
SCIENTIFIC REPORT submitted to EFSA

Applicability of QSAR analysis in the evaluation of developmental and neurotoxicity effects for the assessment of the toxicological relevance of metabolites and degradates of pesticide active substances for dietary risk assessment¹

Prepared by

Andrew Worth, Mojca Fuart-Gatnik, Silvia Lapenna and Rositsa Serafimova

**Systems Toxicology Unit, Institute for Health & Consumer Protection,
European Commission - Joint Research Centre, Ispra, Italy**

Final version of 31/05/11

¹ Question No: EFSA-Q-2010-01258. Accepted for publication on 10 June 2011.

Abstract

This is the final report of a project carried out by the European Commission's Joint Research and sponsored by the European Food Safety Authority with the overall aim of evaluating the potential applicability of computational methods for predicting adverse developmental and neurotoxicity effects in the dietary risk assessment of pesticides.

While the toxicological profile of the parent active substance is fully characterised through the experimental studies required by EU legislation, only very limited toxicological data are usually available for their metabolites and degradates. For reasons of efficiency and animal welfare, computational methods based on structure-activity analysis and read-across are being investigated for their applicability in assessing the toxicological relevance of metabolites and degradates of pesticide active substances. The ability to reliably predict the presence and absence of short-term effects of concern, and in particular developmental toxicity and neurotoxicity, would have a positive impact on the way pesticide risk assessments are currently carried out by reducing the need for toxicity testing on metabolites and degradates as well as the need to conduct short-term exposure assessments.

In this study, the ability of selected Quantitative Structure-Activity Relationship (QSAR) tools to predict developmental and neurotoxicity was analysed, and a stepwise approach based on the use of QSAR analysis and read-across was proposed as possible way of supporting, alongside other non-testing approaches such as the Threshold of Toxicological Concern (TTC) approach, the assessment of pesticide metabolites and degradates in terms of their toxicological relevance. In this stepwise approach, QSAR tools are used in a preliminary step to identify toxic chemicals, while read-across is applied, in cases where a chemical is predicted by QSAR to be non-toxic, as a means of distinguishing between true and false negatives. This approach is shown to improve the overall ability to distinguish between toxic and non-toxic chemicals compared with the use of individual tools.

Summary

This is the final report of a project carried out during 2011 by the European Commission's Joint Research (JRC) and sponsored by the European Food Safety Authority (EFSA) with the overall aim of evaluating the potential applicability of computational methods for predicting adverse developmental and neurotoxicity effects in the dietary risk assessment of pesticides. The work was carried out under the terms of a Service Level Agreement between EFSA and the JRC.

Whereas the toxicological profile of the parent active substance is fully characterised through the experimental studies required by EU legislation, very limited toxicological data are usually available for their metabolites and degradates. For reasons of efficiency and animal welfare, computational methods based on structure-activity analysis and read-across are being investigated for their applicability in assessing the toxicological relevance of metabolites and degradates of pesticide active substances. The ability to reliably predict the presence and absence of short-term effects of concern, and in particular developmental toxicity and neurotoxicity, would have a positive impact on the way pesticide risk assessments are currently carried out by reducing the need for toxicity testing on metabolites and degradates as well as the need to conduct short-term exposure assessments.

In this study, the ability of selected Quantitative Structure-Activity Relationship (QSAR) tools to predict developmental and neurotoxicity was analysed, and a stepwise approach based on the use of QSAR analysis and read-across was proposed as possible way of supporting, alongside other non-testing approaches such as the Threshold of Toxicological Concern (TTC) approach, the assessment of pesticide metabolites and degradates in terms of their toxicological relevance. In this stepwise approach, QSAR tools are used in a preliminary step to identify toxic chemicals, while read-across is applied, in cases where a chemical is predicted by QSAR to be non-toxic, as a means of distinguishing between true and false negatives. This approach is shown to improve the overall ability to distinguish between toxic and non-toxic chemicals compared with the use of individual tools. Provided that the general approach is considered acceptable, the short-term prospects for applying it will depend on acceptance criteria that will need to be established by EFSA.

AVAILABILITY OF QSAR TOOLS AND DATABASES

A limited range of software tools and databases were identified as potentially useful for developmental toxicity and neurotoxicity prediction (Sections 2 and 5). QSAR tools for predicting developmental toxicity include CAESAR, Derek, HazardExpert, Leadscope, PASS and TOPKAT. Tools for predicting neurotoxicity include Derek, HazardExpert and PASS. A few models and tools for predicting ADME characteristics (placental barrier transfer and blood-brain barrier passage) were also investigated.

