormier S, Satterstrom FK and Bridges T, 2009. Weight-of-evidence evaluation in environmental assessment: review of qualitative and quantitative approaches. Science of the Total Environment, 407, 5199–5205. Epub 2009 Jul 19. PMID: 19619890. https://doi.org/10.1016/j.scitotenv.2009.05.004

Louisse J, Alewijn M, Peijnenburg A, Cnubben N, Heringa M, Coecke S and Punt A, 2020. Towards harmonization of test methods for in vitro hepatic clearance studies. Toxicology in Vitro, 63, 0887–2333. https://doi.org/10.1016/j.tiv.2019.104722

Ma B, Eisenhandler R, Kuo Y, Rearden P, Li Y, Manley PJ, Smith S and Menzel K, 2017. Prediction of metabolic clearance for low-turnover compounds using plated hepatocytes with enzyme activity correction. European Journal of Drug Metabolism and Pharmacokinetics, 42, 319–326. https://doi.org/10.1007/s13318-016-0336-3

Marx U, Akabane T, Andersson TB, Baker E, Beilmann M, Beken S, Brendler-Schwaab S, Cirit M, David R, Dehne E-M, Durieux I, Ewart L, Fitzpatrick SC, Frey O, Fuchs F, Griffith LG, Hamilton GA, Hartung T, Hoeng J, Hogberg H, Hughes DJ, Ingber DE, Iskandar A, Kanamori T, Kojima H, Kuehnl J, Leist M, Li B, Loskill P, Mendrick DL, Neumann T, Pallocca G, Rusyn I, Smirnova L, Steger-Hartmann T, Tagle DA, Tonevitsky A, Tsyb S, Trapecar M, van de Water B, van den Eijnden-van Raaij J, Vulto P, Watanabe K, Wolf A, Zhou X and Roth A, 2020. Biology-inspired microphysiological systems to advance patient benefit and animal welfare in drug development. ALTEX - Alternatives to Animal Experimentation, 37, 365–394. https://doi.org/10.14573/altex.2001241.

Mateus A, Treyer A, Wegler C, Karlgren M, Matsson P and Artursson P, 2017. Intracellular drug bioavailability: a new predictor of system dependent drug disposition. Scientific Reports, 7, 1–12. https://doi.org/10.1038/srep43047

Montesissa C, Anfossi P, Van't Klooster G and Mengelers M, 1996. The use of cultured hepatocytes from goats and cattle to investigate xenobiotic oxidative metabolism. Veterinary Research Communications, 20, 449–460. https://doi.org/10.1007/BF00419182.

Muratov EN, Bajorath J, Sheridan RP, Tetko IV, Filimonov D, Poroikov V, Oprea TI, Baskin II, Varnek A, Roitberg A, Isayev O, Curtalolo S, Fourches D, Cohen Y, Aspuru-Guzik A, Winkler DA, Agrafiotis D, Cherkasov A and Tropsha A, 2020. QSAR without borders. Chemical Society Reviews, 49, 3525–3564.

Nallani GC, ElNaggar SF, Shen L and Chandrasekaran A, 2017. In vitro metabolism of [^{14}C]-benalaxyl in hepatocytes of rats, dogs and humans. Regulatory Toxicology and Pharmacology, 84, 26–34. Epub 2016 Dec 18 PMID: 27993653. https://doi.org/10.1016/j.yrtph.2016.12.006

NTP (National Toxicology Program), 2019. Handbook for Conducting a Literature-Based Health Assessment Using OHAT Approach for Systematic Review and Evidence Integration. Office of Health Assessment and Translation (OHAT). Division of the National Toxicology Program. National Institute of Environmental Health Sciences, Research Triangle Park. Available online: https://ntp.niehs.nih.gov/ntp/ohat/pubs/handbookmarch2019_508.pdf

Nyffeler J, Willis C, Lougee R, Richard A, Paul-Friedman K and Harrill JA, 2020. Bioactivity screening of environmental chemicals using imaging-based high-throughput phenotypic profiling. Toxicology and Applied Pharmacology, 389, 114876. https://doi.org/10.1016/j.taap.2019.114876

Obach RS, Kalgutkar AS and Dalvie DK, 2016. *In vitro* methods for evaluation of drug metabolism. In SL Iverson and DA Smith (eds.). Metabolite Safety in Drug Development. https://doi.org/10.1002/9781118949689.ch4

OECD (Organisation for Economic Co-operation and Development), 1998. *Test No. 409: Repeated Dose 90-Day Oral Toxicity Study in Non-Rodents*, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris. https://doi.org/10.1787/9789264070721-en

OECD (Organisation for Economic Co-operation and Development), 2004. *Test No. 428: Skin Absorption: In Vitro Method*, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris. https://doi.org/10.1787/9789264071087-en

OECD (Organisation for Economic Co-operation and Development), 2007a. Test No. 501: Metabolism in Crops, OECD. Guidelines for the Testing of Chemicals, Section 5, OECD. Publishing, Paris. https://doi.org/10.1787/9789264061835-en

OECD (Organisation for Economic Co-operation and Development), 2007b. Guidance on Grouping of Chemicals. Series on Testing and Assessment Number 80.ENV/JM/MONO(2007)28. OECD Publishing, Paris. Available online: https://www.oecd.org/

OECD (Organisation for Economic Co-operation and Development), 2010. Guidelines for the testing of chemicals (No. 417): Toxicokinetics. Adopted 22 July 2010, OECD, Paris. https://doi.org/10.1787/9789264070882-en

OECD (Organisation for Economic Co-operation and Development), 2011. *Test No. 456: H295R Steroidogenesis Assay*, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris. https://doi.org/10.1787/9789264122642-en

OECD (Organisation for Economic Co-operation and Development), 2014. Guidance on Grouping of Chemicals, OECD Series on Testing and Assessment, No. 80. OECD Publishing, Paris, https://doi.org/10.1787/9789264085831-en

OECD (Organisation for Economic Co-operation and Development), 2015. Test No. 493: Performance-Based Test Guideline for Human Recombinant Estrogen Receptor (hrER) In Vitro Assays to Detect Chemicals with ER Binding Affinity, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris. https://doi.org/10.1787/9789264242623-en

OECD (Organisation for Economic Co-operation and Development), 2016. Test No. 455: Performance-Based Test Guideline for Stably Transfected Transactivation In Vitro Assays to Detect Estrogen Receptor Agonists and Antagonists, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris. https://doi.org/10.1787/9789264265295-en

OECD (Organisation for Economic Co-operation and Development), 2018a. Guidance Document on Good In vitro Method Practices (GIVIMP), OECD Series on Testing and Assessment, No. 286. OECD Publishing, Paris. https://doi.org/10.1787/9789264304796-en

OECD (Organisation for Economic Co-operation and Development), 2018b. Guidance document on the determination of in vitro intrinsic clearance using cryopreserved hepatocytes (RT-HEP) or liver S9 sub-cellular fractions (RT-S9) from rainbow trout and extrapolation to in vivo intrinsic clearance, OECD Environment, Health and Safety Publications, Series on Testing and Assessment No. 280.

OECD (Organisation for Economic Co-operation and Development), 2018c. Test No. 319A: Determination of in vitro intrinsic clearance using cryopreserved rainbow trout hepatocytes (RT-HEP), OECD Guidelines for the Testing of Chemicals, Section 3. OECD Publishing, Paris. https://doi.org/10.1787/9789264303218-en

OECD (Organisation for Economic Co-operation and Development), 2018d. Test No. 319B: Determination of in vitro intrinsic clearance using rainbow trout liver S9 sub-cellular fraction (RT-S9), OECD Guidelines for the Testing of Chemicals, Section 3, OECD Publishing, Paris. https://doi.org/10.1787/9789264303232-en

OECD (Organisation for Economic Co-operation and Development), 2020a. Overview of Concepts and Available Guidance related to Integrated Approaches to Testing and Assessment (IATA), OECD Series on Testing and Assessment, No. 329, Environment, Health and Safety, Environment Directorate, OECD.

OECD (Organisation for Economic Co-operation and Development), 2020b. Draft OECD Guideline Defined Approaches for Skin Sensitisation. Available online: https://www.oecd.org/env/ehs/testing/GL%20DASS_22Sep2019v2.pdf

OECD (Organisation for Economic Co-operation and Development), 2020c. *Test No. 458: Stably Transfected Human Androgen Receptor Transcriptional Activation Assay for Detection of Androgenic Agonist and Antagonist Activity of Chemicals*, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris. https://doi.org/10.1787/9789264264366-en

OECD (Organisation for Economic Co-operation and Development), 2021. Guidance document on the characterisation, validation and reporting of Physiologically Based Kinetic (PBK) models for regulatory purposes, OECD. Series on Testing and Assessment, No. 331, Environment, Health and Safety, Environment Directorate, OECD.

Olesti E, González-Ruiz V, Wilks MF, Boccard J and Rudaz S, 2021. Approaches in metabolomics for regulatory toxicology applications. Analyst, 146, 1820–1834. https://doi.org/10.1039/D0AN02212H

OpenFoodTox. Available online: https://www.efsa.europa.eu/en/data/chemical-hazards-data

Pal R, Mamidi MK, Das AK, Gupta PK and Bhonde R, 2012. A simple and economical route to generate functional hepatocyte-like cells from hESCs and their application in evaluating alcohol induced liver damage. Journal of Cellular Biochemistry, 113, 19–30. https://doi.org/10.1002/jcb.23391

Palmgrén JJ, Mönkkönen J, Korjamo T, Hassinen A and Auriola S, 2006. Drug adsorption to plastic containers and retention of drugs in cultured cells under *in vitro* conditions. European Journal of Pharmaceutics and Biopharmaceutics, 64, 369–378. https://doi.org/10.1016/j.ejpb.2006.06.005

Park BK, Boobis A, Clarke S, Goldring CE, Jones D, Kenna JG, Lambert C, Laverty HG, Naisbitt DJ, Nelson S, Nicoll-Griffith DA, Obach RS, Routledge P, Smith DA, Tweedie DJ, Vermeulen N, Williams DP, Wilson ID and Baillie TA, 2011. Managing the challenge of chemically reactive metabolites in drug development. Nature Reviews Drug Discovery, 10, 292–306. https://doi.org/10.1038/nrd3408

Partosch F, Mielke H, Stahlmann R, Kleuser B, Barlow S and Gundert-Remy U, 2014. Internal threshold of toxicological concern values: enabling route-to-route extrapolation. Archives of Toxicology, 89, 941–948.

Pawar G, Madden JC, Ebbrell D, Firman JW and Cronin MT, 2019. In silico toxicology data resources to support read-across and (Q) SAR. Frontiers in Pharmacology, 10, 561.

Pelkonen O, Pasanen M, Tolonen A, Koskinen M, Hakkola J, Abass K, Laine J, Hakkinen M, Juvonen R, Auriola S, Storvik M, Huuskonen P, Rousu T and Rahikkala M, 2015. Reactive metabolites in early drug development: predictive in vitro tools. Current Medicinal Chemistry, 22, 538–550. https://doi.org/10.2174/0929867321666141012175543

Pelkonen O, Tolonen A, Rousu T, Tursas L, Turpeinen M, Hokkanen J, Uusitalo J, Bouvier d'Yvoire M and Coecke S, 2009. Comparison of metabolic stability and metabolite identification of 55 ECVAM/ICCVAM validation compounds between human and rat liver homogenates and microsomes - a preliminary analysis. Altex, 26, 214–222. https://doi.org/10.14573/altex.2009.3.214

Pomponio G, Savary CC, Parmentier C, Bois F, Guillouzo A, Romanelli L, Richert L, Di Consiglio E and Testai E, 2015. *In vitro* kinetics of amiodarone and its major metabolite in two human liver cell models after acute and repeated treatments. Toxicology in Vitro, 30, 36–51. https://doi.org/10.1016/j.tiv.2014.12.012

Renwick AG, 1993. Data-derived safety factors for the evaluation of food additives and environmental contaminants. Food Additives and Contaminants, 10, 275–305. https://doi.org/10.1080/02652039309374152. PMID: 8359312.

