TECHNICAL REPORT

APPROVED: 26 March 2020

doi:10.2903/sp.efsa.2019.EN-1837

Outcome of the pesticides peer review meeting on general recurring issues in mammalian toxicology

European Food Safety Authority (EFSA)

Abstract

The technical report reflects the outcome of the discussion during the mammalian toxicology experts' meeting on general recurring issues, which took place in Parma in October 2019. The issues discussed and identified during the EFSA peer review of pesticide active substances under Regulation (EC) No 1107/2009 were mainly related to: the assessment of substances with endocrine disrupting properties, the assessment of impurities, isomers and metabolites, the genotoxicity of mixtures, the use of *in silico* methods for the assessment of genotoxicity and the use of the benchmark dose instead of the NOAEL. In addition, EFSA provided an update of the ongoing activities such as the in vitro metabolism, dermal absorption and developmental neurotoxicity.

Key words: pesticides, mammalian toxicology, endocrine disruption, impurities, metabolites, genotoxicity, benchmark dose

Requestor: EFSA

Question number: EFSA-Q-2020-00170

Correspondence: pesticides.peerreview@efsa.europa.eu

Acknowledgements: EFSA wishes to thank the following experts for the support provided to this scientific output: Laura Ruggeri, Tamara Coja and Carsten Kneuer.

Suggested citation: EFSA (European Food Safety Authority), 2020. Technical report on the outcome of the pesticides peer review meeting on general recurring issues in mammalian toxicology. EFSA supporting publication 2020:EN-1837. 26 pp. doi:10.2903/sp.efsa.2020.EN-1837
ISSN: 2397-8325

Summary

During the EFSA peer review of pesticide active substances under Regulation (EC) No 1107/2009, several aspects in the area of mammalian toxicology were identified by EFSA that needed discussion with experts from national Authorities in order to enhance the harmonisation of the risk assessment of active substances. To this purpose a general meeting was organised in October 2019 (Pesticide Peer Review Meeting 17, 16-18 October 2019).

EFSA presented the experience gained so far on the assessment of endocrine disrupting properties by applying the EFSA/ECHA Guidance and the feedback received from stakeholders on the use of such guidance. A list of relevant points for the potential update of the Guidance was provided and discussed. The possibiliy to use the ECHA Guidance on the assessment of the relevance of impurities for pesticides was also discussed together with general issues related to the assessment of isomers and metabolites found as residues or in groundwater. Genotoxicity was also discussed in particular in regard to mixtures and *in silico* methods potentially predicting specific genotoxicity endpoints. In addition, the experts examined the possibility to use the benchmark dose approach in toxicological studies where a NOAEL cannot be set and a LOAEL is derived instead. Finally, EFSA presented an update on the ongoing activities in regard to the *in vitro* interspecies comparative metabolism, the OECD project on dermal absorption, developmental neurotoxicity and adverse outcome pathway (AOP) development.

There are several general issues that were not discussed but should be taken into consideration for future discussion. These include, but are not limited to, the use of historical control data.

Recommendations were compiled on the basis of the discussions and conclusions achieved at the meeting and will be considered for implementation.

Table of contents

1. Introduction

During the EFSA peer review of pesticide active substances under Regulation (EC) No 1107/2009[1], EFSA identified several general recurrent issues in the area of mammalian toxicology which deserved experts' consultation and agreement in order to enhance the harmonisation of the risk assessment of active substances. Recommendations will be listed for further consideration.

To this purpose a general meeting was organised in October 2019 (Pesticide Peer Review Meeting 17, 16-18 October 2019). Member States representatives with expertise in toxicology attended this meeting. One member of the EFSA Scientific Panel on Plant Protection Products and their Residues (PPR Panel) participated as an external expert in her private capacity.

EFSA proposed to discuss the experience gained so far on the assessment of endocrine disrupting (ED) properties applying the recent ECHA/EFSA guidance (2018) and the feedback received from stakeholders on the use of such guidance.

Considering that a guidance for the assessment of impurities as regards the equivalence of technical material is currently available (European Commission, 2012) and that a guidance on the assesment of the relevance of the impurities in technical material is still missing for pesticides, EFSA also proposed to discuss the possibiliy to use the ECHA Guidance (still to be published) on the assessment of the relevance of the impurities in biocides for pesticides.

Other general issues proposed for discussion were related to the assessment of isomers and metabolites found as residues or in groundwater, the genotoxicity of mixtures and the use of *in silico* methods for predicting specific genotoxicity endpoints.

In addition, EFSA identified the need to discuss the possibility to use the benchmark dose (BMD) approach for the assessment of active substances during approval or renewal of approval and asked to Member States (MSs) to provide examples to be discussed on the application of the BMD approach.

Moreover, EFSA proposed to present an update on the ongoing activities in regard to the *in vitro* interspecies comparative metabolism, the OECD (Organisation for Economic Co-operation and Development) project on dermal absorption, developmental neurotoxicity and adverse outcome pathway (AOP) development.

Finally, some recommendations on the basis of the discussions and conclusions achieved at the meeting have been formulated and will be considered for implementation.

In addition, the comments received from Member States on the draft technical report following the written procedure launched from 19 December 2019 to 15 January 2020 are available as background document of this technical report. It is noted that the written procedure was performed with the purpose to enhance readability and correct possible inconsistencies. Since the scope of this technical report was to reflect the meeting discussions and conclusions, the commenting round was not meant to reopen the discussions or to change the outcome of the meeting.

2. Points of discussion and meeting's conclusions

2.1. Assessment of Endocrine Disrupting (ED) properties

Background

EFSA presented an overview of the experience gained by EFSA regarding the ED assessment of active substances since the implementation of criteria through the application of the EFSA/ECHA guidance document for the identification of endocrine disruptors under Regulations (EU) No 528/2012[2] and (EC)

[1] Regulation (EC) No 1107/2009 of the European Parliament and of the Council of 21 October 2009 concerning the placing of plant protection products on the market and repealing Council Directives 79/117/EEC and 91/414/EEC. OJ L 309, 24.11.2009, p. 1-50.

[2] Regulation (EU) No 528/2012 of the European Parliament and of the Council of 22 May 2012 concerning the making available on the market and use of biocidal products Text with EEA relevance. OJ L 167, 27.6.2012, p. 1–123.

No 1107/2009 (ECHA/EFSA, 2018). The feedback received from stakeholders on the application of the guidance during this period was also presented. A list of selected items related to the use of the Guidance was provided and discussed.

EFSA points for discussion

- Rapporteur Member State (RMS) to present the ED assessment under the Volume 1 of the draft (renewl) assessment report (DAR/RAR), Chapter 2.10.
- RMS to present the study summaries for individual mammalian toxicological and ecotoxicological ED studies in the Volume 3, B.6 (mammalian toxicology) and B.9 (ecotoxicology).
- RMS to present the conclusions of the weight of evidence (WoE) assessment (including the mode of action (MoA) analysis) and, where necessary, a proposal for a further testing strategy, as stated in the guidance.
- The submission of the excel spreadsheet is not mandatory, though strongly recommended in compliance with the template in the Appendix E.1 of the EFSA/ECHA guidance. The RMS should check the spreadsheet provided by the applicant and, where necessary, amend the information reported. The excel spreadsheet should be then submitted as an Annex to the Volume 1 as a stand-alone document.