In the case of developmental toxicity, the US EPA's ToxRef Database is a potentially useful reference database for the development of new models and the application of grouping and read-across. In the case of neurotoxicity, there is a lack of freely available QSAR tools, as well as a suitable public reference database for the development of new models and the application of grouping and read-across. A general caveat, which needs to be considered irrespective of the endpoint being predicted and the QSAR or reference database used, is that different regulatory bodies may apply different criteria in the evaluation of raw data. In order to take such differences

into account, it is important that the conclusions (positive or negative toxicity) are accompanied by a description of the underlying effects at the organ, tissue and cellular levels. The ideal situation to meet EFSA's needs in predictive toxicology would be to develop an in-house database of relevant evaluated data, in which the conclusions are based on guideline criteria, and are linked to the underlying findings in the original study reports.

When compiling and searching chemical databases, the stereochemistry of molecules can be encoded into their SMILES strings. This could, in principle, be important in the prediction of developmental toxicity and neurotoxicity. Such information can in principle be encoded into QSAR models and structural alert-based rulebases. For example, the Derek knowledgebase includes several alerts that are sensitive to stereochemistry.

PREDICTIVE PERFORMANCE OF SELECTED QSAR TOOLS FOR DEVELOPMENTAL TOXICITY

Based on the performances of selected QSAR tools for developmental toxicity (Section 6), a number of conclusions can be drawn.

A literature-based model for placental transfer shows no tendency to distinguish between developmental toxicants and non-toxicants. This is not unexpected since the passage of a chemical across the placental barrier is not a sufficient or even necessary condition for developmental toxicity.

To predict the absence of developmental toxicity, the PASS models for embryotoxicity and teratogenicity appear to be the best stand-alone models in terms of their negative predictivities (44-45%) when assessed against the EFSA Extended Test Set. The combined use of two models led to a marginal increase in negative predictivity to 48%. With negative predictivities less than 50%, none of the models investigated, and no two-model combination, is expected to be adequate for use.

Some QSAR tools, such as Derek, HazardExpert and PASS, might be useful for the identification of developmental toxicants (due to their high positive predictivities of 81-96% when assessed against the EFSA Extended Test Set). In particular, such models could be useful in the context of a stepwise assessment strategy in which the use of QSAR to identify positives is followed by the use of read-across to identify negatives.

When evaluating the performances of QSAR models, care should be taken in the choice of test set, since different criteria for discriminating between positives and negatives may be used by different regulatory bodies or database providers. For example, when the developmental toxicity models were assessed against the US EPA's ToxRefDB dataset, their performances were strongly dependent on how the Lowest Effect Levels (LELs) for developmental effects were compared with the LELs for maternal effects.

PREDICTIVE PERFORMANCE OF SELECTED QSAR TOOLS FOR NEUROTOXICITY

On the basis of the performances of selected QSAR tools for neurotoxicity (Section 6), it is concluded that to predict the absence of neurotoxic potential, no individual model, and no two-model combination, appears adequate for use (since the negative predictivities are less than 50%).

Conversely, the statistics indicate that some tools, such as Derek and HazardExpert, might be useful for the identification of neurotoxicants (due to their high positive predictivities between

90-100%). In particular, such software tools might be useful in the context of a stepwise assessment strategy in which the use of QSAR to identify positives is followed by the use of read-across to identify negatives. This possibility could not be explored in this study, due to the lack of a suitable reference database for the read-across exercise. To investigate the applicability of this assessment strategy, it will be necessary to develop such a database.

In view of the lack of available tools for predicting the absence of neurotoxic potential, the most pragmatic consideration in the risk assessment of pesticide metabolites/degradates is to apply the hypothesis that non-neurotoxic parent substances do not generate neurotoxic (bio)transformation products. Based on its experience of evaluating pesticide dossiers, EFSA could not identify any evidence that refutes this hypothesis. In addition, the only evidence we could find of non-neurotoxic parents (but not of pesticides) giving rise to products with neurotoxic effects were a few papers describing *in vitro* / mechanistic findings, which are not necessarily relevant to the *in vivo* effects of pesticides.