RPA, Hsl, IEH, 2017. Human biomonitoring data collection from occupational exposure to pesticides – Final Report. EFSA supporting publication 2017;EN-1185, 207 pp.

Rusyn I and Roth A, 2021. Editorial overview of the special issue on application of tissue chips in toxicology. Toxicology, 450, 152687. https://doi.org/10.1016/j.tox.2021.152687

Ryu S, Tess D, Chang G, Keefer C, Burchett W, Steeno GS, Novak JJ, Patel R, Atkinson K, Riccardi K and Di L, 2020. Evaluation of fraction unbound across 7 tissues of 5 Species. Journal of Pharmaceutical Sciences, 109, 1178–1190.

SCCS (Scientific Committee on Consumer Safety), 2021. The SCCS Notes of Guidance for the Testing of Cosmetic Ingredients and their Safety Evaluation 11th revision, 30-31 March 2021, SCCS/1628/21.

Schultz TW, Amcoff P, Berggren E, Gautier F, Klaric M, Knight DJ, Mahony C, Schwarz M, White A and Cronin MTD, 2015. A strategy for structuring and reporting a read-across prediction of toxicity. Regulatory Toxicology and Pharmacology, 72, 586–601. https://doi.org/10.1016/j.yrtph.2015.05.016

Schultz TW, Richarz AN and Cronin MT, 2019. Assessing uncertainty in read-across: questions to evaluate toxicity predictions based on knowledge gained from case studies. Computational Toxicology, 9, 1–11. https://doi.org/10.1016/j.comtox.2018.10.003

Serras AS, Rodrigues JS, Cipriano M, Rodrigues AV, Oliveira NG and Miranda JP, 2021. A critical perspective on 3D liver models for drug metabolism and toxicology studies. Frontiers in Cell and Developmental Biology, 9, 626805. https://doi.org/10.3389/fcell.2021.626805

Shah I, Tate T and Patlewicz G, 2021. Generalized Read-Across prediction using genra-py. Bioinformatics, 37, 3380–3381.

Si-Tayeb K, Noto FK, Nagaoka M, Li J, Battle MA, Duris C, North PE, Dalton S and Duncan SA, 2010. Highly efficient generation of human hepatocyte-like cells from induced pluripotent stem cells. Hepatology, 51, 297–305. https://doi.org/10.1002/hep.23354

Smeraldi C, Giarola A, Aggett PJ, Moldeus P and Gundert-Remy U 2020. Use of mechanistic information to derive chemical-specific adjustment factors–Refinement of risk assessment. Regulatory Toxicology and Pharmacology, 117, 104776. https://doi.org/10.1016/j.yrtph.2020.104776.

Smith CM, Nolan CK, Edwards MA, Hatfield JB, Stewart TW, Ferguson SS, Lecluyse EL and Sahi J, 2012. A comprehensive evaluation of metabolic activity and intrinsic clearance in suspensions and monolayer cultures of cryopreserved primary human hepatocytes. Journal of Pharmaceutical Sciences, 101, 3989–4002. https://doi.org/10.1002/jps.23262

Spaggiari D, Mehl F, Desfontaine V, Grand-Guillaume PA, Fekete S, Rudaz S and Guillarme D, 2014a. Comparison of liquid chromatography and supercritical fluid chromatography coupled to compact single quadrupole mass spectrometer for targeted in vitro metabolism assay. Journal of Chromatography A, 1371, 244–256. https://doi.org/10.1016/j.chroma.2014.10.055

Spaggiari D, Geiser L and Rudaz S, 2014b. Coupling ultra-high-pressure liquid chromatography with mass spectrometry for in-vitro drug-metabolism studies. Trends in Analytical Chemistry, 63, 129–139. https://doi.org/10.1016/j.trac.2014.06.021

Stachulski AV, Baillie TA, Park BK, Obach RS, Dalvie DK, Williams DP, Srivastava A, Regan SL, Antoine DJ, Goldring CE, Chia AJ, Kitteringham NR, Randle LE, Callan H, Castrejon JL, Farrell J, Naisbitt DJ and Lennard MS, 2013. The generation, detection, and effects of reactive drug metabolites. Medicinal Research Reviews, 33, 985–1080. https://doi.org/10.1002/med.21273

Testai E, Bechaux C, Buratti FM, Darney K, Di Consiglio E, Kasteel EEJ, Kramer NI, Lautz LS, Santori N, Skaperda Z-V, Kouretas D, Turco L and Vichi S, 2021. Modelling human variability in toxicokinetic and toxicodynamic processes using Bayesian meta-analysis, physiologically-based modelling and in vitro systems. EFSA Supporting Publications 2021;18(4):6504, 44 pp. https://doi.org/10.2903/sp.efsa.2021.EN-6504

Tolonen A and Pelkonen O, 2015. Analytical challenges for conducting rapid metabolism characterization for QIVIVE. Toxicology, 332, 20–29. https://doi.org/10.1016/j.tox.2013.08.010

Tracking System for Alternative methods towards Regulatory acceptance (TSAR), Test Method number 2009-13: Cytochrome P450 (CYP) enzyme induction in vitro method using the cryopreserved human hepatocytes. Available online: https://tsar.jrc.ec.europa.eu/test-method/tm2009-13

Tracy TS, Chaudhry AS, Prasad B, Thummel KE, Schuetz EG, Zhong XB, Tien YC, Jeong H, Pan X, Shireman LM, Tay-Sontheimer J, Lin YS, 2016. Interindividual Variability in Cytochrome P450-Mediated Drug Metabolism. Drug Metabolism and Disposition, 44, 343–351. https://doi.org/10.1124/dmd.115.067900. Epub 2015 Dec 17. PMID: 26681736; PMCID: PMC4767386.

Turpeinen M, Tolonen A, Chesne C, Guillouzo A, Uusitalo J and Pelkonen O, 2009. Functional expression, inhibition and induction of CYP enzymes in HepaRG cells. Toxicology in Vitro, 23, 748–753. https://doi.org/10.1016/j.tiv.2009.03.008

UK-DEFRA (Department for Environment Food and Rural Affairs), 2003. Final Report: Differential Metabolism of Chiral Compounds – PS2501. Available online: https://sciencesearch.defra.gov.uk/Default.aspx?Menu=Menu&Module=More&Location=None&Completed=0&ProjectID=11787

US EPA (United States Environmental Protection Agency), 2021. List of Alternative Test Methods and Strategies (or New Approach Methodologies [NAMs]) Second Update: February 4th, 2021. Available online: https://www.epa.gov/sites/production/files/2021-02/documents/nams_list_second_update_2-4-21_final.pdf

Van de Kerkhof EG, de Graaf IA, de Jager MH, Meijer DK and Groothuis GM, 2005. Characterization of rat small intestinal and colon precision-cut slices as an *in vitro* system for drug metabolism and induction studies. Drug Metabolism and Disposition, 33, 1613–1620. https://doi.org/10.1124/dmd.105.005686

Van de Kerkhof EG, de Graaf IAM, de Jager MH and Groothuis GMM, 2007. Induction of phase I and II drug metabolism in rat small intestine and colon *in vitro*. Drug Metabolism and Disposition, 35, 898–907. https://doi.org/10.1124/dmd.106.014563

Wang DQ, Tazuma S, Cohen DE and Carey MC, 2003. Feeding natural hydrophilic bile acids inhibits intestinal cholesterol absorption: studies in the gallstone-susceptible mouse. American Journal of Physiology-Gastrointestinal and Liver Physiology, 285, G494–G502. https://doi.org/10.1152/ajpgi.00156.2003

Wang Q, Ye C, Jia R, Owen AJ, Hidalgo IJ and Li J, 2006. Inter-species comparison of 7-hydroxycoumarin glucuronidation and sulfation in liver S9 fractions. In Vitro Cellular and Developmental Biology. Animal, 42, 8–12. https://doi.org/10.1007/s11626-006-0004-z

Weidolf L and Wilson ID, 2016. Understanding drug metabolism in humans. In: SL Iverson and DA Smith (eds.). Metabolite Safety in Drug Development. https://doi.org/10.1002/9781118949689.ch6

Whalley PM, Bartels M, Bentley KS, Corvaro M, Funk D, Himmelstein MW, Neumann B, Strupp C, Zhang F and Mehta J, 2017. An in vitro approach for comparative interspecies metabolism of agrochemicals. Regulatory Toxicology Pharmacology, 88, 322–327. https://doi.org/10.1016/j.yrtph.2017.03.020

Wilk-Zasadna I, Bernasconi C, Pelkonen O and Coecke S, 2015. Biotransformation *in vitro*: An essential consideration in the quantitative *in vitro*-to-*in vivo* extrapolation (QIVIVE) of toxicity data. Toxicology, 332, 8–19. https://doi.org/10.1016/j.tox.2014.10.006

Williams AJ, Lambert JC, Thayer K and Dorne JLC, 2021. Sourcing data on chemical properties and hazard data from the US-EPA CompTox Chemicals Dashboard: a practical guide for human risk assessment. Environment International, 154, 106566.

World Health Organization & International Programme on Chemical Safety, 1994. Assessing human health risks of chemicals: derivation of guidance values for health-based exposure limits / published under the joint sponsorship of the United Nations Environment Programme, the International Labour Organisation, and the World Health Organization. World Health Organization. Available online: https://apps.who.int/iris/handle/10665/40675

Worth A, Asturiol D and Paini A, 2018a. EURL ECVAM Rodent In Vitro Intrinsic Clearance Database, European Commission, Joint Research Centre (JRC) [Dataset]. PID. Available online: https://data.europa.eu/89h/jrc-eurl-ecvam-rodent-in-vitro-intr-clear-db

Worth A, Asturiol D and Paini A, 2018b. EURL ECVAM Rodent In Vivo Biotransformation Database. European Commission, Joint Research Centre (JRC) [Dataset] PID. Available online: https://data.europa.eu/89h/jrc-eurl-ecvam-rodent-in-vivo-biotrans-db

Wu B, 2011. Substrate inhibition kinetics in drug metabolism reactions. Drug Metabolism Reviews, 43. 440–456. https://doi.org/10.3109/03602532.2011.615320

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 L 93, 3.4.2013, p. 1–84. Available online: https://data.europa.eu/eli/reg/2013/283/oj

Ye J, Zhao M, Niu L and Liu W, 2015. Enantioselective environmental toxicology of chiral pesticides. Chemical Research in Toxicology, 28, 325–338. https://doi.org/10.1021/tx500481n

Zanger UM and Schwab M, 2013. Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation. Pharmacology Theory, 138, 103–141. https://doi.org/10.1016/j.pharmthera.2012.12.007

Zhao Y, Zhu X, Jiang W, Liu H and Sun B, 2021. Chiral recognition for chromatography and membrane-based separations: recent developments and future prospects. Molecules, 26, 1145. https://doi.org/10.3390/molecules26041145