Please also refer to the instructions in the Administrative guidance on submission of dossiers and assessment reports for the peer-review of pesticide active substances (EFSA, 2019c).

Meeting's discussion and agreed conclusion

- EFSA presented the ED assessments done so far since the entering in force of the EFSA/ECHA guidance within the peer review process (43 substances at the time of the meeting) and a general interest was expressed by MSs regarding the database compiled by EFSA.
- ToxCast: it was agreed to have an annex to the DAR/RAR with a screenshot or a stand-alone pdf-printout of all the relevant ToxCast, including the date of data collection. This will be necessary to keep track of any possible changes in ToxCast over time. At least a summary containing all relevant information from ToxCast provided by the applicant and checked by the RMS, should be included in the Volume 3 together with an evaluation by the RMS. Some MS experts expressed the need for more training on the use of ToxCast.
- Waiving: EFSA presented an overview of the cases for which the ED assessment was waived. The most frequent rational for waiving was based on: the intrinsic physico-chemical and the toxicological properties of the substances. Specific examples include potent AChE inhibitors, dose-limiting local irritant effects and dose-limiting induction of methemoglobinemia. In line with the EFSA/ECHA guidance, EFSA clarified that the waiver should be scientifically justified and made on a case-by-case basis. It was further clarified that it is expected that the available information should be summarised and evaluated and the reason for waiving the ED assessment should be contextualized in accordance with the EFSA/ECHA guidance. Remaining uncertainties should be therefore discussed in light of the above mentioned rationale.
- Completeness of the EAS (estrogen, androgen, steroidogenic) dataset for adversity: according to the EFSA/ECHA guidance, the dataset for EAS-mediated adversity for a specific substance is considered sufficient only when studies according to the OECD TG (test guideline) 416 (latest version from 2001) or OECD TG 443 (including the F2 generation) are available (level 5 studies). It was agreed that the dataset can be considered sufficiently investigated also in the case the old version (before 2001) of the OECD TG 416 was applied providing that all relevant parameters, foreseen to be measured according to the new version of OECD TG 416, were measured.
- A scientific discussion was held to debate the sensitivity of the endocrine mediated endpoints included in the current OECD TG 416. There was a general consensus on the lacking of a clear ranking order of sensitivity for endocrine endpoints and, considering that identification of endocrine mediated adverse outcomes should be based on a consistent pattern of evidence, a number of additional endpoints should be considered by default when conducting a study in line with the OECD TG 416. This was considered as a best scientific practice and would allow a comprehensive evaluation of a relevant level 5 study. As a post-meeting action and here noted, EFSA also consulted with ECHA, and the following parameters were considered as a default best

scientific practice to be included in the protocol of the study carried out according to the OECD TG 416 i.e. the following parameters should be measured and reported in the study report and then in the DAR/RAR:

- anogenital distance of each F1 and F2 pups,
- presence and number of nipples/areolae in all male F1 and F2 pups,
- histopathological assessment of the mammary gland in P0 and F1 adult males and females,
- sperm parameters measured always by default regardless if they have also been tested in the 90-days.

- Sufficiency of T (thyroid) dataset: comments where received from some MSs if the availability of THs (thyroid hormones) measurements should also be used to evaluate the sufficiency of the adversity related to the T-modality. EFSA clarified that in the old versions of the OECD TGs the measurment of thryoid hormones was optional. Therefore, in these cases, the lack of THs measurment cannot be used to conclude that the dataset for adversity is not complete. However, it should be noted that in the new versions of OECD TGs, THs measurment is mandatory. In addition, the dataset for thyroid can be considered complete on a case-by-case basis, pending whether the duration and doses selection allow a proper assessment of the thyroid histology (thyroid histopathology is generally considered more sensitive and informative than thyroid weight).
- Carcinogenicity and ED assessment: the adequacy of using the carcinogenicity studies in the ED assessment when the only effect observed in the available dataset is the tumorigenic effects on endocrine organ/s was discussed. It was agreed that a MoA analysis should be provided and data should be sufficient to demonstrate that the effect is, or is not, consequent to an endocrine mode of action and therefore, differentiate a hormonal carcinogen from a non-hormonal carcinogen. These data must be provided by the applicant. It was noted that, in the absence of a MoA analysis, the effects on many endocrine organs are considered EATS-mediated and therefore, in line with the EFSA/ECHA guidance, indicative of endocrine disruption.
- Hormones measurement: it was discussed if the current testing strategy should be updated by including the measurement of sex hormones to support the MoA analysis. EFSA indicated that liquid chromatography with tandem mass spectrometry (LC-MS/MS) represents the most accurate and specific method for the measurement of THs unless another method is justified by the applicant. One MS pointed out that immunochemical methods are well established and economic and should be considered equally acceptable if adequate quality control (e.g. calibration, successful round robin testing) is demonstrated. This was considered a plausible approach; however, the detection method should be fit-for-purpose and in case of THs evaluation, the assay should be sensitive enough to allow hormones estimation also in the pups which are considered a relevant sensitive population for THs disruption. In this case the LC-MS/MS is considered the most sensitive method.
- Uncertainty analysis: in the current EFSA/ECHA guidance the uncertainty analysis is assessed only for the MoA analysis. It was therefore questioned if uncertainty analysis should always be performed. It was agreed that the inclusion of a qualitative uncertainty analysis is highly recommended for a more transparent and comprehensive assessment of the WoE. The structure of the excel template for reporting the available information relavant for ED assessment would be amended accordingly.
- Assessment of thyroid disruption: based on the experience gained by EFSA in regard to ED assessment, EFSA discussed in more details how to assess thyroid disruption in line with the approach described in the EFSA/ECHA guidance. Among the substances assessed until now and identified as EDs, the conclusion in most of the cases was based on evidence of thyroid disruption.
  The assessment of thyroid disruption is complex since several mechanisms are involved. Thyroid adversity is relatively common in experimental toxicity studies, particularly in the rat. Therefore, the evaluation of the MoA leading to thyroid toxicity is relevant to conclude on ED properties.