CHEMICAL SPACE ANALYSIS

Chemical space analysis can be used to explore and define the applicability domains of statistically-based models if their training sets (including structures and biological data) are available. It can also be used to inform model development by identifying areas of chemistry that are not adequately covered in existing models.

There are different approaches to building applicability domains, but two of the most commonly used approaches are based on structural fragments and molecular descriptors. Model applicability domains can be used to rationalise the predictions made for test set chemicals (for which the toxicological effects are known) and to help determine the reliability of prediction for untested chemicals. However, the interpretation is not straightforward. If a chemical is outside the applicability domain of a model, it does not necessarily mean that its predicted toxicity is wrong, but simply that the prediction cannot be made with as much confidence. Conversely, when a chemical is within the applicability domain, it does not necessarily follow that the predicted toxicity will be accurate, but simply that the prediction can be made with a defined level of confidence. Furthermore, there is no absolute definition of a model applicability domain – different interpretations may be useful for different purposes. Some software tools provide their own assessment of prediction reliability based on applicability domain considerations, whereas other software tools do not. In practice, the definition and interpretation of applicability domains is not a trivial exercise.

In this study, only the CAESAR and Leadscape models were amenable to chemical space analysis (Section 7). However, the results were not particularly informative in terms of understanding the reliability of prediction, which reinforces the view that such analyses should be regarded as indicative rather than conclusive.

In addition to exploring the applicability domains of QSAR models, chemical space analysis can be used to compare the test sets with the “universe” of pesticides, as represented by the Plant Protection Products (PPP) inventory. The developmental toxicity test sets used in this study were found to largely cover the space of the PPP inventory, while the (smaller) neurotoxicity test set was less diffuse and chemically diverse. In future efforts to build a more extensive neurotoxicity test set, as a means of providing a more comprehensive chemical challenge to available models,

it would be useful to search for chemicals in these areas of the PPP. However, there is no guarantee that reference chemicals with adequate data will be found.

PREDICTIVE ABILITY OF READ-ACROSS

Compared with QSAR models, the predictive performance of the grouping and read-across approach cannot be generalised as easily, since this is an *ad hoc* approach in which a number of subjective choices are made. For example, decisions have to be made concerning the choice of reference database(s) for analogue searching, the similarity criteria used to identify close analogues, and the assessment of the relevance and reliability of the analogue data, and the interpretation of positive, negative and inconclusive outcomes. Nevertheless, the read-across approach is equally well-suited to the identification of positive and negative chemicals, and for this reason, it is proposed in this study as a means of clarifying the negative predictions resulting from the application of QSAR.

Developmental toxicity and neurotoxicity are complex endpoints, which are only partially understood in mechanistic terms. In the absence of endpoint-specific profilers for analogue identification, analogues can be identified based on the presence of organic functional groups.

The usefulness of read-across in the assessment of developmental toxicity is illustrated in Section 8.

PRACTICAL USE OF EVALUATED MODELS

An evaluation of the practical usefulness of QSAR and read-across tools, for the purposes of pesticide metabolite assessment, is not a straightforward task. Such an evaluation needs to take into account the availability and cost of the software, the expertise required to use the software, as well as the validation characteristics and the regulatory context in which the models are being used.

When establishing the pesticide residue definition for risk assessment, any decision on which models / software tools are fit-for-purpose should be taken by EFSA, ideally on the basis of a transparent set of acceptance criteria (which could be developed, for example, by the PPR Panel). In particular, EFSA needs to decide on the acceptable false positives and false negatives in the use of the models, taking into account that models with different strengths and weaknesses can be combined in a stepwise strategy that optimises the overall predictive performance. The main considerations that would help in setting these criteria are provided in Section 15.

QSAR models are generally designed to predict toxicity by identifying structural features associated with the molecular interactions that lead to toxicological outcomes. Structural features are rarely associated with the absence of toxicity, unless they are structural groups that have a mitigating effect on the properties of another group (e.g. steric hindrance, or alteration of chemical reactivity via electronic polarisation). To some extent, QSARs may also capture the absence of toxicity to the extent that they implicitly encode ADME characteristics, such as limitations in bioavailability due to molecular size or hydrophobicity. In contrast, the read-across approach is equally suited to the identification of toxicants and non-toxicants, provided that a sufficient number of analogues can be found with adequate experimental data. In this respect, in the absence of a more specific mechanistic understanding of toxicity, analogue searching by organic functional groups is particularly useful, since analogues that contain different (and potentially reactive) functional groups to the chemical of interest can be excluded.