Abbreviations

ADME	See TK
AEP	Aggregated exposure pathway
AOP	Adverse Outcome Pathway
BMDL	Benchmark Dose Level
CL int; in vitro intrinsic clearance	CL int, *in vitro* time-dependent decrease of the test item (μL/min per mg protein or μL/min per 10^6 cells) in microsomes or hepatocytes respectively
CONTAM	EFSA Panel on Contaminants in the Food Chain
CSAF	Chemical-Specific Adjustment Factor
CYP	Cytochromes P450
Cytotoxicity assessment	General cytotoxicity (or basal cytotoxicity) is the result of toxic effects on structures and functions common to all cells of the body, such as DNA, chromosomes, mitochondria, the cytoskeleton and various membranes (OECD, 2018)
DEFRA	Department for Environment Food & Rural Affairs
DHM; Disproportionate human metabolite(s)	Metabolite(s) produced by both human- and animal hepatocytes, but at a significantly higher amount with the human hepatocytes than in animal species.
Distal metabolite	Produced from two or more reactions from the test item. This term has been used in this opinion. In xenobiotic metabolism secondary metabolite is used to define this kind of metabolite, however because this term is used in a different meaning in plant or microorganism sciences, a word ' distal' has been often used in this opinion

EURL ECVAM	EU Reference Laboratory for alternatives to animal testing
FDA	Food Drug Administration
FP7 Predict IV	Seventh Framework Programme. Profiling the toxicity of new drugs: a non-animal-based approach integrating toxicodynamics and biokinetics.
GSH	Glutathione
HBGV	Health-based guidance values
HPLC	High-pressure liquid chromatography
HPLC-UV	High-pressure liquid chromatography-ultraviolet
HRMS	High Resolution MS
IATA	integrated approach to testing and assessment
IMS	Ion mobility
INCHI	International Chemical Identifier
IPCS	International Programme on Chemical Safety
IUPAC	International Union of Pure and Applied Chemistry
KCN	Potassium Cyanide
Km	Michaelis constant
LC-MS	liquid chromatography-mass spectrometry
livestock	Cattle, horses, poultry, and similar animals kept for domestic use but not as pets, especially on a farm or ranch.
LDH	Lactate dehydrogenase
LOAEL	Low observed adverse effect level
LOE	Line of evidence
Metabolic competence	Capacity of a test system to metabolise a test item.
MAO	monoamino oxidase
Mass balance	The term 'mass balance' refers to attempting to collect as much of the radiolabel administered as possible. The radiolabel is collected while monitoring the exposure of radioactivity in biological samples and/or accessible tissues of interest.
MetaPath	Metabolic Pathways
Met-ID	Metabolite identification
MS	mass spectrometry
MS/MS	Tandem mass spectrometry
MSE/DDA	Data dependent Acquisition
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
NADPH	Nicotinamide adenine dinucleotide phosphate
NAMs	New Approach Methodologies
NB	Nota bene
NMR	nuclear magnetic resonance
NOAEL	No observed adverse effect level
Non-specific binding	Attachment of the test item to either the macromolecules of the cell medium or sorption of the test item to the plastic labware or physical sequestration of the test item in the matrix (e.g. collagen scaffolds)
OECD	Organisation for Economic Co-operation and Development
OHT	OECD Harmonised Templates
PB(T)K models	Physiologically Based (Toxico)Kinetic models. Physiologically based (toxico) kinetic, or alternatively referred to as physiologically based pharmacokinetic or biokinetic models, are quantitative descriptions of absorption, distribution, metabolism, and excretion (ADME, possibly including toxicity as ADMET) of synthetic or natural chemical substances in humans and other animal species. PBTK models are increasingly being used as an effective tool for designing toxicology experiments and for conducting extrapolations essential for risk assessments (e.g. in pharmaceutical research and drug development, and in health risk assessment for cosmetics or general chemicals) (OECD, 2021).
Phase 1 metabolism	Oxidation, reduction or hydrolysis reactions
Phase 2 metabolism	Conjugation reactions, i.e. glucuronidation, acetylation, glutathione conjugation and sulfation

Pivotal species	Laboratory animal species used in toxicological studies that have been used for setting human health based guidance values.
PPR	Plant Protection Products and their Residues
PREV	Pesticide Peer Review
Primary metabolite	Produced from a single biotransformation reaction from the test item (Dalvie et al., 2009)
QIVIVE	Quantitative in vitro–in vivo extrapolation
QSAR	Quantitative structure–activity relationship
RAM	Radioactive detector
REACH	Registration, Evaluation, Authorisation and Restriction of Chemicals
Slow metabolism	Less than 20% decrease of the concentration of the test item (and its related conversion into metabolites) during 4 h in suspended hepatocytes
Secondary metabolite	Produced from two or more reactions from the test item (Dalvie et al., 2009). Because this term is used in a different meaning in plant or microorganism sciences, a word ' distal' has been often used in this opinion.
TG	Test Guideline
TD	Toxicodynamics
Test item	Synonym of the active substance
3Rs principle	Replacement, reduction and refinement of animal experimentation
TK (or ADME)	Toxicokinetics. Study of the absorption, distribution, excretion, and metabolism (ADME) of substances (OECD TG 417)
TLC	Thin layer chromatography
TOF	Time-of-flight
TTC	Threshold of toxicological concern. The threshold of toxicological concern (TTC) approach is a pragmatic, scientifically valid methodology to assess the safety of substances of unknown toxicity found in food
UF	Uncertainty Factor
UHM, Unique human metabolite(s)	Metabolite(s) only formed/detected *in vitro* in human hepatocytes and not in laboratory animal species.
US EPA	United States Environmental Protection Agency
Vmax	Maximum velocity
WG	Working group
WHO	World Health Organization

Appendix A – *In vitro* systems for studying hepatic metabolism

A.1. Metabolism of xenobiotics

A large number of enzymes participate in the metabolism of compounds foreign to the body ('xenobiotics' including pesticide active substances) and traditionally these biotransformations are grouped into two phases. Phase I (i.e. oxidations, reduction, hydrolysis, etc.) is responsible for either detoxification or for bioactivation resulting in the formation of more toxic metabolites. Phase II metabolism (conjugations a parent or a phase I metabolite with glucuronic acid, sulfate, glutathione, etc.) mainly facilitates the excretion of chemicals by making them more hydrophilic and generally makes reactive metabolites inactive by, e.g. glutathione conjugation. However, it is known that some phase II metabolites can be toxic (Glatt, 2000), indicating that it is relevant to also cover all Phase II metabolites, including sulfo-conjugation for which the responsible enzymes are lacking in liver microsomes, as well as glutathione conjugation which takes predominantly place in the soluble fraction in the cell.

The number of metabolites produced from a single substance may be from none or a few to more than hundreds. Often, one or a couple major and several minor metabolites are formed from a single compound. However, the PPR Panel noted that primary metabolites are usually further metabolised to more distal metabolites and thus amounts of any single metabolite could change as a function of time, either decreasing, if further metabolism is very active, or increasing as happens to many more distal metabolites, e.g. glucuronides and other conjugates. It should also be remembered that *in vivo* metabolites move through the body to be ultimately excreted via urine or bile (faeces). Thus, actually the measurement of metabolites in any *in vitro* test system is a moving target, a snapshot of the continuous dynamic process, and by *in vitro* systems, it is possible to obtain only a partial, incomplete and time point specific view of the metabolite profile of a substance.

A.2. *In vitro* test systems

A large variety of *in vitro* test systems to study metabolism are currently available ranging from tests to study single or a few metabolising enzymes (e.g. recombinant enzymes) to organs-on-a-chip techniques intended to mimic metabolically active tissues/organs in more physiological arrangements (e.g. organoids). These *in vitro* hepatic test systems are used for the production, identification and quantitation of metabolites, but they all have definitive advantages and disadvantages in this respect. Subcellular fractions (e.g. liver microsomes) and isolated hepatocytes have been used for decades for metabolism studies and there is a huge database on using them for specific purposes in metabolism studies. There are a large number of more elaborated test systems (see Table A.1), many of them commercially available, but these systems are still under development and there is insufficient experience and validation.

Because the goal of this opinion is to compare metabolite profiles between humans and test species with the aim to detect both UHM and DHM, to be applied for human risk assessment based on animal regulatory toxicity studies, the outcome of the test should mimic the *in vivo* situation as much as possible. In the context of this opinion, it is appropriate to state that there exists a common consensus among scientists and regulators that hepatocyte *in vitro* test system is the best option when there is a need to produce a metabolite profile resembling an *in vivo* profile. The main reasons to select hepatocytes have been elaborated in Section 3.

A.3. Short-term incubation: hepatocytes in suspension

As explained in Section 3, the opinion recommends using in the first place a pool of hepatocytes in suspension, unless there are reasons as elaborated below to employ cultured hepatocytes, and to perform incubations at three time points within a 4-h time frame. If conversion of the parent is less than 20% (and if this is not increased to 20% or more upon increasing the hepatocyte concentration (e.g. to 2 million hepatocytes/mL)), or existing information suggests a low clearance for that chemical, the PPR Panel recommends to perform a longer incubation by using plated hepatocytes exposed to the chemical for 24 h.

A.4. Long-term incubation: hepatocytes plated

For chemicals that are slowly metabolised, the PPR Panel recommends to use plated primary hepatocytes, since suspended hepatocytes have a limited viability at longer incubation times (more than 4 h). In the context of this opinion, a chemical (parent compound) is considered slowly metabolised when its concentration decreases less than 20% after 4 h incubation using the recommended cell concentration of 1 million hepatocytes/mL. Only when the slowly metabolised test item is known to have a half-life lower than 24 h, the long-term incubation can be skipped. Various long-term hepatic test systems are available, as being described in Bernasconi et al. (2019). It is considered sufficient to use plated hepatocytes in a 2D configuration. It is acknowledged that various culture techniques using 3D configurations are rapidly evolving and currently available from commercial developers (see Table A.1). Advantages of 3D configurations include a better differentiation state and longer longevity. On the other hand, one should consider that chemical diffusion into the liver cells, and related availability of chemical to the metabolising enzymes, may be lower in 3D configurations. The use of 3D configurations is in this stage not recommended because the aim of this opinion is to have simple study designs, and since various adequate 2D models are available. Furthermore, the chemical applicability domain of hepatic 3D models has not been evaluated so far, so their appropriateness for comparative metabolism studies with pesticide active substances is not known.

Table A.1: *In vitro* hepatic test systems used for the production, identification and quantitation of metabolites: advantages and disadvantages

Test system	Main advantage	Main disadvantage	References
***In vitro* systems with extensive experience and current use for regulatory purposes** (Modified from Wilk-Zasadna et al. (2015) and Gouliarmou et al. (2018))			
Recombinant enzymes	Study of iso-enzyme specific metabolism	One or few enzymes in isolation in an 'artificial' environment	Foti et al. (2010), Buratti et al. (2011)
Liver microsomes	Major phase I enzymes present (CYPs, etc.)	Insufficient activities for most phase II enzymes	Buratti et al. (2013), Buratti and Testai (2015)
Liver S9 or homogenate	Major phase I and II enzymes present	Subcellular preparation	Pelkonen et al. (2009)
* **Primary hepatocytes, in suspension**	Major phase I and II enzymes present and intact cell system: the current 'gold standard' for incubation periods up to 4 h	Viable for limited time	Smith et al. (2012)
* **Primary hepatocytes, plated**	Major phase I and II enzymes present and intact cell system: the current 'gold standard' for incubation periods longer than 4 h	Gradual decline of metabolic functions in time	Smith et al. (2012)
HepaRG® cell line	Metabolic competence 'hepatocyte-like'	Represents one donor only	Bell et al. (2017)
Human liver slices	Major phase I and II enzymes present and intact cell system, basic hepatic architecture preserved	Specialised skills needed and Difficult to obtain	de Graaf et al. (2010)
Test systems based on isolated hepatocytes and/or supporting cells (examples) Currently investigated, but not yet established, test systems for assessing (human) liver metabolism (see e.g. Godoy et al., 2013; Wilk-Zasadna et al., 2015; Lauschke et al., 2016; Serras et al., 2021)			
Primary cell co-cultures	Relevant interactions with other cell types	Limited information on performance available	Guguen-Guillouzo and Guillouzo (2010)
Micropatterned plated cell cultures (e.g. Hepatopac)	Improved stability and functionality Commercially available	Limited information on performance available	Chan et al. (2013)
HμREL® Biochip (microfluidic flow)	Improved stability and functionality Commercially available	Limited information on performance available Relative complex set-up	Burton et al. (2018)

Test system	Main advantage	Main disadvantage	References
Spheroid scaffold-free cultures	Improved stability and functionality Relatively simple setup	Limited information on performance available	Bell et al. (2017)
Organoids based on primary hepatic material	Close resemblance to *in vivo* physiological situation	Limited information on performance available	Gough et al. (2021)
Hepatocyte-like cells based on pluripotent (embryonic or induced) stem cells	Easy access to material from different donors (for induced pluripotent stem cells)	Limited information on performance available Functionality in doubt	Si-Tayeb et al. (2010), Pal et al. (2012)

*: Recommended test system for testing comparative metabolism.