Indeed, as outlined in the criteria to identify EDs set in the Regulations (EU) 2018/605[3] and (EU) 2017/2100[4], endocrine mediated adverse effects that are secondary to other toxicities (including liver toxicity) should not be considered for concluding that ED criteria are met. In this case, it is necessary to demonstrate by means of comparative MoA analysis that thyroid toxicity is secondary to e.g. liver toxicity. In the comparative MoA analysis, a MoA for thyroid toxicity and one for liver toxicity should be postulated in a comparative manner. The applicant should transparently tabulate the data in order to evaluate the dose- and temporal-response. The RMS should ultimately evaluate if the thyroid effects are secondary to the liver toxicity.
The EFSA/ECHA guidance also includes a specific approach when the option is to support a non-human relevance of the observed thyroid effect.
The assessment of human relevance is mainly applicable to those cases where the T-mediated effect is through a liver-mediated mechanism i.e. liver enzyme induction resulting in an increase of THs clearance. In this case, three pieces of information should be provided to evaluate whether the thyroid findings are likely or not to be human relevant: 1) analysis of T3, T4 and TSH in the repeated dose studies; 2) *in vitro* comparative studies to evaluate liver enzyme induction in the tested species (i.e. rat, mouse and dog) and humans; 3) evaluation of other potential *in vitro* mechanisms involved in the thryoid disruption. Finally, all the available evidences should be weighed, including interspecies differences and lack of any concomitant molecular initiating event.
Ultimately, the EFSA/ECHA guidance, in the Appendix A, provides further considerations to be applied when interpreting data from experimental animals, particularly in relation to potential neurological developmental effects related to thyroid adversity. In details, the Appendix A proposes to consider that: a) substances inducing histopathological changes in the thyroid, with or without changes in the circulating levels of THs, would pose a hazard for the human thyroid hormone insufficiency in adults as well as pre- and post-natal neurological development of offspring; b) substances that alter the circulating levels of THs without histopathological findings would still present a potential concern for neurodevelopment. During discussions, EFSA confirmed that a CAR/PXR-mediated MoA that can also be expected to be functional in humans, leading to an increased clearance of THs would be considered relevant.
Therefore, EFSA clarified that to further investigate the thyroid disruption, particularly when perturbations of THs in absence of histological changes are observed, is necessary to consider the hazard assessment in the most sensitive population as pups and offspring. In such cases, further testing (e.g. developmental neurotoxicity (DNT) study) should be perfomed to assess the likelihood of a substance to induce developmental neurotoxicity. Alternatively, a special study developed by US EPA, Guidance for Thyroid Assays in Pregnant Animals, Foetuses and Postnatal animals, and Adult Animals[5], can be performed to generate mechanistic data to confirm or refute the observed changes in circulating THs. This is stated in the current guidance.

- A possible update of the EFSA/ECHA guidance was also discussed but without precise timeline.

2.2. Guidance on the assessment of the relevance of impurities from ECHA Biocides: discussion on possible application for pesticides

Background

A definition of a "relevant impurity" is provided in ECHA Biocidal Product Regulation (BPR) guidance on information requirements and ECHA Guidance on applications for technical equivalence. However, different interpretations of the definition have been applied in biocides active substance approval among Member States. ECHA together with the Member States developed a guidance document with the aim

[3] Commission Regulation (EU) 2018/605 of 19 April 2018 setting out scientific criteria for the determination of endocrine disrupting and amending Annex II to Regulation (EC) 1107/2009. OJ L 101, 20.4.2018, p. 33–36.
[4] Commission Delegated Regulation (EU) 2017/2100 of 4 September 2017 setting out scientific criteria for the determination of endocrine-disrupting properties pursuant to Regulation (EU) No 528/2012 of the European Parliament and Council. OJ L 301, 17.11.2017, p. 1–5.
[5] Guidance for Thyroid Assays in Pregnant Animals, Fetuses and Postnatal Animals, and Adult Animals. Office of Pesticide Programs Health Effects Division Washington DC (October 24, 2005) Available online: https://www.epa.gov/sites/production/files/2015-06/documents/thyroid_guidance_assay.pdf

of clarifying the definition of relevant impurities and providing the principles and practical guidance for the identification of the relevant impurit(y)ies. The guidance provides consistent principles towards impurities that have a non-threshold mode of action and, for other types of substances, establishes concentration limits to not significantly contribute to the (eco)toxicity of the active substance. The document will be published soon on the ECHA website.

EFSA points for discussion

- Currently the toxicological relevance of impurities in pesticides risk assessment is performed by following the European Commission guidance document on the assessment of the equivalence of technical materials (European Commission, 2012). Considering that the process of testing the applicability of the ECHA Guidance to pesticides is still onging, for the assessment of the toxicological relevance of impurities for pesticides it is recommended to continue using the European Commission guidance document.

Meeting's discussion and conclusion

- The ECHA definition of the relevance considers not only the hazard (as currently done under pesticides) but also the amount of the impurity, in line with FAO (JMPS).
- The experts recognised that a consistent approach between ECHA and EFSA would be needed and discussed the possibility to apply the guidance also to pesticides: some MS experts agreed to apply the ECHA guidance, while others expressed some reservations. Overall, MSs agreed that additional examples would be needed to further discuss the possible application of the ECHA guidance to pesticides. In addition, MSs should further discuss the following points:
    1) what to do when no harmonised classification is available for an impurity,
    2) when to consider that "sufficient information" is available,
    3) how to proceed when a relevant impurity is not covered by the batches used in toxicological studies,
    4) whether the Appendix V of the European Commission guidance document should be amended by including the decision tree to support the assessment of the relevance of impurities provided in the ECHA guidance,
    5) the pros and cons of the ECHA approach in particular with regard to the assessment of the intrinsic hazard properties (as currently done for pesticides) regardless of the content, versus the consideration of the amount of impurity. As pointed out by several MSs, this approach could consider an impurity as not relevant because the concentration is below the threshold set for the relevance (even though it has more severe hazardous properties as the technical active substance), and consequently this impurity would not be monitored,
    6) the potential issues and possible solution regarding the assessment of the technical specification and the equivalence of the technical material when applying the ECHA guidance.
- The experts were asked to further reflect on the possible use of the ECHA guidance to pesticides and give feedback to their PAFF representatives, since the use of the guidance in the pesticides area should be agreed by risk managers and the Commission.

2.3. Guidance on isomers

Background

The EFSA Guidance on risk assessment for active substances of plant protection products that have stereoisomers as components or impurities (EFSA, 2019a) was presented: recommendations were provided on how to best address and assess the data requirements for active substances containing stereoisomers.

Meeting's discussion and conclusion

According to the guidance, changes of ≥10% of stereoisomeric excess in the relevant residue, with respect to the composition of the mixture tested, are considered relevant for the general toxicity assessment. However, EFSA explained that the 10% trigger should not be considered as a "hard trigger" given that there might be variations of the percentage of the stereoisomers. These variations can be attributed to different situations (e.g. analytical methods used, matrices' effects, etc…) and therefore these variations should be taken into account before deciding that general toxicity assessment is required. In the case of potential differences in the isomer composition, it was pointed out that ADME data as well as *in vitro* comparative metabolism studies can provide useful information for bridging/read-across. For racemic mixtures, which undergo degradation (isomeric shift), the scientific literature search provided by the applicant can be used to collect the data for bridging and risk assessment purposes.

2.4. Groundwater metabolites

Background

The guidance document on the assessment of the relevance of metabolites in groundwater proposed an *in vitro* test battery consisting of an Ames test, an *in vitro* mammalian gene mutation assay and an *in vitro* chromosome aberration assay (European Commission, 2003). This test battery is not properly assessing aneugenicity potential according to the current scientific knowledge (EFSA Scientific Committee, 2011).

EFSA points for discussion

EFSA recommends Member States to request an *in vitro* micronucleus test instead of an *in vitro* chromosome aberration test to properly cover aneugenicity in the case of assessment of groundwater metabolites. In the case an *in vitro* micronucleus test has not been submitted for a specific metabolite under evaluation, a data gap will be set in the EFSA conclusion.