The software models evaluated were found to have good abilities to identify positive chemicals (positive predictivities greater than 80%) but poor abilities to identify negative chemicals (negative predictivities less than 50%). The strengths of these QSAR models can be exploited in a stepwise strategy in which QSARs are used in a preliminary step only for the identification of positive chemicals (in other words, the positive predictions are trusted, but no confidence is attached to the negative predictions), whereas a subsequent step based on grouping and read-across is used to discriminate between the true and false negatives generated by QSAR. In Section 9, this is shown to be an effective strategy for the prediction of developmental toxicity. The concept could not be tested for the prediction of neurotoxicity due to a lack of a suitable reference database for read-across. Nevertheless, one would expect a similar stepwise approach to neurotoxicity prediction to be more effective than the use of QSAR models alone.

The stepwise non-testing approach could be used, along with computational methods for other toxicological endpoints (e.g. genotoxicity) and the TTC approach, into a decision tree for evaluating the toxicological relevance of metabolites and degradates.

RECOMMENDATIONS FOR FUTURE RESEARCH

Based on the findings of this study, a number of recommendations are made with a view to improving the use of computational methods to identify the presence and absence of short-term effects of high concern, such as neurotoxicity and developmental toxicity. These recommendations, given in Sections 10-16, focus on prospects for the short-term (<2 years), mid-term (2-5 years), and long-term (>5 years).

Key words: acute reference dose, alternative method, degrade, dietary risk assessment, exposure assessment, in silico, metabolite, pesticide active substance, QSAR, read-across, structural alert, toxicity, TTC

Table of Contents

Abstract	2
Summary	3
Table of Contents	8
Background	10
Terms of reference	10
Acknowledgements	10
Introduction and Objectives	11
Materials and Methods	14
1. Datasets and treatment of chemical structures	14
1.1. Developmental toxicity datasets	14
1.2. Neurotoxicity dataset	15
2. QSAR analysis	20
2.1. Introduction to QSAR analysis	20
2.2. Statistical measures of the performance of classification models	21
2.3. QSAR models for ADME prediction	22
2.3.1. Literature model for placental barrier passage	22
2.3.2. ADMET Predictor	23
2.3.3. Accelrys ADME (add-in for Excel)	23
2.4. QSAR models for toxicity prediction	23
2.4.1. CAESAR	24
2.4.2. Derek	24
2.4.3. HazardExpert	25
2.4.4. Leadscape Model Applier	26
2.4.5. PASS	26
2.4.6. TOPKAT	27
3. Grouping and read-across	27
3.1. Introduction to grouping and read-across	27
3.2. OECD QSAR Toolbox	28
4. Chemical space analysis	29
4.1. Structural space analysis	30
4.2. Molecular descriptor space analysis	30
Results	31
5. Available QSAR tools and databases	31
5.1. Software tools for developmental toxicity prediction	31
5.2. Databases on developmental toxicity	32
5.3. Software tools for neurotoxicity prediction	34
5.4. Databases on neurotoxicity	35
5.5. Conclusions on the availability of predictive tools and databases	36
6. Performance of selected QSAR tools	36
6.1. Assessment of a literature-based QSAR for placental transfer	36
6.2. Assessment of models for developmental toxicity prediction	36
6.2.1. Predictive performance of developmental toxicity models (used alone)	36
6.2.2. Predictive performance of batteries of developmental toxicity models	37
6.2.3. Assessment of the developmental toxicity models against the ToxRefDB dataset	37
6.3. Assessment of models for neurotoxicity prediction	37
6.3.1. Predictive performance of neurotoxicity models (used alone)	37
6.3.2. Predictive performance of batteries of neurotoxicity models	38
6.4. Conclusions on the performance of QSAR models and tools	38
6.4.1. Placental transfer and developmental toxicity	38