Appendix B – Test item and reference compounds

B.1. Purity of the test item

The PPR Panel recommends to specify chemical purity for the test item. For reference compounds, the PPR Panel recommends purity as high as possible and certified material.

The test item to be used in the comparative *in vitro* metabolism study can be both non-radiolabelled or radiolabelled material. The choice mainly depends on the sensitivity of the analytical method employed for metabolite identification analysis. The advances in analytical methods allow the use of non-radiolabelled material still ensuring good sensitivity, although special attention should be paid, when using non-radiolabelled material and developing the analytical method, to avoid the possibility that potentially relevant metabolites are not detected because present at levels below the limit of detection of the analytical method.

Radiolabelled material can balance the selectivity issue of the analytical method. The use of radiolabelled material could be from a practical use more straightforward since it is normally used when conducting animal (rat) TK studies. When radiolabelled material is used, in line with the OECD TG 417 (item 18 and 19 OECD, 2010), the radiopurity of the radioactive test substance should be the highest attainable for a particular test substance (ideally it should be greater than 95%). Information should be provided on the type of radionuclide used.

B.2. Isotope labelling of the test item

The PPR recommends to follow the same principles as in OECD TG 501 (OECD, 2007a) where the test item should be labelled so that the degradation pathway can be traced as far as possible. Particularly, the radiolabel should be positioned in the molecule so that all significant moieties or degradation products can be tracked. If multiple ring structures or significant side chains are present, separate studies reflecting labelling of each ring or side chain will normally be required if it is anticipated that cleavage between these moieties may occur. In general, the position of the radiolabel when conducting a comparative *in vitro* metabolism study should reflect what is done when performing *in vivo* ADME studies according to OECD TG 417 (item 18 and 19 in the OECD TG 417; OECD, 2010).

B.3. Assess solubility, stability, volatility

The assessment of solubility, volatility and stability is highly relevant to determine the highest concentration of the test item that can be present in the medium containing the test system and to properly characterise the test item.

In line with OECD GD on GIVIMP (2018a), the PPR recommends to determine solubility, stability, volatility and non-specific binding (see Appendix C.1.3 for more detailed discussion on non-specific binding). These parameters can affect the concentration and the bioavailability of the test item and reference compound added to hepatocytes when performing a comparative *in vitro* metabolism study.

The concentration of the test item added to the test system (by diluting a stock solution in solvent to the medium) is called nominal concentration. Once the stock solution of the test item is added to the cell medium containing the test system (hepatocytes), the concentration of the test item available to the hepatocytes may be lower than the nominal concentration. This concentration available for the cells is called actual concentration. An actual concentration lower than the nominal concentration can be due to one or more processes related to solubility, stability, volatility and non-specific binding, which depend on the specific chemical–physical properties of the test item. Low actual concentration of the test item might cause that some relevant metabolites are formed at low levels which can eventually be not detected because below the limit of detection of the analytical method. Therefore, considering that the experimental layout is designed to maximise the chances of detecting all relevant metabolites by reducing any source which can lead to a low actual concentration of the test item.

Poor solubility and/or high volatility of the test item contributes to a low actual concentration. Moreover, testing a test item with poor solubility or that is highly volatile can be technically difficult.

Most probably, solubility of the test item has been assessed earlier in the testing process (*in vivo* TK study and toxicology studies), before performing the comparative *in vitro* metabolism study. However, the PPR recommends to always evaluate solubility before conducting a comparative *in vitro* metabolism study, as also indicated in GIVIMP. This is necessary because the solvent used to prepare the stock solution and then the cell medium for the working solution are most likely different from those used in the *in vivo* studies and solubility may be different.

It is suggested to prepare, if solubility allows, stock solutions that are 1,000 times higher than the desired nominal concentrations. OECD GIVIMP provides technical guidance on how to measure solubility, both in stock solutions and in working solutions. The easiest method relies on visual inspection although the accuracy is less than when using appropriate instrumentation as nephelometry analysis. This method is outlined in a recent validation study with human hepatocytes by EURL-ECVAM (Bernasconi et al., 2019).

Solubility has to be always measured in both the stock and working solutions. In case of apparent insolubility in the chosen solvent or in cell medium, dissolution can be attempted by trying other potential solvents or by incremental twofold dilution as necessary.

Stability refers to the degradation of the test item either under storage conditions and/or in test conditions which is not due to metabolism (e.g. chemical hydrolysis in aqueous solutions). The stability of the stock solution can be assessed to cover the storage duration period (days, weeks, months depending on the conditions of storage). Stability should also be assessed in the cell medium of the working solution at 37°C in medium by measuring the test item pre- and post-incubation (i.e. 4 h for short-term incubation and 24 h for long-term incubation) including centrifugation for observation of any precipitation (pellet residue).

B.4. Preparation of stock and working solutions

The PPR recommends to follow the principles described in GIVIMP (OECD, 2018a) to prepare stock and working solutions

When preparing the stock solution, different solvents can be used, depending on physical–chemical properties of the test item, to solubilise the test item. Solubility, stability and volatility can be determined for the stock solution as well as the working solution. The PPR recommends to store the stock solution in glass, to avoid unspecific binding, especially for lipophilic compounds. It must be noted, however, that certain chemicals adhere to glass surfaces, indicating that for such chemicals storage in glass may not be recommends.

The working solution is obtained by diluting in one or more steps the stock solution. The working solution must be prepared in the same cell medium used when conducting the comparative *in vitro* metabolism study with hepatocytes. Attention should be paid to the amount of organic solvent which is present in the working solution. The dilution scheme from the stock to the working solution can be designed to ensure that a minimum of organic solvent is present in the working solution (i.e. < 0.5% DMSO). This is to ensure that the organic solvent does not interfere with the test system and thus with the performance of the *in vitro* experiment.

B.5. Positive control

The purpose of the positive control chemical is to check the proper performance of the test system (GIVIMP). Because a response in the *in vitro* comparative metabolism assay is the temporal and quantitative metabolite profile produced by biotransformation enzymes in hepatocytes – which is not known beforehand with a test substance – preferable positive controls would be compounds with several metabolites produced by several phase I and II enzymes, which would provide a required proof that hepatocytes under study contain a complete and functioning biotransformation machinery. This last requirement is rather apparent with human hepatocytes because of their variable contents of biotransformation enzymes. For this reason, pools based on several, preferably at least 20 different donors are regularly used to level off considerable interindividual variability (Tracy et al., 2016). Pooled human hepatocytes consisting of 10–100 individuals could be purchased from several companies.

As hepatocytes from maximally five species (human, rat, mouse, dog, rabbit) will be compared, a question about suitable positive controls becomes important. In this case, simple compounds with qualitatively and quantitatively well-defined metabolism in various species, i.e. testosterone 6β-hydroxylation or 7-ethoxycoumarin O-deethylation, and subsequent glucuronidation and sulfation, would be feasible to use (Arlotto et al., 1991; Honkakoski and Negishi, 1997; Wang et al., 2003, 2006; van de Kerkhof et al., 2005, 2007). Alongside with historical values, they provide information on the viability and metabolic competence of the hepatocyte pool used. However, given large and often poorly understood qualitative and quantitative interspecies differences in xenobiotic-metabolising enzymes, it is not possible to derive any enzyme-specific information from their use as positive controls.

If specific enzymes are deemed to play a major role in the metabolism (e.g. as estimated based on rat *in vivo* TK study or other appropriate evidence), more specific positive controls may also be

considered to be included. Such chemicals are, e.g. CYP-selective substrates and metabolites of melatonin (CYP1A1/2), coumarin (CYP2A6), bupropion (CYP2B6), amodiaquine (CYP2C8), tolbutamide (CYP2C9), omeprazole (CYP2C19), dextromethorphan (CYP2D6), chlorzoxazone (CYP2E1), midazolam (CYP3A4) and testosterone (CYP3A4), all of which have been compared in human hepatocytes at least up to 48 h of culturing (Turpeinen et al., 2009). Potential multienzyme multimetabolite substrates have not been characterised as potential reference chemicals. Further guidance and examples can be found in the OECD guidance document (OECD, 2018).

B.6. Solvent control

For hepatocytes, the experiment should include a solvent (vehicle) control. Solvent control samples provide insight into possible background signals in the analytical measurements, pointing to the possible presence of signals that are not related to biotransformation of the test item.

B.7. Acceptance criteria

The documentation provided by the test system supplier should also be taken into account together with historical data, when available, and used to establish *in vitro* method acceptance criteria (OECD and GIVIMP, OECD, 2018a).

Appendix C – Experimental test conditions

C.1. Preliminary experiments

To ascertain that the comparative metabolism study is performed in the best way possible, preliminary experiments need to be performed. These include assessment of the cytotoxicity of the test item, determination of the metabolic competence of the test system (if no information on the specific batch of the test system is available) and determination of non-specific binding.

C.1.1. Cytotoxicity assessment

Various methods exist that can be used to assess chemical-induced cytotoxicity, including methods that determine the number of dead cells (such as the trypan blue exclusion assay), methods that determine the activity of the cells (such as the WST-1 or resazurin reduction assay) or methods that measure the release of cellular constituents, indicating possible cell death (such as the Lactate dehydrogenase (LDH) leakage assay). The latest schemes of cytotoxicity assessments in cells in culture have been described by Bernasconi et al. (2019), and some in-depth considerations elaborated by OECD GIVIMP (OECD, 2018a). Potential cytotoxicity of the test item for hepatocytes is determined starting from the highest soluble concentration, followed by a pre-experimented dilution scheme. The incubation time should follow the conditions used for the metabolism studies (i.e. 4 h for short-term incubation and 24 h for long-term incubation).

The PPR recommends to test for cytotoxicity a highest concentration of 100 μmol/L in cell medium, provided that this concentration is soluble. In case 100 μmol/L is not soluble, further dilutions by using a fixed dilution factor of 2 can be carried out to obtain the highest soluble concentration that will be tested for cytotoxicity. A cytotoxicity study will be carried out using 100 μmol/L (or the highest soluble concentration) as the highest concentration, including also lower concentrations using a fixed dilution factor of 2 (final concentrations amounting to 100 μmol/L, 50 μmol/L, 25 μmol/L, 12.5 μmol/L, 6.25 μmol/L, 3.125 μmol/L, etc.). Based on the results of the cytotoxicity study, the highest concentration to be used for the comparative metabolism study will be selected, being the highest non-cytotoxic concentration (causing no more than 20% cell death). The cytotoxicity assessment can be performed by using the hepatocytes from all species used in the *in vitro* comparative metabolism study (i.e. human, rat, mouse, dog, rabbit).