Meeting's discussion and conclusion

The majority of the MS experts agreed that the European Commission guidance (European Commission, 2003) is out of date; the possible need of an update should be brought to the attention of the European Commission. EFSA indicated that a harmonised approach to assess groundwater metabolites would be welcome and needed also by ECHA. A data gap for an *in vitro* micronucleus test should be set in the EFSA conclusion when such information has not been provided for the groundwater metabolites assessed during the evaluation of a specific active substance. The MS experts pointed out that it may be difficult to set a data gap for an *in vitro* micronucleus test while the European Commission guidance referring specifically to the *in vitro* chromosome aberration test is still in force. However, EFSA clarified that the data gap is reflecting the current scientific state of the art for genotoxicity testing using experimental data (the *in vitro* micronucleus test has been shown to be as sensitive as the chromosome aberration tests for detection of clastogens and has the additional advantage of detecting aneugenic substances; Corvi et al., 2008). QSAR and read-across for the assessment of the genotoxicity is proposed as an alternative to experimental data for the residues metabolites, given that there are not specific data requirements for residues metabolites. This is not the same situation for groundwater metabolites for which the European Commission guidance requests experimental data.

2.5. Genotoxicity of mixtures

Background

In 2018, the Scientific Committee of EFSA addressed how to assess the genotoxic hazard of substances in chemical mixtures present in food and feed (EFSA Scientific Committee, 2019). The approach starts from the identification of the substances in the mixture as far as possible and is followed by the genotoxicity assessment of the substances of the mixture. The Scientific Committee discussed how to deal with mixtures where a fraction of substances has not been chemically identified and the possible limitations of *in vivo* genotoxicity testing of chemical mixtures. The Scientific Committee noted that it may be possible to deviate from the approaches proposed, if it can be scientifically justified. In the area of pesticides, very few active substances are mixtures. In addition, when dealing with pesticide active

substance mixtures, the majority are well defined mixtures (i.e. uncharacterised fraction is small). The current approach for assessing pesticide active substance mixtures is the whole mixture approach (WMA) even if they are well defined mixtures while according to the Scientific Committee statement (EFSA Scientific Committee, 2019) a component based approach (CBA) should be done.

EFSA points for discussion

EFSA recommends Member States to consider the statement of the Scientific Committee when dealing with pesticide active substance mixtures. An important consideration is that some of the pesticide active substance mixtures are botanicals for which special considerations have been given by the European Commission given their natural origin (European Commission, 2014) and this should be also taken into account by Member States.

Meeting's discussion and conclusion

The MS experts agreed to use the statement of the Scientific Committee on the genotoxicity of mixtures for the assessment of pesticide active substance mixtures. As a practical approach, the toxicologists of the Pesticide Peer Review (PREV) Unit will ask the support from the physico-chemical team whenever it is needed to double check whether the active substance under evaluation should be considered a mixture (including characterisation of the mixture). The approach proposed by the EFSA Scientific Committee statement will then be applied accordingly. The same approach can be followed by the RMS during the discussion with the applicant in the pre-submission meetings. It was clarified that the statement is not applicable to plant protection products.

2.6. Use of *in silico* methods for the assessment of genotoxicity

Background

To facilitate the implementation of the guidance on the residue definition for dietary risk assessment, EFSA outsourced (2017-2019) an evaluation of the applicability of *in silico* models (quantitative structure activity relationships, QSARs and read-across) for predicting the genotoxicity of pesticides and their metabolites and analysis of the impact of the metabolic structural changes on genotoxicity. A wide range of commercial and publicly available QSAR models were applied to the EFSA pesticides genotoxicity database. A detailed analysis on the performance and on the reliability of the predictions and possibility of combination of predictions from different models in order to improve the performance was done. The performance of different methodology of read-across for prediction of genotoxicity (point mutations and *in vitro* chromosomal aberrations) for about 60 case studies was investigated. The study included "automatic" 1:1 and many to one read-across based on structural similarity, mechanistic read-across and read-across based on weight of evidence. The impact of the structural changes in the molecule in result of metabolic or degradation processes to the genotoxic potential (point mutations only) of the substances was evaluated. This resulted in lists of structural changes that may, or may not influence the Ames mutagenicity. It was suggested the knowledge on these structural factors to be used in complementary to the knowledge on the structural alerts related with AMES mutagenicity in the assessment of the genotoxicity (point mutation) of metabolites. A stepwise approach using all the methodology investigated for assessment of genotoxicity of pesticides was developed. External scientific report presenting the above mentioned project is available on the EFSA website (Benigni R et al., 2019).

EFSA points for discussion

To perform a follow-up project to further explore the applicability of *in silico* tools to non-Ames endpoints and possibly also to other toxicological endpoints.

Meeting's discussion and conclusion

The *in silico* models (single models or used in different combinations) used in the EFSA outsourced project for prediction of genotoxicity showed good performance in the case of prediction of point

mutations, but poor performance in the case of prediction of other genotoxicity endpoints. The availability of experimental data and the quality of the experimental protocols, which affect the quality of the data used as training sets, were proposed as possible explanation of the worse performance of the models for non-Ames endpoints. This should be further explored in follow-up activities related to collection and curation of data for non-Ames endpoints. The experts agreed that the predictions for endpoints different than Ames still could be used as a part of the weight of evidence approach but not individually, considering their limited reliability. A recommendation was done to analyse and use all available and additional information provided by the model/software in order to evaluate the reliability and the acceptability of the prediction. The experts agreed that, in case of positive predictions testing should be done; in case of negative predictions, a case-by-case approach should be applied. Overall, the regulatory acceptance of the models for risk assessment was discussed and it is still considered as pending.

2.7. Toxicological assessment of metabolites found as residues

Background

In 2016, the guidance on the establishment of the residue definition for dietary risk assessment was developed by the Panel on Plant Protection Products and their Residues (PPR Panel) (EFSA PPR, 2016). The document provides directions for determining the metabolites that require hazard identification and characterisation using QSAR models, grouping, read-across, threshold of toxicological concern (TTC) and available data in combination, and for developing an appropriate testing strategy for these compounds, if needed. Although the guidance has not been taken note of by the European Commission and MSs, some of the scientific elements of the guidance are currently used by the applicants and MSs. However, the data included in the dossiers have not been consistently and/or sufficiently reported.

EFSA points for discussion

EFSA proposes a template for assessing QSARs reports in the DAR/RAR to allow an independent peer review (Appendix A). Examples are also provided in the Appendix B of this Technical Report. The template is based on the OECD principles for validation of QSARs and on the ECHA guidance used in the framework of REACH (ECHA, 2008). EFSA also proposes a table for summarising and integrating the evidence for genotoxicity (Appendix C) and an additional table for summarising all available data on metabolites in residues of plant and animal origin (Appendix D).