6.4.2. Neurotoxicity	39
7. Chemical space analysis	56
7.1. Structural space analysis	56
7.1.1. Comparison of the PPP inventory with the developmental test sets	56
7.1.2. Comparison of the EFSA developmental test set with the model training sets	56
7.1.3. Comparison of the PPP inventory with the neurotoxicity test set	57
7.1.4. Comparison of the EFSA neurotoxicity test set with the model training sets	57
7.2. Molecular descriptor space analysis	57
7.2.1. Comparison of the PPP inventory with the developmental test sets	57
7.2.2. Comparison of the EFSA developmental test set with the model training sets	57
7.2.3. Comparison of the PPP inventory with the neurotoxicity test set	57
7.2.4. Comparison of the EFSA neurotoxicity test set with the model training sets	57
7.3. Conclusions on the chemical space analysis	58
8. Case studies on the read-across assessment of developmental toxicity	76
8.1. Read-across prediction of positive (development toxicant)	77
8.2. Read-across prediction of negative (non-development toxicant)	78
8.3. Read-across inconclusive due to contradictory evidence	79
8.4. Conclusions on grouping and read-across	80
9. Use of QSAR and read-across in the assessment of toxicological relevance	81
9.1. Stepwise assessment scheme based on combined use of QSAR and read-across	81
9.2. Conclusions on the use of QSAR and read-across	82
Conclusions and Recommendations	88
10. Availability of QSAR tools and databases	88
11. Predictive performance of selected QSAR tools for developmental toxicity	88
12. Predictive performance of selected QSAR tools for neurotoxicity	89
13. Chemical space analysis	89
14. Predictive ability of read-across	90
15. Practical use of evaluated models	91
16. Future developments in predictive toxicology	92
References	97
Appendices	101
Glossary / Abbreviations	102

Background

See Introduction

Terms of reference

See Objectives

Acknowledgements

This contract/grant was awarded by EFSA to: European Commission Joint Research Centre, Institute for Health & Consumer Protection, Ispra, Italy

Contract/grant title: Applicability of QSAR analysis in the evaluation of developmental and neurotoxicity effects for the assessment of the toxicological relevance of metabolites and degradates of pesticide active substances for dietary risk assessment

Contract/grant number: SLA/EFSA-JRC/PPR/2010/01

The authors very much appreciate the contributions and advice provided by EFSA staff (Luc Mohimont, Istvan Sebestyen, Hans Steinkellner, Manuela Tiramani) and members of the EFSA Steering Group (Claudia Bolognesi, André Muller, Maria Tasheva, Christiane Vleminckx). In addition, the authors would like to acknowledge: Mark Hewitt (Liverpool John Moores University, UK) for providing QSAR predictions of placental barrier passage; Vladimir Poroikov (Institute of Biomedical Chemistry, Russian Academy of Medical Sciences) for providing a trial version of the PASS software; as well as Aldert Piersma and André Muller (RIVM, NL) for providing a dataset of chemicals classified for developmental toxicity. The authors are also grateful to JRC colleagues for advising on and reviewing this work, in particular Elisabet Berggren, Paolo Castello, Bibiana Scelfo and Maurice Whelan.

Introduction and Objectives

INTRODUCTION

In the European Union (EU), the provisions for the evaluation and authorisation of plant protection products (PPPs) are laid down by *Council Directive 91/414/EEC* (EC, 1991). For active substances in plant protection products a comprehensive risk assessment is required, including identification of metabolites and degradates which, after application of PPPs to crops, can be present as residues in food commodities. Metabolites and degradates of active substances may arise in plant but also livestock metabolism (after ingestion of treated plants as feed), through microbial activity in soil, chemical degradation processes and food and feed processing.

Within the frame of the evaluation of PPPs usually two definitions of residue are established (OECD, 2006). One for monitoring/enforcement of MRLs (Maximum Residue Levels), which has to meet analytical practicalities and therefore usually only including the active substance and one for risk assessment which should quantitatively and qualitatively represent the actual toxicological burden of residues in the food commodity, and therefore including also relevant metabolites and degradates of the active substance.

While the toxicological profile of the active substance is fully characterised through the studies required by *Directive 91/414/EEC* only very limited toxicological data are usually available for their metabolites and degradates. A full toxicological characterisation of relevant metabolites and degradates is not feasible as toxicological studies should be restricted to the extent possible to minimise use of test animals.

Therefore the European Food Safety Authority (EFSA) commissioned three projects to explore alternative (non-animal) methods for the toxicity assessment of metabolites and degradates. The final reports from these projects are available on the EFSA webpage. Explored were the suitability of the Threshold of Toxicological Concern (TTC) concept (CRD, 2010), the applicability of Quantitative Structure Activity Relationship (QSAR) analysis (JRC, 2010) and the impact of metabolism on toxicity (AGES, 2010). These reports form the basis for the work of EFSA's Panel on Plant Protection Products and their Residues (PPR Panel) on an opinion on the toxicological relevance of metabolites and degradates of pesticide active substances for dietary risk assessment.