Since the purpose of this study is to identify DHM and UHM, it is important to minimise any source of experimental variability that could confound the comparison of metabolic profiles among species. For this reason, the PPR recommends that the highest concentration (as well as the lower concentrations) should be the same across the different species.

C.1.2. Metabolic competence

The PPR recommends to characterise the metabolic competence of the test system by including reference chemicals/positive controls in hepatocytes of all the tested species (see Appendix B.5). This would allow to show that the test system adequately works when assessing the metabolism of the test item. With commercially available hepatocytes, suppliers typically provide for the batches they sell information on rates of metabolite formation related to various substrates, providing insights into the metabolic capacity of the batch used. If this information is not available, it is recommended to assess the metabolic competence of the test system by assessing time-dependent metabolite formation for a selection of reference items (see Appendix B.5). It is also recommended to perform such a study when a specific batch is used for the first time in the lab, even if information on the metabolic competence has been provided by the supplier.

The PPR considers sufficient to use a reference chemical that provides insight into the activities of those enzymes that are known to often lose their activities in *in vitro* incubations (e.g. CYPs) (Appendix B.5).

C.1.3. Non-specific binding

Once the highest concentration of the test item to be tested has been identified, based on cytotoxicity and solubility investigations, the PPR recommends to assess or estimate (e.g. *in silico*) non-specific binding in the preliminary experiment phase at any incubation time, since it is one of the processes affecting the freely dissolved concentration in the medium, limiting the amount of the

bioavailable test item, especially for a number of lipophilic chemicals, and lead to an overestimation of the intracellular concentration (Palmgrén et al., 2006; Kramer et al., 2015).

Non-specific binding refers to the binding of the test item to either the macromolecules of the cell medium (e.g. proteins in case cell medium-containing serum is used), to sorption to the plastic labware and to physical sequestration in the matrix (e.g. collagen scaffolds) (this last condition being limited to cases when 24-h incubation is necessary and plated hepatocytes in sandwich configuration are used). Results in the literature indicated that chemicals can be adsorbed to plastics depending on the lipophilicity in a dose- and time-dependent manner, possibly leading to binding saturation at high concentrations of the chemical. It has also been shown that serum constituents and plastics compete for binding the chemical (Pomponio et al., 2015). The relationship between lipophilicity and sequestration by cell matrices is not as clear as for the binding (Kramer et al., 2015).

When not assessed/estimated before starting the main experiment, whenever the recovery obtained summing up levels of the parent and the metabolite in the cell fraction and cell medium is less than 90% $\pm$ 5%, considered as the minimum recovery, the assessment of non-specific binding has to be performed.

While binding to macromolecules can be reduced by using low percentage or serum-free cell medium, in case the test item is highly adsorbed to plastic, glass labware could be used to overcome this issue or other considerations can be evaluated, such as plastic coating, to reduce the non-specific binding.

Methods to be used are described in 'Appendix H. Biokinetics and xenobiotic bioavailability' of the OECD GIVIMP Document, and indicated in a number of publications whose results are reviewed and cited in Kramer et al. (2015), reporting results of the FP7 Predict-IV project.

Non-specific binding can be assessed by using the same experimental conditions, which is used to perform the *in vitro* comparative metabolism study. Furthermore, the PPR recommends to assess non-specific binding for the different species if different cell media and labware is used for different species. Overall, the recovery should be similar across all species used in the main experiment. The reason, similarly to the cytotoxicity assessment, is to minimise any source of variability (in this case due to difference in chemical bioavailability) which could confound the comparison of metabolic profile among species, especially to identify and evaluate DHM. This means that in case a recovery around 80% is obtained, this could be acceptable providing the value is similar across all species. On the other hand, submitting data with low recovery are strongly discouraged, since it implies a low actual concentration of the test item which can therefore impact on metabolite identification (because metabolite(s) formed may not be detected due to low levels which are below the limit of detection of the analytical method).

It is recommended to use the same solvent, cell media, plastic device across species when conducting the comparative *in vitro* metabolism studies to minimise interference of non-specific binding, to enable a reliable estimation of tissue binding and intracellular free active item concentrations (Mateus et al., 2017; Barr et al., 2019; Ryu et al., 2020; Hsu et al., 2021).

C.2. Pilot *in vitro* clearance study

Having information on *in vitro* intrinsic hepatic clearance (CL_{int}, *in vitro*) of the chemical may help to better design the experiment on metabolite formation, e.g. to decide whether a 4-h incubation with hepatocytes in suspension will suffice to reach the minimal required conversion of 20% of the parent chemical, or whether a 24-h incubation with plated hepatocytes would be required. To that end, one may perform a pilot clearance study first before the main study on metabolite formation. Figure C.1 shows the simulated time-dependent levels of parent chemical remaining in *in vitro* incubations applying 1 million hepatocytes/mL for chemicals with different *in vitro* intrinsic hepatic clearance values (CL_{int}, *in vitro*), assuming that all chemical is available for clearance (so no non-specific binding). It shows that it can be expected that for chemicals with CL_{int}, *in vitro* values $\leq$ 1 µL/min per million hepatocytes, the percentage of remaining parent is $\geq$ 80%, indicating that the criterion of $\geq$ 20% conversion of parent is not met and a long (24-h) incubation would be needed. An increase in hepatocyte concentration may increase the percentage of parent conversion, although one must take into account that kinetics may not be linear between 1 and 2 million hepatocytes/mL and conversion may not necessarily increase twofold (especially for chemical with high non-specific binding to hepatocytes).

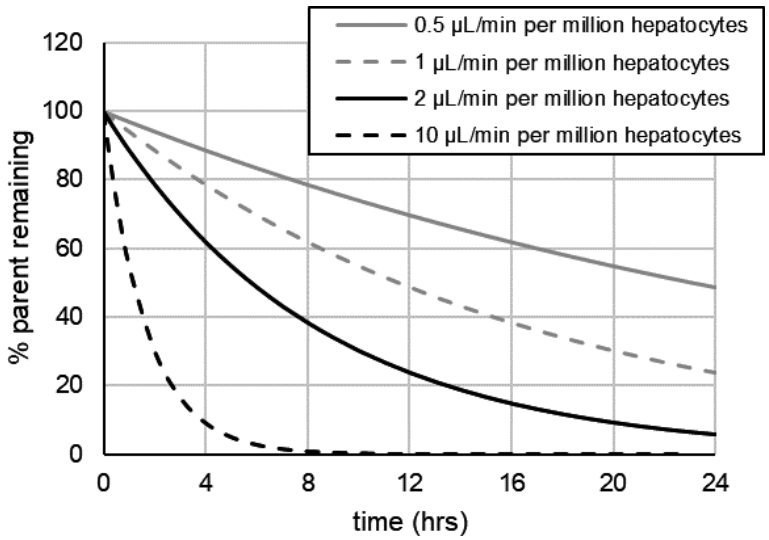

Figure C.1: Time-dependent simulated concentrations of parent chemical remaining in an in vitro incubation using 1 million (functional) hepatocytes/mL of chemicals with intrinsic in vitro hepatic clearance (CL int) values of 0.5, 1, 2 or 10 µL/min per million hepatocytes. NB: in this example, substrates are considered to be present at concentrations at which metabolic conversion is not saturated (i.e. are below Km). Also, this graph shows the estimated parent chemical levels assuming that all chemical is available for clearance. As the chemical may bind to plastic and hepatocytes, the available concentration for clearance will be lower and remaining parent level will be higher (if plastic and hepatocyte-bound chemical is recovered during sample preparation for analytical analysis)

C.3. Main Comparative *In vitro* metabolism study: experimental layout

When the test item has been properly characterised (by assessing solubility, volatility and stability) and the preliminary experiments conducted to identify the highest concentration to be tested (based on cytotoxicity and solubility), to verify the metabolic competence of the test system and to assess possible non-specific binding, the main experiment can be performed.

In the following sections, the experimental conditions on how to perform a comparative *in vitro* metabolism study are described to ensure that reliable data are obtained and to support their transparent interpretation.

It is important to note that in the following sections, minimum requirements on how to perform these types of studies are provided, proposing fixed values for the number of cells to be used (cell density) and the incubation time. Both these experimental parameters (cell density and incubation time) play an important role in the formation of metabolites. Therefore, depending on the specific chemical-specific properties of the test item to be tested, both parameters can be optimised. Reliable kinetic data are obtained in conditions in which kinetics are linear with respect to cell concentration and incubation time, which can be chemical-specific.

If studies are not performed to optimise conditions, standard conditions are recommended (i.e. 1.0 million cells/mL and 1, 2 and 4 h for chemicals with rapid conversion, and 24 h for chemicals with slow conversion).

Figure C.2 provides a general overview of the main comparative *in vitro* metabolism study, also summarising all the steps to be carried out during preliminary experiments and relevant information related to the test item.

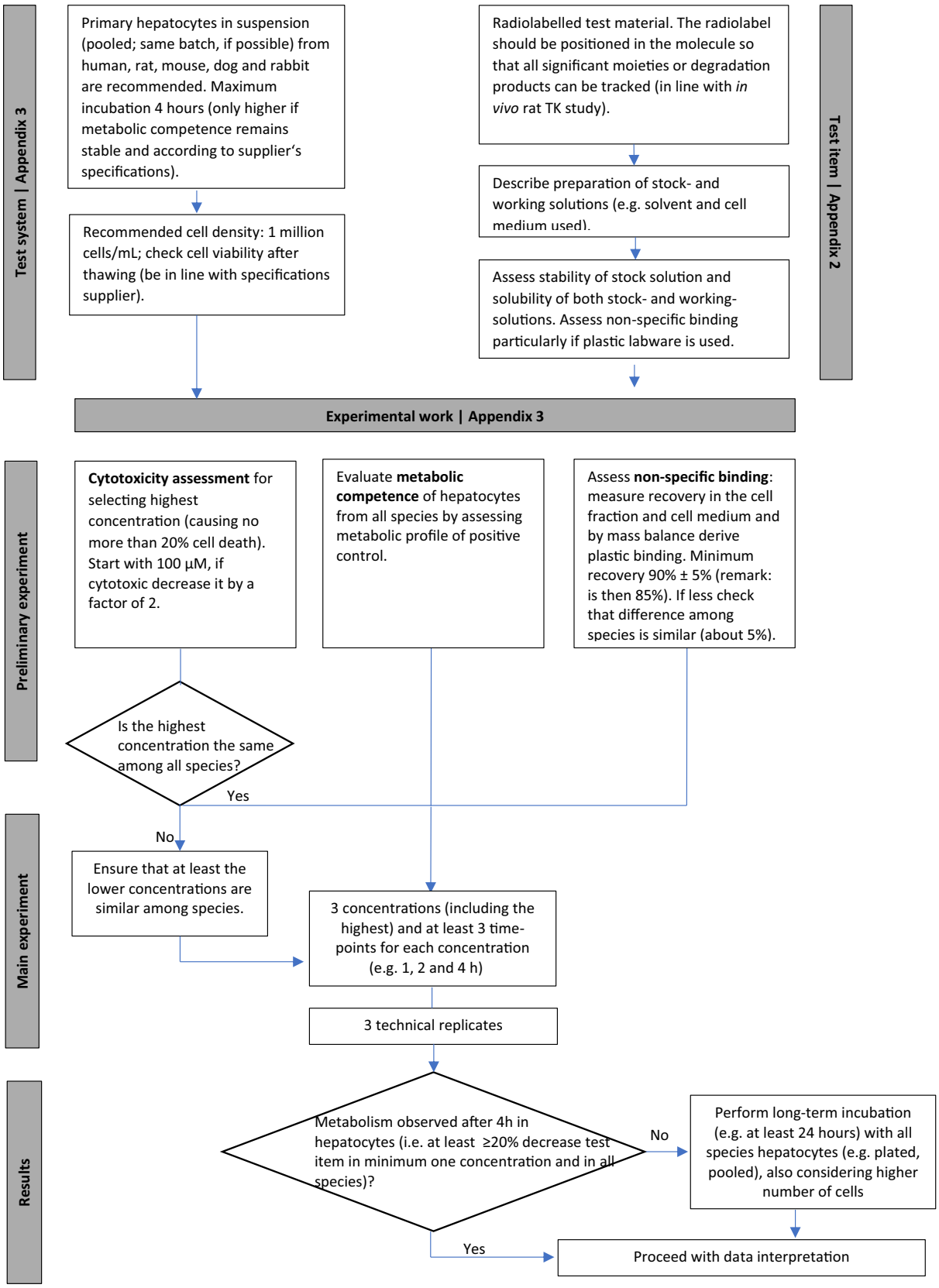

Figure C.2: Overview of the study design and experimental strategy as described in Appendixes B and C