Meeting's discussion and conclusion

EFSA clarified that there is no need to run QSAR analysis and read-across when experimental data are available or when the metabolite is covered by the parent. When no experimental data are available and when the metabolite is not covered by the parent, the three genotoxicity endpoints should be covered (gene mutation-Ames and chromosome aberration-clastogenicity and aneugenicity) by two independent and reliable QSAR model for each endpoint. EFSA indicated that often the following is lacking in the DAR/RAR: information regarding how the models are independent from each other, summary details of the model and robust assessment of the reliability of the prediction for each QSAR, summary table integrating the evidence of genotoxicity and similarity's assessment of active substance and metabolite when they shared same structural alerts and/or prediction. In addition, often the grouping and its strategy are not clearly documented, and, even if specific experimental data are available, comparison to parent is not reported. Finally, a table summarising all available information is often missing. Therefore, EFSA presented templates (see Appendices A to D) that could be used by the applicant and assessed by RMS to assess QSAR reports and to collect the most relevant information (experimental data, QSAR analysis or read-across) and overall conclusions for the evaluation of metabolites. The feedback provided by MS experts on the use of such templates is the following:

1) some MSs expressed their willing to ask the applicant to use these templates for the assessment of metabolites in the next DARs/RARs;

2) concerning general toxicity it would be helpful to have an additional table similar to the summary table for genotoxicity to show the information with more structured details;

3) the template should also contain information on the reliability of the experimental data and QSARs, which is essential for the overall conclusion on the endpoint;

4) some MSs proposed to re-consider naming/terminology of endpoints indicated in the table "summary for integrating experimental evidence for genotoxicity" (to use gene mutation in bacteria, gene mutation in mammalian cells, clastogenicity *in vitro*, aneugenicity *in vitro*, clastogenicity/aneugenicity *in vivo* instead of gene mutation-Ames test and chromosome aberration-micronucleus); these tables available in the Appendices have been amended after the experts' meeting to reflect comments received by MSs.

5) to re-word the column on ADME data as percentage of the applied/absorbed dose in excreta, tissues and downstream metabolic pathway;

6) to include in the overview table under the "General toxicity" also information on the reference values (with uncertainty factors (UFs) and basis for derivation of the reference values);

7) to include these templates in the EFSA Administrative guidance on submission of dossiers and assessment reports for the peer-review of pesticide active substances as an Appendix (EFSA, 2019c).

EFSA confirmed that MSs are already invited to make available the (draft) template to applicants in order to improve the quality of current dossiers.

2.8. Use of benchmark dose (BMD) instead of NOAEL

Background

Traditionally, in the European pesticide peer review the critical no-observed-adverse-effect levels (NOAELs) are identified from the available dataset and acceptable reference values (e.g. acceptable daily intake (ADI)) are derived. In cases where adverse effects are observed in a critical study at all doses, a lowest-observed-adverse-effect level (LOAEL) is used as the starting point for risk assessment purposes and additional uncertainty/safety factor (mostly a factor of 3) is added to derive the reference value.

One of the flaws of the NOAEL/LOAEL concept is that the observation of the critical effect is restricted to one of the predefined dose levels used in the study. Additionally, studies with low power (e.g. small group sizes) are able to detect only relatively large effects, which tend to result in higher NOAELs.

Benchmark Dose (BMD) approach (EFSA, 2017b) estimates the dose that corresponds to a low, but measurable change in response (benchmark response - BMR; e.g. 5% increase in relative organ weight (continuous data) or a 10% increase in the incidence of a tumour (quantal data)). The choice of 10 or 5% or any other agreed % of relevant biological change is dependent on the toxic effect and agreed response level and is calculated by fitting mathematical models to the dose-response data.

A number of models is run, potentially averaged, and a lower 95% confidence limit calculated. The lower bound (BMDL) is needed as a potential reference (starting) point, and the upper bound (BMDU) is needed for establishing the BMDU/BMDL ratio, reflecting the uncertainty in the BMD estimate.

The two main models are the US EPA BMD software (https://www.epa.gov/bmds/about-benchmark-dose-software-bmds) and the National Institute for Public Health and the Environment of the Netherlands (RIVM) software for dose-response modelling and benchmark dose analysis (PROAST, https://proastweb.rivm.nl/). PROAST is also integrated in the statistical models platform at EFSA homepage (https://shiny-efsa.openanalytics.eu/app/bmd).

The presentation focused on some examples from the peer review process where BMD has been already applied, in parallel to the setting of NOAEL/LOAEL.

EFSA points for discussion

To implement more regularly the use of the BMD analysis, in particular in the case where no NOAEL is set for a specific parameter but a LOAEL is set instead, or also on critical endpoints to derive toxicological reference values.

Meeting's discussion and conclusion

Experts discussed some specific examples provided by two MSs on the application of the BMD approach to active substances during approval or renewal of approval. Pros and cons were highlighted. Experts agreed that BMD is not meant to replace the NOAEL for the time being, but NOAEL and BMD should be considered in parallel. In addition, more experience would be needed to run the model and for interpreting the results also for merged studies i.e. studies with combined results from different studies. The experts agreed that the most critical part is the selection of the critical effects size (CES), i.e. the BMR level used, which is depending on the parameter analysed but also on the control population. One MS proposed to use 5% as BMR for whatever parameter and increase it, if necessary. Overall, the experts agreed that the choice of the BMR is on a case-by-case basis and should be of 5% or lower for some types of effects or 10% or higher for other effects (e.g. for early precursor effects). In any case, the choice of the BMR should always be explained and well documented. The experts agreed that the application of the BMD modelling depends upon the available data and choice of parameters and should be carefully considered when conducting data analysis.

Some MSs indicated that they evaluated the possibility to use the approach for the active substances they are evaluating, but they concluded their substance was not an ideal candidate because NOAEL had already been set for all endpoints, and most studies have been conducted with three doses (which according to the consideration behind this method is not ideal for applying the BMD approach). The experts discussed the possibility to use the BMD approach on critical endpoints in the study used for setting reference values. A consensus was reached to apply the BMD analysis for the critical endpoints which are the basis of the reference values setting and also when no NOAEL is set for a specific parameter and a LOAEL is set instead.

EFSA reminded that the EFSA Standing working group will be able to provide support to MSs in BMD analysis issues, if needed. Finally, a pilot phase on the use of BMD was evoked, however, considering capacity limitations and lack of experience, no MS expressed its availability for a possible pilot phase. Some MSs voiced a need for training instead.

During the meeting, MSs and EFSA identified a need to exchange practical experience with the methodology, its challenges and solutions thereof in a dedicated workshop. Some MSs expressed their interest in supporting EFSA in organising such an event and EFSA will further consider this possibility.

2.9. Report of the *in vitro* metabolism comparative study workshop

Background

In 2018, EFSA held a workshop with stakeholders on the use of *in vitro* interspecies comparative metabolism studies in the pesticide risk assessment (EFSA, 2019b).

Particularly, the workshop focused on the use on the *in vitro* comparative studies to: 1) identify major/unique human metabolites in a line with the information requirement set in the Regulation (EC) No 1107/2009 and 2) role of *in vitro* metabolism studies to assess human relevance of toxicological animal data, specifically for endocrine-mediated thyroid toxicity based on the approach outlined in Appendix A of the EFSA/ECHA guidance to identify EDs.

EFSA presented the main outcomes of the workshop that are available in the report on *in vitro* comparative metabolism studies (EFSA, 2019b) and explained that the discussion on the use of these studies to identify major/unique human metabolites provided useful input for the development of an EFSA guidance. The draft guidance should be published at the end of 2020.

Meeting's discussion and conclusion

It was criticised that the report appears to state that testing one concentration of the active substance only would be sufficient. EFSA clarified that the workshop report reflects the opinion of the participants from industry, academia and regulators. It is not binding and an EFSA guidance is beeing developed.

2.10. OECD project on dermal absorption

Background

EFSA (acting on behalf of the European Commission) is co-leading with BfR the OECD Expert Group on Dermal Absorption for the revision of the Guidance Notes (GN) n. 156 on the setting of dermal absorption values to be used in pesticides/biocides risk assessments (OECD, 2011). The update is based on the analysis of data from human *in vitro* studies conducted with plant protection products (PPPs), that have driven the recent revision of the EFSA Guidance on Dermal Absorption (EFSA, 2017a).