C.3.1. Determination of cell density to be applied

It is important to use an appropriate cell density in the *in vitro* incubations. The optimal cell density may be chemical-specific and is preferably optimised for the chemical of interest. Cell density can be optimised in a first study, in which the test item is incubated using different concentrations of hepatocytes (e.g. 0.25, 0.5, 1 and 2 million cells/mL). Based on the results of such a first study, an optimal hepatocyte concentration can be selected for the final study. Ideally, a condition is selected within the linear frame regarding decrease of parent compound and increase of metabolite(s) related to cell number.

As the performance of optimisation studies for all species may not be feasible, a default cell concentration of 1.0 million hepatocytes/mL is recommended to be used for all species, allowing sufficient (more than 20%) conversion of the parent compounds for chemicals with clearance values > 1 μL/min per million hepatocytes (see Appendix C.2 for related estimations). As indicated before, in order to ensure that the test item is metabolised, a minimum of 20% decrease in parent compound concentration has to be observed at the end of incubation time (i.e. 4 h for short-term exposure or 24 h for long-term exposure) for one of the concentrations tested. If metabolic conversion is too low, the cell density may be increased in order to meet this acceptance criterion. One must realise that upon increase of cell number, non-specific binding of the chemical to the cells may increase, thereby decreasing the fraction of the parent chemical available for metabolism. Therefore, in case the cell density is changed, the PPR Panel recommends to re-evaluate non-specific binding.

C.3.2. Cell viability assessment

The PPR recommends to assess the quality of the hepatocytes used in the incubations by determining their viability. The percentage of viable cells after thawing should be around the suppliers' recommendations. This can be done before and after the experiment. At the end of the study, cell viability can be checked to ensure that the number of cells is not decreased and could be affected by cytotoxicity. In the absence of the test guideline, the PPR recommends that cell viability is not decreased more than 20% compared to the start of the study, since a larger decrease may indicate a rapid decrease in quality of the cells during the incubation.

The PPR recommends to use a test that quickly determines the number of viable cells as a fraction of the total cell numbers, such as a trypan blue exclusion assay or similar. Other methods, such as MTT/WST-based or resazurin-based assays, are less suitable, since these require a relative long incubation time and provide a relative and not an absolute readout for viability.

C.3.3. Concentrations to be tested

In the final study, it is recommended to test three concentrations (including the highest one) that are the same for the different species. The PPR recommends to cover a reasonable concentration-spacing range and considering also the analytical sensitivity.

The testing of more concentrations is based on the following rationales:

1) enhance the possibility to detect all metabolites formed by minimising the risk of the analytical sensitivity issues.

The study design is optimised to maximise the chances of identifying all relevant metabolites, particularly those that are formed at low levels and therefore that might not be detectable below the limit of detection of the analytical method. This study design is deemed necessary considering the challenges of the analytical method.

2) improve the evaluation of disproportionate metabolites.

Since the purpose of this study is also to identify DHM, it is important to minimise any source of experimental variability which could confound the comparison of metabolic profiles among species. For this reason, testing more than one concentration would allow to compare the concentration-dependent metabolic profiles of the different species, providing insights into possible saturation of metabolic pathways, which may differ from one species to the other.

C.3.4. Time points to be tested

For short-term exposure, the PPR recommends to expose hepatocytes for a maximum of 4 h, providing that based on the preliminary experiment, there is no cytotoxicity and that the cells are still viable and metabolically competent.

For long-term exposure, the PPR recommends to expose hepatocytes for 24 h, providing that based on the preliminary experiment, there is no cytotoxicity and that the cells are still viable and metabolically competent.

It is highly recommended to include more sampling time points in order to assess whether changes in metabolite levels are time dependent (indicating adequate functioning of the *in vitro* test system during the whole test duration). Furthermore, information on time-dependent changes in metabolite levels provides more insight into the kinetics of formation of primary and more distal metabolites, i.e. facilitating to have a complete picture on metabolic pathways and better evaluation of DHM in humans compared to animals.

C.3.5. Pooling of hepatocytes

The PPR recommends to use pooled hepatocytes for the comparative metabolism studies. For humans, both genders (with half females and half males) are recommended to use. For the animal species used for toxicity testing (i.e. rat, rabbit, mouse and dog), use of pooled hepatocytes of the sexes used in the toxicity studies is recommended, unless there are reliable evidence to use gender-mixed hepatocytes (i.e. *in vivo* studies). For rat, it can be considered, depending on the outcome of *in vivo* rat TK study (if available), to separately perform studies with female and male hepatocytes, since certain biotransformation enzymes are differently expressed between sexes. For rabbit, it is preferable to use only female hepatocytes (since only rabbit female is used in the developmental study).

Regarding the number of donors to be present in the hepatocyte pools, it is recommended to use a number of donors as high as possible. Regarding human hepatocytes (in suspension) pools of up to hundreds of donors are commercially available. As a minimum, 20 donors are recommended, to cover intraspecies differences in metabolism. For laboratory animals, less intraspecies differences are expected, and a pool of at least three donors is recommended to be used.

C.3.6. Technical replicates

The PPR recommends to run for each condition (time point and concentration), three technical replicates by using the same batch of pooled hepatocytes.

C.3.7. Long-term exposure: 24 h

Although isolated primary hepatocytes in solution perform well up to 4-h incubation time in terms of xenobiotic metabolism capacity, longer incubations lead to a rapid deterioration of enzyme activities. To preserve the hepatocellular phenotype of the isolated cells over this limit, cryopreservation is currently a routine way to get the desired cells in required amounts for either immediate short-term experiments ('in solution') or metabolism-competent hepatocyte cultures for extended incubations ('plated'). Plated hepatocytes are especially useful for the study of compounds with low metabolism (e.g. Ma et al., 2017). For long-term culturing, various strategies have been used to maintain the differentiation status of the isolated hepatocytes, but the main focus is still on the strategy that employs differentiation-promoting compounds in the culture medium. An overview on various antidifferentiation procedures and cryopreservation agents has been published (Gouliarmou et al., 2015). Metabolically competent cryopreserved human primary hepatocyte culture used for the validation of CYP induction has been recently published (Bernasconi et al., 2019) and protocols are available (TSAR, TM2009-13), where also most important primary literature sources have been referenced. Both short-term and long-term primary hepatocyte cultures are widely used also in pharmaceutical companies and contract research organisations.

C.3.8. Acceptance criteria

The acceptance criteria for experimental setting are described in detail into the OECD GIVIMP 2018a.

Appendix D – Chemical Analysis

D.1. Comparative metabolite profiles: usefulness of radioactive techniques

- When using radioisotopes (mostly with the ^{14}C isotope), if possible, the radiolabel should be located in a core portion of the molecule which is metabolically stable. The radiopurity of the radioactive test substance should be the highest attainable for a particular test substance (ideally it should be greater than 95%) in order to establish with the highest accuracy, the concentration used within the *in vitro* experiments.
- All significant moieties or degradation products can be tracked. A scientifically based rationale may be submitted in lieu of conducting studies with multiple radiolabels if no cleavage is anticipated. However, if cleavage of the molecule is evident, the applicant may be required to conduct an additional study with a radiolabel that tracks the portion of the molecule that is cleaved.
- The appropriate amount for injection onto the HPLC-RAM system is based on the need for a minimum of discriminant signal per peak in the chromatogram. Radiolabels, which ideally do not affect the physicochemical properties of the compound should allow both the qualitative and quantitative determination of the parent compound and its metabolites. When different concentrations of test substance are used, the total radioactivity percentage converted to metabolites at these concentrations should be consistent.
- Incubations conditions (e.g. million cells/mL and protein concentration) should be carefully adjusted to prevent excessive non-specific binding in the assay. In addition to protein removal, the addition of organic solvent to the sample disrupts any binding between parent compound and/or metabolites and the proteins present in solution, and thus, the analyte concentrations obtained represent total concentration, which is equivalent to the sum of bound and unbound (free) analyte concentrations. If there is a considerable fraction of non-extractable radioactivity, it could be due to binding of reactive intermediates to macromolecules and may need further investigations.
- The ability of the procedure to recover radioactivity added should be demonstrated. Recovery is typically inclusive of test material recovered in the extraction solvent, pellet residue and appropriate vial washes. Overall recovery should ideally be quantitative for applied radioactivity (i.e. $>$ 90%). If not, further investigation may be required to ascertain the cause of the loss.
- When a PBK model is considered to be used and taken into account that the analytes under study depends, e.g. of an active transport process, concentrations of the parent compounds and/or the metabolites in the intracellular space may be higher or lower than the extracellular space. In this case, and to be efficiently used, the intracellular concentration of compound must be carefully estimated to be incorporated in the model.
- The questions about the use of HRMS methods to circumvent the need for radioactive material in identifying metabolites in complex mixtures are controversial. It appears that radioactive and non-radioactive approaches are generally complementary, especially in the context of metabolite identification.

D.2. Identification and Characterisation

Identification refers to the exact structural determination of components of the total radioactive residue. Typically, identification is accomplished either by co-chromatography of the metabolite with known standards using two different systems or by techniques capable of positive structural identification such as MS, NMR, etc. In the case of co-chromatography, chromatographic techniques utilising the same stationary phase with two different solvent systems (e.g. various pH) could be considered to be an adequate two-method verification of metabolite identity, if the methods separation could be demonstrated as orthogonal. Identification by co-chromatography should be obtained using two dissimilar, analytically independent systems such as reversed-phase, HILIC or ion-exchange HPLC, as example.

- For routine metabolite identification (Met-ID), the complete LC–MS workflow can be greatly supported by dedicated Met-ID application managers implemented in the MS data acquisition and data processing software of the LC–MS instrument. Data processing is based on the comparison between data files acquired for biological samples and for blank/control samples,

thus searching for differences in the two data sets that can be correlated to the parent compound or its metabolites.