Meeting's discussion and conclusion

EFSA presented an update of the activities related to the project: a revised version of GN 156 has been finalised in October 2019 (☐300 experts' group comments addressed) and a second consultation of the experts' group has been conducted. During the Annual Meeting in April 2020, the WNT (Working Group of the National Coordinators for the Test Guidelines Programme) will be updated on the status of the project and will be consulted on the way forward for the finalisation of the revised GN 156.

2.11. Developmental neurotoxicity (DNT)

Background

EFSA presented an update of the ongoing activities related with the working group dealing with the "Scientific Opinion of the PPR Panel for developing Integrated Approaches to Testing and Assessment (IATA) case studies on developmental neurotoxicity (DNT) risk assessment". In particular, EFSA updated on the work in progress in regard to *in vitro* studies (*in vitro* testing battery), the OECD Guidance and the development of the Scientific Opinion.

2.12. Developmental of an AOP relevant for the identification of substances having ED properties

Background

EFSA presented an update on the development of Adverse Outcome Pathway (AOP) relevant for the identification of substances having endocrine disruptor properties. EFSA indicated that AOPs for both EATS and non-EATS endocrine adverse outcomes can provide a practical answer for the two main scientific issues and can represent a regulatory valid approach for the scientific implementation of the ECHA/EFSA guidance for the identification of endocrine disruptors. A working group of the PPR Panel will draft a Scientific Opinion in which 4 AOPs dealing with the uterine adverse outcomes will be developed.

2.13. Top dose selection

Background

The top dose used in certain chronic toxicity, carcinogenicity and reproductive toxicity studies is not always high enough to support all regulatory implementations for these data in Europe, particularly for classification and labelling. Proper guidance is still missing and current TGs offer too much flexibility. A webinar series has been organised by the OECD in order to facilitate greater understanding amongst member countries of the regulatory frameworks in various jurisdictions where chronic toxicity data are used and how.

Meeting's discussion and conclusion

The experts discussed the need to set the highest dose level on the basis of kinetic data including metabolic saturation (i.e. at the dose showing that the plateau of saturation has been reached). The majority of the experts agreed that without an understanding of the saturation, toxicity at the highest dose tested should be observed. One expert commented that plateau of saturation can be consequent

to a formulation which is not absorbed that is not representative for the active substance as such and that this is one aspect that needs to be considered when saturation is observed. It was also commented that the evidence of a plateau often is lacking, and that only a flexion of the kinetic curve is considered as enough evidence that the plateau has been reached. This is considered not acceptable. Most of the experts agreed that the current proposal of using kinetic data for the high dose selection is only justifiable for hazard characterisation and risk assessment, but not for hazard identification.

2.14. AOB

- One expert pointed out a potentially emerging issue related to the exposure of children to common metabolites of pyrethroids or to alkyl phosphates (common metabolites of organophosphate substances). Recent studies from a Danish research group indicated that the presence of alkyl phosphates in the urine of children would be associated with higher risk of neurological impairment in children. However, the toxicity of alkyl phosphates, which can be present also in the food, is still unclear. The cumulative effects of alkyl phosphates are unknown and EFSA indicated that the report of cumulative risk assessment is currently under public consultation. Meaningful data on the DARs/RARs are not available for these alkyl phosphates and, considering no reference values have been set, a risk assessment for such substances is quite difficult to be performed. In addition, considering these metabolites are at very low levels, no toxicological data have been required in the past. EFSA proposed to highlight this issue to the European Commission. Another MS expert indicated that RIVM is also working on these metabolites, but they do not have toxicological data. Immediately after this experts' meeting EFSA checked whether information on metabolites common to organophosphate substances might be retrieved from MRL applications, but unfortunately no toxicological data were available.
- One expert expressed the need to have a clear guidance on the use of historical control data (HCD) to ensure consistency in the peer review. In many cases, misuse of HCD has been observed and different acceptability criteria are present. Therefore, current information from different regulatory frameworks should be considered. There was wide support from other MSs. The experts proposed e.g. a workshop to be organised. The outcome of the workshop could be used for further development of criteria for usage of HCD in a more structured manner. This would ensure consistency in the DAR/RAR and would allow to have a better understanding of the cases where the HCD are needed and acceptable or not, as well as how to use those data. One expert pointed out that HCD may also be discussed in the context of defining the CES for benchmark dose modelling.

3. Overall EFSA Conclusions and Recommendations

A majority view was reached in the discussed points. On the basis of current legal data requirements, the following recommendations reflect EFSA's view shared by the majority of the experts. EFSA will take care of appropriate communication and implementation of the proposed measures.

- Assessment of substances ED potential: different actions and recommendations were agreed and listed under 2.1.
- Possible application of the ECHA Guidance for the assessment of impurities also for pesticides: a consistent approach between ECHA and EFSA would be needed, but several issues requiring further discussions have been identified (see also 2.2).
- Groundwater metabolites: the current guidance document on the assessment of the relevance of metabolites in groundwater is out-of-date and the need for an update should be brought to the attention of the European Commission. A data gap for *in vitro* micronucleus test to properly cover aneugenicity should be set when such data has not been provided for groundwater metabolites assessed during the evaluation of a specific active substance (see also 2.4).
- Genotoxicity of mixtures: when dealing with mixtures of active substances, it is encouraged to use the statement of the Scientific Committee (see also 2.5).
- Use of *in silico* methods for the assessment of genotoxicity: the EFSA outsourced project for prediction of genotoxicity showed good performance in the case of prediction of point mutations only. Therefore, EFSA would recommend to perform a follow-up project to further explore the applicability of *in silico* tools to non-Ames endpoints and possibly also to other toxicological endpoints.
- Toxicological assessment of metabolites found as residues: the residues section is encouraged to provide a priori a list of candidate metabolites that are considered relevant for the assessment of the toxicological profile. The EFSA template proposed for collecting information on the experimental data, QSAR analysis, read-across and overall conclusions for the analysis of metabolites have to be included in the DAR/RAR, pending some modifications (see also 2.7).
- Use of BMD instead of NOAEL: to be used, if supported by data, in case no NOAEL is set for a specific parameter and a LOAEL is set instead. A need for training was identified by MSs. MSs indicated that EFSA should organise a dedicated workshop on this with the help of volunteering MSs (see also 2.8).
- AOB: a potential emerging issue was indicated to be the exposure of children to alkyl phosphates since the presence of alkyl phosphates in urine has been associated with higher risk of neurological impairment in children. Considering that toxicological data are lacking for these metabolites, it is quite difficult to perform the risk assessment. This issue should be brought to the attention of the European Commission. The experts also agreed upon the need to have a clear guidance on the use of historical control data to ensure that such data is used in a consistent and structured way in the peer review processes. A workshop can be organised to start the preparation of such guidance.

References

Benigni R, Battistelli CL, Bossa C, Giuliani A, Fioravanzo E, Bassan A, Fuart Gatnik M, Rathman J, Yang C, 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; EN-1598. https://www.efsa.europa.eu/en/supporting/pub/en-1598

Corvi R, Albertini S, Hartung T, Hoffmann S, Maurici D, Pfuhler S, van Benthem J, Vanparys P, 2008. ECVAM retrospective validation of *in vit*ro micronucleus test (MNT). Mutagenesis. 23(4):271-83.

ECHA (European Chemicals Agency), 2008. Guidance on information requirements and chemical safety assessment. Chapter R.6: QSARs and grouping of chemicals. May 2008.

ECHA (European Chemicals Agency), 2018. Guidance on the Biocidal Products Regulation - Volume V, Guidance on applications for technical equivalence - Version 2.0 July 2018. https://echa.europa.eu/documents/10162/23036412/guidance_applications_technical_equivalence_en.pdf/18f72d37-98b6-47c8-98bb-941afeff6968

ECHA (European Chemicals Agency) and EFSA (European Food Safety Authority) with the technical support of the Joint Research Centre (JRC), Andersson N, Arena M, Auteri D, Barmaz S, Grignard E, Kienzler A, Lepper P, Lostia AM, Munn S, Parra Morte JM, Pellizzato F, Tarazona J, Terron A and Van der Linden S, 2018. Guidance for the identification of endocrine disruptors in the context of Regulations (EU) No 528/2012 and (EC) No 1107/2009. EFSA Journal 2018;16(6):5311,135 pp. https://doi.org/10.2903/j.efsa.2018.5311. ECHA-18-G-01-EN.

EFSA (European Food Safety Authority), 2017a. Guidance on dermal absorption. EFSA Journal 2017;15(6):4873. https://www.efsa.europa.eu/en/efsajournal/pub/4873

EFSA (European Food Safety Authority), 2017b. Update of the guidance on the benchmark dose approach in risk assessment. EFSA Journal 2017;15(1):4658. https://www.efsa.europa.eu/it/efsajournal/pub/4658

EFSA (European Food Safety Authority), 2019a. Guidance of EFSA on risk assessments for active substances of plant protection products that have stereoisomers as components or impurities and or transformation products of active substances that may have stereoisomers. EFSA Journal 2019;17(8):5804. https://www.efsa.europa.eu/en/efsajournal/pub/5804

EFSA (European Food Safety Authority), 2019b. EFSA Report on in vitro comparative metabolism studies in regulatory pesticide risk assessment. EFSA Supporting publication 2019;EN-1618.

EFSA European Food Safety Authority, 2019c. Administrative guidance on submission of dossiers and assessment reports for the peer-review of pesticide active substances, EFSA supporting publication 2019:EN-1612. 49pp. doi:10.2903/sp.efsa.2019.EN-1612

EFSA PPR Panel (European Food Safety Authority 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. doi:10.2903/j.efsa.2016.4549

EFSA Scientific Committee, 2011. Scientific opinion on genotoxicity testing strategies applicable to food and feed safety assessment. EFSA Journal 2011;9(9):2379. https://www.efsa.europa.eu/en/efsajournal/pub/2379

EFSA Scientific Committee, 2019. Genotoxicity assessment of chemical mixtures. EFSA Journal 2019;17(1):5519. https://www.efsa.europa.eu/it/efsajournal/pub/5519

European Commission, 2003. Guidance Document on the assessment of the relevance of metabolites in groundwater of substances regulated under Council Directive 91/4147EEC. Sanco/221/200-rev.10-final, 25 February 2003.

European Commission, 2012. Guidance Document on the assessment of the equivalence of technical materials of substances regulated under Regulation (EC) No 1107/2009, SANCO/10597/2003 –rev. 10.1, 2012.

European Commission, 2014. Guidance document on botanical active substances used in plant protection products. SANCO/11470/2012-rev.8, 20 March 2014.

OECD (Organsation for Economic Co-operation and Development), 2011. Guidance notes for the Estimation of Dermal Absorption Values, OECD Series on Testing and Assessment, No. 156. OECD, Paris.

Glossary and Abbreviations

ADME	administration, distribution, metabolism and excretion
ADI	acceptable daily intake
AOP	adverse outcome pathway
BMD	benchmark dose
BMDL	benchmark dose lower bound
BMDU	benchmark dose upper bound
BMR	benchmark response
CAR	constitutive androstane receptor
CBA	component based approach
CES	critical effects size
DAR	Draft Assessment Report
DNT	Developmental neurotoxicity
EATS	estrogen, androgen, thyroid, steroidogenic
EC	European Commission
ECHA	European Chemicals Agency
ED	Endocrine disruptor
EFSA	European Food Safety Authority
EMA	European Medical Agency
GN	Guidance notes
HCD	Historical control data
IATA	Integrated Approaches to Testing and Assessment
JMPR	Joint FAO/WHO Meeting on Pesticide Residues
JMPS	Joint Meeting on Pesticide Specifications
LC-MS/MS	liquid chromatography with tandem mass spectrometry
LOAEL	Low observed adverse effect level
MoA	Mode of action
NOAEL	No observed effect level
MS	Member State
OECD	Organisation for Economic Co-operation and Development
PAFF	Standing Committee on Plants, Animals, Food and Feed
PPP	Plant Protection Product
PPR Panel	Panel of Plant Protection Products and their Residues
PREV	Pesticides Peer Review
PXR	pregnane X receptor
QSAR	Quantitative structure activity relationship

RAR	Renewal Assessment Report
RMS	Rapporteur Member State
SA	Structural alert
SC	Scientific Committee
SPSF	Standard project submission form
SVM	Support Vector Machines
T3	Triiodothyronine
T4	Thyroxine
TG	test guideline
TH	thyroid hormone
TSH	thyroid stimulating hormone
TTC	threshold of toxicological concern
UF	Uncertainty factor
WMA	whole mixture approach
WNT	Working Group of the National Coordinators for the Test Guidelines Programme
WoE	weight of evidence

Appendix A – Template of a summary assessment report of a QSAR

B.6.8.1 Studies on metabolites

Metabolite M1

A- Administrative data:

Report Author and date	Risk assessor, year.
Title	QSAR analysis on metabolite M1
Report No	Reference to the dossier where the full report is available
Guideline	ECHA REACH Guideline on QSAR R.6 (ECHA, 2008)
GLP	No

B- Material and Methods:

1. Substance identity:
    a. Chemical name:.
    b. Smiles: .
    c. Code: M1.
    d. Structure:

2. Data model source: reference and QMRF protocol: *This information is applicable to all metabolites assessed by the model for a specific endpoint.*

3. Information on the QSAR Model. *This information is applicable to all metabolites assessed by the model for a specific endpoint.*
    a. Prediction endpoint
    b. Algorithm of the model
    c. Statistics:
        i. Internal.
        ii. External against the EFSA Genotoxicity Pesticides Database
    d. Domains: general description, coverage, describe if mechanistic or metabolic interpretation were given.
    e. Analogues

C- **Results and Discussion:** *This is up to the risk assessor, this is specific for each metabolite and endpoint.*

1. Prediction value:
2. Applicability domain, analogues and uncertainties:

D- **Conclusions:** *This is up to the risk assessor, this is specific for each metabolite and endpoint.*

Conclusions including adequacy of the prediction.