- It is generally believed that a single chemical shift, *m/z* value or other singular chemical parameter is insufficient for metabolite identification. The latter traditionally involves extraction, isolation and purification followed by elemental analysis, accurate mass measurement, ion mass fragmentation patterns, NMR (1H, 13C, 2D) and other spectral data such as IR, UV or chemical derivatisation.
- The accurate mass data together with high spectral accuracy (isotopic pattern determination) provide key information for molecular formula assignment. Characteristic *m/z* shifts that occur in Phase I and Phase II metabolisms could be anticipated. From an MS point of view, the biotransformation of the analyte results in metabolic m/z shifts relative to the m/z of the parent compound, e.g., a shift of −14.052 Da due to demethylation and +15.995 Da due to hydroxylation, N-oxide or sulfoxide formation. Therefore, from the *m/z* shifts in a metabolite, one can often investigate more carefully which biotransformation reaction has occurred.
- Metabolites often share the same elemental composition and molecular mass, therefore without the use of standards, the discrimination of isomeric compounds remains challenging. Fragmentation patterns of the profiled metabolites are necessary for their unambiguous identification by elucidating the biotransformation.
- To obtain this additional structural information, molecular ion species can be fragmented in tandem mass spectrometers using MS/MS or MSn experiments. MS/MS or MSn data can be acquired in the same analytical run as accurate m/z full scan profiling (on-line MS/MS), by signal-dependent precursor ion selection (MS/MS) or non-selective fragmentation of all ions (MSE/DDA) or a two-stage CID system integrated with ion mobility mass spectrometry (IMS-MS).

D.3. Stereoisomers

Currently, about 30% of the marketed pesticide active substances contain a chiral centre (Ye et al., 2015). The majority of chiral test substances exhibit slight differences in metabolism; however, some notable examples are available where major differences in metabolism of enantiomers occur (UK DEFRA, 2003). In these cases, the analytical workflow described in the EFSA guidance (EFSA, 2019b) can be followed in the comparative *in vitro* metabolism studies (Nallani et al., 2017; EFSA, 2019b).

When an accurate identification of isomers is necessary, the development of direct or indirect chiral chromatographic methods that unambiguously separate different stereoisomers is required (Zhao et al., 2021). Here, NMR spectroscopy can be very powerful in determining structural configurations of stereoisomers. If there are differences in metabolism of stereoisomers, the toxicological significance of these differences can be further considered as defined by the EFSA guidance (EFSA, 2019b; see also de Albuquerque et al., 2018).

Appendix E – Reactive metabolites

Reactive unstable metabolites and metabolism-connected free radicals are significant initiators of tissue injuries and toxicities, e.g. genotoxicity (Stachulski et al., 2013). Their detection is necessary especially in connection with chemicals exposing humans and other living organisms to a significant extent. The first successful *in vitro* test to screen and predict potentially genotoxicity of chemicals after biotransformation was the Ames test, which is based on the use of metabolic activation system, induced rat liver S9 fraction (see OECD 471). The potential production and presence of reactive metabolites in early metabolism tests for toxicological risk assessment of plant protection products could be anticipated on the basis of the approaches listed briefly below (Park et al., 2011; Pelkonen et al., 2015). There are no regulatory approved strategies to assess the potential role of reactive metabolites; in a case of suspected toxicological significance, a scientifically justified research strategy can be developed.

Initial *in silico/in vitro* studies

- Pre-screening for potential structural alerts using *in silico* methodology against existing databases (e.g. QSAR, Read-across, etc.).
- Use of various trapping agents in *in vitro* systems based on subcellular preparations or isolated/cultures cells, especially hepatocytes.
- Detection of covalently (tightly) bound radioactivity in *in vitro* systems (i.e. coming from comparative *in vitro* metabolism studies, see Section 5)
- Tentative identification of trapped/bound reactive metabolites by mass spectrometric techniques.
- Formation of a reactive metabolite in human cellular systems.
- Identification of metabolising enzymes producing reactive metabolites using several types of *in vitro* systems.
- Biomarker-based screening of potential (available) *in vitro* targets for identified reactive metabolites.

It is also possible to search metabolite adducts in the target tissue (e.g. liver 2D-protein gel electrophoresis) or plasma (e.g. albumin adduct) or urine (distal products of a reactive metabolite).

As an example of specific adverse reaction, a major hypothesis used to explain chemical-induced hepatotoxicity is that many toxic chemicals become metabolically activated in the liver to reactive metabolites that bind covalently to tissue proteins (Baillie and Rettie, 2011). The most frequently used method involves performing *in vitro* incubations with liver microsomes (usually from rat and human) in the presence of the cofactor for cytochrome P450 function (NADPH) and the nucleophile Glutathione (GSH) (10 mmol/L) or other nucleophilic trapping reagents such as potassium cyanide (KCN), and semi- carbazide (see e.g. Lassila et al., 2015; Pelkonen et al., 2015). The capturing of reactive metabolites with GSH (or other trapping agent), followed by structural characterisation by LC/MS/MS and NMR, is a very useful strategy because it does not require the synthesis of radioactive drug analogues.

Appendix F – OpenFoodTox

OpenFoodTox[5] contains hazard data from all EFSA chemical assessments since its creation in 2002 and is updated on a yearly basis. It currently includes nearly 5,000 unique substances (4,950 substances) in over 2,040 Scientific Opinions, Statements and Conclusions through the work of its scientific Panels, Units and Scientific Committee. For regulated products and contaminants, these risk assessments have been performed by six scientific panels and five supporting units. In details, it contains data for 2,600 food ingredients, 1,200 pesticides, 350 contaminants, 800 feed additives and 300 food contact materials. Overall, nearly 11,000 toxicological endpoints are available in the database for human health (2,050), non-target and target animal species (approx. 3,000 and 350), species in the water compartment (~ 3,100), species in the soil compartment (~ 2,500). In the pesticide area, reference points for the critical toxicological effect are available from test species (rat, rabbit, mice, dog). Current developments for 'OpenFoodTox 2.0' include (1) enriching the data model to include more chemical properties such as physico-chemical properties, TK (metabolism, bioaccumulation), mechanisms of toxicity, (2) exposure data and (3) linking the database to computer models used to predict TK and toxicity of chemicals. It is foreseen that in the future *in vitro* comparative metabolism data for pesticide active substances will be included and can support inputs for modelling pesticide active substance kinetics using generic PBK models within the TK plate of EFSA. *In silico* models using OpenFoodTox have been developed for ecological risk assessment (bees and rainbow trout) and human risk assessment (rat toxicological data) (Dorne et al., 2021). These *in silico* models provide alternative means to animal experiments for the hazard identification and characterisation of chemicals, and are becoming of increasing interest in the risk assessment community to deliver the 3Rs (replacement, reduction, refinement) (particularly since the banning of animal testing for the approval of cosmetics as consumer products (Regulation (EC) No. 1223/2009, Art.18(2)).

[5] https://www.efsa.europa.eu/en/data/chemical-hazards-data

Appendix G – Robust Study Summary Report

The PPR Panel proposed the following robust study summary report based on existing OECD Harmonised Templates (OHT) 201 and 58, the Guidance Document on Good *In vitro* Method Practices (GIVIMP), OECD Series on Testing and Assessment, No. 286, OECD Publishing, Paris (OECD, 2018a); and the experimental layout recommends in this opinion (i.e. hepatocytes as test systems). Although not comprehensive, the robust study summary can include the following information:

Data source

Reference

Method used

METHOD USED (description)

Name/identifier of the method (if available)

Deviations

Short description of the method and how it is relevant to the endpoint being investigated

Parameters analysed/observed

Applicability domain of the test method

Guideline, quality system/standards or guidance followed (e.g. OECD Guidance Document no. 286 on Good In vitro Method Practices (GIVIMP), scientific opinion)
GLP compliance.

Test system

DESCRIPTION

Test system: primary hepatocytes (suspended or plated)

Description of the number of passages, mass, volume, dimensions, if applicable

If purchased:

commercial source

supplier

species (i.e. human, rabbit, dog, mouse, rat)

pooled (half females and half males for humans; sex used in toxicity studies for animals; n. of donors, health status of donor; same batch, if possible)

media used and incubation conditions (cell culture medium characteristics, incubation temperature, pH, humidity, CO_2)

METABOLIC COMPETENCE OF THE TEST SYSTEM

Describe the knowledge about the metabolic competence of primary hepatocytes (i.e. Phase I and/or II biotransformation capacity) of the test system under remarks

Test and control items (positive and negative control)

CHARACTERISATION OF THE TEST AND CONTROL ITEMS

Chemical name and chemical identification (e.g. IUPAC, CAS number)

Structural formula and molecular weight

Source (i.e. manufacturer or supplier)

Lot/batch number and expiration date

Purity including information on contaminants, isomers

Volatility

Stability and storage condition

Analytical method for quantification of the test chemical

RADIOLABELLING INFORMATION (IF APPLICABLE)

In line with the *in vivo* rat TK study

Structural formula with the radiolabel

Radiochemical purity

Specific activity

Expiration date of radiochemical substance

FORM AS APPLIED IN THE TEST (if different from that of starting material)

Specify the relevant form characteristics if different from those in the starting material, such as state of aggregation, shape of particles or particle size distribution

OTHER SPECIFICS

Other relevant information needed for characterising the tested item, e.g. adjustment of pH, osmolality and precipitate in the culture medium to which the test chemical is added

Detection method

DETECTION METHODS FOR METABOLITE PROFILING AND IDENTIFICATION

Analytical method: semi-quantitative, relative quantification

Instrument characteristics (e.g. type, model)

Complete description of the analytical method including details on methods for analytical separation (e.g. chromatography) and analytical detection (e.g. mass spectrometry), limits of detection and quantification, variability and recovery efficiency, matrix used for standard preparations, internal standard, etc.

Statistical method

Description and statistical method used for exclusion of time points and/or runs

Test Design

PRELIMINARY EXPERIMENTS (description)

Cytotoxicity (e.g. trypan blue exclusion assay, WST-1 or resazurin reduction assay, LDH leakage assay)

Metabolic competence of hepatocytes from all species (e.g. incubation with positive control)

Non-specific binding assessment (recovery in the cell fraction and cell medium, by mass balance derive plastic binding) (Appendix H, GIVIMP)

Pilot *in vitro* clearance study

TEST AND CONTROL ITEMS PREPARATION

Test item

Concentration selection of the test item (e.g. 0 μmol/L, 1 μmol/L, 10 μmol/L and 100 μmol/L)

Method of preparation of stock and working solutions

Vehicle/solvent and medium characteristics

Dilution steps/dose intervals

Control Items: indicate whether solvent/vehicle controls, negative controls and/or positive controls (i.e. positive reference substances) were tested concurrently

EXPERIMENTAL CONDITIONS (main experiment – to be compiled for all species)

Number of replicates

Technical replicates (e.g. 3).

Experimental conditions

Concentration of the test system (e.g. 1 million cells/mL)

Cell viability of hepatocytes suspension after thawing (at least 80%)

Concentration of test and reference/control items (at least three, including the highest one)

Test set-up

Incubation conditions

Incubation temperature, CO_2 concentration, humidity level, etc.

Incubation time

Short term (e.g. 4 h).
Long term (e.g. 24 h).

Sampling time/time points (e.g. at least 3)

Data analysis for the test item results

Acceptance criteria: criteria for when results can be accepted, i.e. a set of well-defined parameters describing aspects of the method such as range for positive and negative controls (GIVIMP, OECD, 2018a).

Data calculation and statistics (parametric or non-parametric): provide the method used to calculate the results from raw data to the parameters calculated, such as normalisation, use of calibration curve, subtraction of control values, calculation of averages, Standard deviations etc.

Results and discussion

RESULTS FROM THE PRELIMINARY EXPERIMENTS

In vitro clearance (e.g. tabulated results of individual cells, standard deviations). Please, note that the information reported should allow to conclude on whether a 4-h incubation with hepatocytes in suspension will sufficient to reach the minimal required conversion of 20% of the parent chemical, or whether a 24-h incubation with plated hepatocytes would be required.