Appendix B – Example of summary assessment report of a QSAR

B.6.8.1 Studies on metabolites

Metabolite M1

Example 1

A- Administrative data:

Report Author and date	Risk assessor, 2019.
Title	QSAR analysis on M1
Report No	Muta Caesar Report
Guideline	ECHA REACH Guideline on QSAR R.6 (ECHA, 2008)
GLP	No

B- Material and Methods:

1. Substance identity:
    a. Chemical name:
    b. Smiles:
    c. Code: M1
    d. Structure:

2. Data model source: CAESAR Mutagenicity Model v 2.1.12 - implemented in the VEGA software (v 1.0.8). QMRF protocol: Q15-410-0008.
3. Information on the QSAR Model.
    a. Prediction endpoint: Ames mutagenicity assay (gene mutation).
    b. Algorithm of the model: mutagenicity classifier integrating two different techniques: a machine learning algorithm from the Support Vector Machines (SVM) collection, then an ad-hoc expert system based on known structural alerts (SAs) (Benigni-Bossa rule base), tailored to refine its predictions.
    c. Statistics:
        i. Internal (accuracy around 92% for the training set and 82% for the test set (Ferrari, et al. 2010).
        ii. External against the EFSA Genotoxicity Pesticides Database (around 60% sensitivity, around 85% specificity, see Figure 21 in Benigni et al., 2019).
    d. Domains: The model was built using dataset of 4337 molecular structures with corresponding Ames test data (2401 mutagens and 1936 non-mutagens) The model is applicable to heterogenous chemicals and defined an applicability domain index. The applicability domain index is a multicomponent definition including structural similarity to the training set compounds, accuracy and concordance of the predictions for 6 most similar substances. Mechanistic (Benigni-Bossa rules) are included. Metabolic interpretation is not given.
    e. Analogues: The software provides six most similar substances from the training set with their experimental and predicted values and it is part of the applicability domain index.

C- Results and Discussion:

3. Prediction value: non-mutagenic, according to 3 scale classification scheme (non-mutagenic; mutagenic and suspect mutagenic). The risk assessor interpreted non-mutagenic as negative.
4. Applicability domain, analogues and uncertainties: the compound is not part of the training set. The results indicated that it could be out of the applicability domain. Regarding analogues, the reported 6 most similar substances from the training set have similarity in the range of 0. 84 to 0.91. However, the experimental values of some of them disagreed with the predicted value (accuracy of 67%).

D- Conclusions:

Metabolite M1 is predicted as non-mutagenic. The risk assessor considered the reliability of the prediction as medium given that it could be out of the applicability domain (accuracy of 67% for similar molecules) and the 60% sensitivity of the model compared to EFSA Genotoxicity Pesticides Database. The QSAR model outcome should be used in a weight of evidence approach.

Example 2

A- Administrative data:

Report Author and date	Risk assessor, 2019.
Title	QSAR analysis on M1
Report No	Derek_report
Guideline	ECHA REACH Guideline on QSAR R.6 (ECHA, 2008)
GLP	No

B- Material and Methods:

1. Substance identity:
    a. Chemical name:.
    b. Smiles: .
    c. Code: M1.
    d. Structure:

2. Data model source: Derek Nexus: 6.0.1, Nexus: 2.2.2. QMRF protocol: Q19-761-0004.
3. Information on the QSAR Model.
    a. Prediction endpoint: mutagenicity *in vitro* in bacterium (Ames test).
    b. Algorithm of the model: expert derived structural alerts for mutagenicity, physicochemical properties and associated reasoning. Following alert evaluation, Derek evaluates whether non-alerting query compounds contain any features that are either (i) also present in non-alerting mutagens in a large Ames test reference set (misclassified features) or (ii) not present in a large Ames test reference set (unclassified features).
    c. Statistics:
        i. Internal: not available given that it is a knowledge based.
        ii. External against the EFSA Genotoxicity Pesticides Database (around 67% sensitivity, around 89% specificity, see Figure 21 in Benigni et al., 2019).
    d. Domains: The compounds in the dataset are primarily small and medium-sized chemicals and so are representative of the structures used to build the model. The Ames test reference set contains 4757 mutagens and 5210 non-mutagens (v6.0) The scopes of the structure-activity relationships describing the mutagenicity endpoint are defined by the developer to be the applicability domain for the model. Therefore, if a chemical activates an alert describing a structure-activity for mutagenicity it can be considered to be within the applicability domain. If a compound does not activate an alert or reasoning rule then Derek makes a negative prediction. The applicability of the negative prediction to the query compounds can be determined by an expert, if required, by investigating the presence (or absence) of misclassified and/or unclassified features. Misclassified features in the molecule are found in non-alerting mutagens in the Lhasa reference set. The prediction remains negative and the misclassified features are highlighted to enable the negative prediction to be verified by expert assessment. Inactive, contains unclassified features, some features in the molecule have not been found in the Lhasa reference set. The prediction remains negative and the unclassified features are highlighted to enable the negative prediction to be verified by expert assessment. Mechanistic information is detailed in the comments associated with an alert and can include information on both the mechanism of action and biological target.
    e. Analogues: Non-proprietary elements of the training set are available through the references, and illustrated by the examples, within Derek Nexus.

C- Results and Discussion:

1. Prediction value: inactive-no misclassified or unclassified features, according to 13 scale classification scheme (certain, probable, plausible, equivocal, doubted, improbable, impossible, open, contradicted, inactive-no misclassified or unclassified feature, inactive-contains misclassified features, inactive-contains unclassified features, inactive-contains misclassified and unclassified features). The risk assessor interpreted inactive as negative.
2. Applicability domain, analogues and uncertainties: The software indicated that the query structure does not contain misclassified or unclassified features. The risk assessor interpreted that the prediction is in the domain of the Lhasa dataset given that the metabolite does not contain unclassified features, meaning that the it does not contain any feature that has not been considered in the large Lhasa reference set

D- Conclusions:

Metabolite M1 is predicted as inactive-no misclassified or unclassified features. Despite it is a knowledge-based system, the risk assessor considered that the prediction is within the domain of the Lhasa reference set. The risk assessor considered the reliability of the prediction as high given also the 67% sensitivity of the model compared to EFSA Genotoxicity Pesticides Database. The QSAR model outcome should be used in a weight of evidence approach.

Appendix C – Summary table for integrating experimental evidence on genotoxicity for metabolites

Name/ Code	Endpoint	Experimental Data (reliability)	QSAR 1 (reliability)	QSAR 2 (reliability)	Read-across (including ADME) (reliability)	Overall conclusion by endpoint: Gene mutation, clastogenicity/aneugenicity	Overall conclusion on genotoxicity
M1	Gene mutation- in bacteria.						
	Gene mutation in mammalian cells						
	Clastogenicity in vitro						
	Aneugenicity in vitro						
	Clastogenicity/Aneugenicity in vivo						

Appendix D – Summary table on the assessment of the toxicological profile of metabolites

Name, code, and smiles	**Structure**	**Lead compound within a group**	**Origin (groundwater, crop, livestock, etc)**	**Percentage of the applied/absorbed dose in excreta, tissues and downstream metabolic pathway**	**Genotoxicity conclusion and basis (endpoints)**. Experimental Data/QSAR/Grouping and Read-across/Covered by parent	**General toxicity conclusion and basis (endpoints)** Experimental Data/QSAR/Grouping and Read-across/Covered by parent. Reference values (UF and basis)