Cytotoxicity (e.g. tabulated results of individual cells of all selected species, mean and standard deviation for all tested concentrations, controls and time points).

Metabolic competence.

Non-specific binding assessment.

Discussion of the results (e.g. concentration range selected).

RESULTS FROM THE MAIN EXPERIMENT: describe the most relevant results with regard metabolites profiling, identification and characterisation, metabolic pathway, other (e.g. enzymatic activity)

Parameter (e.g. metabolism observed after 4 h in hepatocytes)

Results for the parameter (e.g. tabulated results from individual cells for each concentrations time points and runs; results clearly presented, including negative and failed runs; measurement including peak area)

Observation (e.g. precipitation, cytotoxicity)

Acceptance of the results (based on acceptance criteria)

Interpretation of the results and discussion (e.g. UHM, DHM identification)

Executive summary

Appendix H – Determination of disproportionate human metabolite

As indicated in the opinion, it is recommended to include in each experiment per species at least three time points and three substrate concentrations. Including more than two time points provide insight into the temporal trend for a given metabolite, which better facilitates the identification of a DHM. In Figure H.1, hypothetical data from a theoretical experiment are presented on the time-dependent formation of a metabolite obtained with human hepatocytes (white symbols) and hepatocytes from a test species (grey symbols). In this example, the metabolite would be considered a DHM in two situations:

1) the situation in which animal data are represented by the grey squares (all time points factor difference > 4), and
2) the situation in which the animal data are represented by the grey triangles (1 of 3 timepoints factor difference > 4).

In the situation in which the animal data are represented by the grey circles, the metabolite would not be considered a DHM (all time points factor difference < 4). It must be noted that for the situation in which the animal data are represented by the grey triangles, inclusion of more than two time points has clear added value, as the temporal trend is not linear as for the other situations, and if only one time point would have been included, the metabolite may not have been identified as DHM. In reality, this can be a metabolite that is further metabolised in the test species (e.g. a second oxidation or a conjugation) which may be absent in humans. In that situation, the metabolite has been tested in the related animal study but (relative) exposure may be much lower in the test species than in humans, and therefore, the metabolite is considered as being disproportionate.

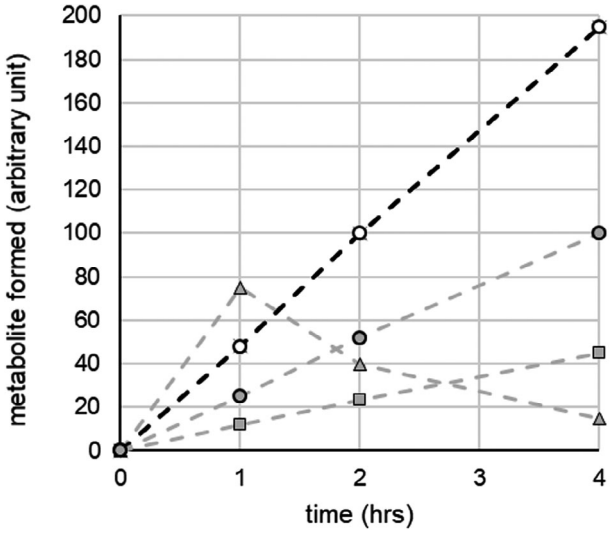

Figure H.1: Hypothetical examples on time-dependent formation of a given metabolite with incubations using hepatocytes of humans (open black circles) and a test species (grey symbols, three different scenarios). In this example, the human metabolite would be considered a DHM in the situation in which the test animal data would be represented by the grey squares or grey triangles, and would not be considered a DHM when the test animal data would be represented by the grey circles

Testing more than one concentration (i.e. at least three) is also important. As indicated in the opinion, a high concentration needs to be included (100 μmol/L when possible, regarding cell viability and solubility), providing the highest chance on detecting potential UHM. However, it must be noted that at these high concentrations, metabolic reactions may get saturated for one species, but possibly not for the other, indicating that for the identification of a possible DHM, inclusion of lower concentrations in which saturation is less likely, is important. In Figure H.2, hypothetical data from a theoretical experiment are presented on the time-dependent formation of a metabolite obtained with human hepatocytes (open black circles) and hepatocytes from a test species (grey circles). In this example, reaction for the formation of the metabolite gets saturated for humans but not for the test species. Therefore, when only including a high concentration (Figure H.2C), this metabolite would not

be identified as a potential DHM, whereas it would be identified as such when also including the lowest concentration (Figure H.2A). In general, including three concentrations (e.g. one low, one intermediate and one high concentration) provides better data to identify a possible DHM, giving insights into possible species-dependent differences in saturation of biotransformation reactions.

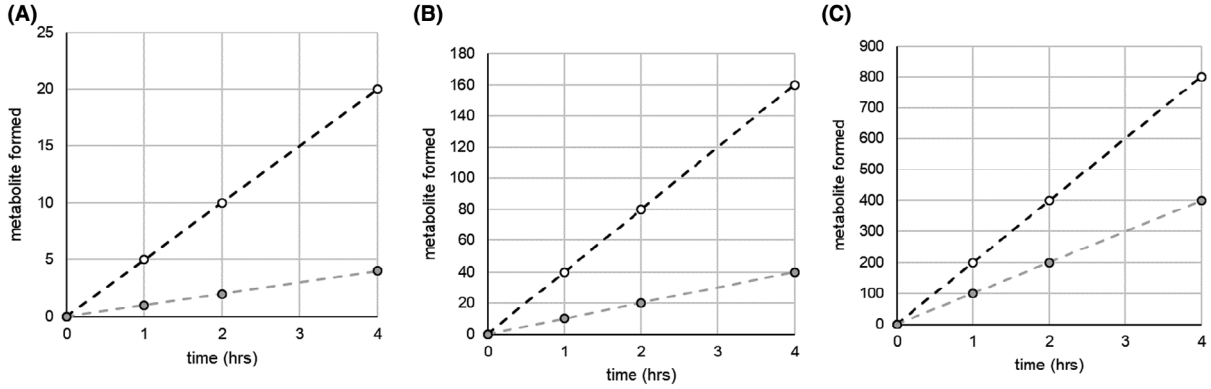

Figure H.2: Hypothetical example on the time-dependent formation of a given metabolite with incubations using hepatocytes of humans (open black circles) and a test species (grey circles) at a low (A), intermediate (B) and high (C) concentration of the test item. In this example, the metabolite would only be identified as DHM upon incubation with the low test item concentration (A)

Appendix I – Weight of evidence reporting table

This appendix proposes a generic example to apply the weight of evidence (WoE) approach on the assessment of comparative *in vitro* metabolism of pesticide active substances between humans and test species. The scientific rationale behind is to evaluate whether all significant metabolites formed in the human *in vitro* test system, as a surrogate of the in vivo situation, are also present at comparable levels in the animal species tested in toxicological studies and, therefore, if their potential toxicity has been appropriately covered by animal studies. In order to address such questions, the assessor needs to evaluate whether unique or disproportionate human metabolite(s) (UHM and DHM) are formed *in vitro* (hepatocytes) and such differences in the profile of metabolites across species can be either qualitative or quantitative. Hence, the generic WoE aims to answer related questions which can be formulated as follows:

Is a UM or DHM produced in humans compared to the test species?
Is the UM or DHM assessed by the toxicological studies in the test species?

All available evidence to identify the UHM or DHM can be assembled in pieces[7] and lines of evidence[8] (LOEs), weighted and integrated using the weight of evidence approach described in detail in the WoE Guidance document from EFSA's Scientific Committee (EFSA Scientific Committee, 2017a). In this context, depending on the evidence available, this will allow to reach a conclusion on the biological relevance of the human metabolites with regard to their toxicity (EFSA Scientific Committee, 2017b):

LOE1: In vitro metabolism in rats or/and other test species.
LOE2: In vitro metabolism in humans.
LOE3: In vivo metabolism and TK in rats or other test species.
LOE4: In silico evidence in rats, test species or humans.

Methods for weighing and integrating hazard evidence have been described elsewhere and include qualitative methods (listing, best professional judgement, semi-quantitative methods (causal criteria, logic); quantitative methods (scoring, indexing and quantification) (Linkov et al., 2009; EFSA Scientific Committee, 2017a). The methods of choice to be applied will depend on data availability, context of the assessment, complexity of the method, time constraints and resources and the assessor should provide a rationale for choosing a particular method. A key aspect for weighing and integrating the evidence is the assessment of the reliability, relevance and consistency of the evidence and the iterative nature of the process (EFSA Scientific Committee, 2017a).

Conclusion and summary of results

Table I.1 summarises the WoE assessment with regard to comparative metabolism of pesticide active substances in humans and test species and its toxicological consequences.

Table I.1: Weight of Evidence reporting table

Question (s)		**Q1: Is a UM produced in humans compared to test species?** **Q2. Is there a DM in humans compared to test species?** **Q3. Is the DHM assessed by the toxicological studies in test species?**
Assemble evidence	**Select evidence**	*In vitro*, *In vivo* and *in silico* evidence, literature, previous assessments, etc.
	Lines of evidence (LOE)	LOE1: *In vitro* metabolism in rats and other test species LOE2: *In vitro* metabolism in humans LOE3: *In vivo* metabolism and TK in rats or other test species LOE4: *In silico* evidence in rats, test species or humans
Weigh the evidence	**Methods**	Semi-quantitative scale (low, moderate, high) Quantitative scale
	Results	Tabular forms for the weighing of each LOE

[7] Piece of evidence: a broad term used to refer to distinct elements of evidence that may be combined to form a line of evidence, e.g. a single study, expert judgement or experience, a model or even a single observation (EFSA Scientific Committee, 2017a).
[8] Line of evidence: set of evidence of similar type (EFSA Scientific Committee, 2017a).

Question (s)		**Q1: Is a UM produced in humans compared to test species?** **Q2. Is there a DM in humans compared to test species?** **Q3. Is the DHM assessed by the toxicological studies in test species?**
Integrate the evidence	**Methods**	Semi-quantitative scale/Expert judgement/Probability scale
	Results	The WoE assessment concludes that: 1) Metabolite A is a UHM compared to test species 2) A DHM is formed compared to test species 3) The DHM is covered by the toxicological studies in test species

Appendix J – Overall Uncertainty Factor Table

Avoiding additional animal testing of a DHM by using an additional uncertainty factor for the derivation of HBGV is described in Section 6.1. Table J.1 shows some examples based on the results of the *in vitro* testing.

Table J.1: Results of the calculation for an additional uncertainty factor (UF) depending on the results of the *in vitro* testing and the resulting overall UF

Quantity of disproportionate metabolite (in % *in vitro*)		Ratio (human/animal)	Additional UF (RQh/a)	Overall UF
In animal	In human			
5	< 20	< 4	1	100
5	20	4	1	100
5	30	6	1.5	150
5	40	8	2	200
5	50	10	2.5	250
5	60	12	3	300
5	70	14	3.5	350
5	80	16	4	400
5	90	18	4.5	450
5	100	20	5	500
10	< 40	< 4	1	100
10	40	4	1	100
10	50	5	1.25	125
10	60	6	1.5	150
10	70	7	1.75	175
10	80	8	2	200
10	90	9	2.25	225
10	100	10	2.5	250
15	< 60	< 4	1	100
15	60	4	1	100
15	70	4.67	1.175	117.5
15	80	5.33	1.33	133
15	90	6	1.5	150
15	100	6.67	1.67	167
20	< 80	< 4	1	100
20	80	4	1	100
20	90	4.5	1.15	115
20	100	5	1.25	125
25	< 100	< 4	1	100
25	100	4	1	100