Based on the promising outcome of the TTC project it became clear already at an early stage of the on-going work on the opinion that the TTC concept would probably (together with the application of certain QSARs and structural comparisons) be the key tool for the assessment strategy of the relevance of pesticide metabolites.

A TTC scheme including several thresholds has been validated for pesticides and is likely to be used as an initial screening tool in order to exempt a significant proportion of pesticide metabolites from detailed risk assessment when the levels of their occurrence in the consumer diet are lower than their respective TTC values.

However acute effects and short term exposure to peak concentrations of pesticide metabolites in the diet are an important issue for the application of the TTC scheme. Although the various TTC thresholds are based on chronic toxicological studies and therefore cover acute effects, the ratios

between short-term and chronic dietary exposures is often greater than the ratio between the acute and chronic toxicological reference values.

In order to increase the validity of the TTC scheme for demonstrating the non-relevance of pesticide metabolites, it should be combined with tools or assessment methodologies demonstrating that they are not of concern in relation to acute toxic effects or toxic effects seen early in short term tests. Within the frame of the peer review of PPPs, the application of an Acute Reference Dose (ARfD) is generally triggered by specific developmental (van Raaij et al, 2003) and, to a lesser extent, neurotoxic effects (Solecki et al, 2008). The final report from the QSAR project (JRC, 2010) identified several software tools that are designed for the identification of developmental and neurotoxic effects. Therefore EFSA concluded that computational tools should be investigated with a view to their use in identifying and/or excluding potential specific developmental and/or neurotoxic effects of pesticide metabolites. Accordingly, EFSA initiated this study with the JRC under the terms of a Service Level Agreement between EFSA and the European Commission's Joint Research Centre (JRC).

Along with the previous EFSA-funded projects (AGES, 2010; CRD, 2010; JRC, 2010), the results of the study described in this report will be used by EFSA to support the development of a scientific opinion which, once adopted, will be the basis for the work on the future guidance document on the establishment of the residue definition for risk assessment in food commodities which should provide scientifically-based criteria and practical advice for the definition of residue for risk assessors and regulators.

OBJECTIVES

The general purpose of this project was to perform an in-depth evaluation of the possible contribution of QSAR analysis for the identification of early-onset specific developmental and neurotoxic effects for the evaluation of the toxicological relevance of metabolites and degradates of active substances of pesticides for dietary risk assessment.

This project is intended as a contribution to the wider objective of reinforcing the robustness and consistency of the assessment of dietary risk resulting from the use of plant protection products, by incorporating to the best possible extent QSAR analysis for the assessment of early-onset specific developmental and neurotoxic effects.

The specific objectives (and terms of reference) of this project were to:

- 1) Identify software tools designed to predict developmental toxicity or neurotoxicity. This excludes models specifically designed for endocrine-receptor binding and metabolism.
- 2) Evaluate the extent to which selected tools can be used to identify the presence or absence of specific developmental toxicity effects and neurotoxicity, or specific effects underlying to these endpoints. This will require the establishment of a test set against which the models can be tested. The chemicals in the test set, and their relevant toxicological effects, will be identified by EFSA. The performance of the selected tools will be expressed in terms of positive/negative predictivity, sensitivity/specificity and concordance. The criteria for determining whether an effect is toxic or non-toxic will be determined by EFSA. For example, developmental toxicity will be related to PPP peer review evaluations or EU or GHS classification criteria.

- 3) Provide, if possible, a description of the applicability domains of the selected models and evaluate their relationship with the chemical space of pesticides. The latter will only be possible in the case of models for which the training sets are well-defined and fully transparent.
- 4) Comment on the practical usefulness of the evaluated tools, taking into account the availability and cost of the software, the expertise required to use the software, and the need for high sensitivity in the prediction of specific developmental and neurotoxic effects.
- 5) Comment on the insertion of such tools in an overall assessment scheme for metabolites and degradates bearing in mind that the TTC concept will be the central entity of such a scheme. The details of the scheme will be developed by EFSA.

Materials and Methods

1. Datasets and treatment of chemical structures

To ensure the suitability of chemicals included in the test sets used to challenge the various QSAR models, the chemicals were carefully selected by EFSA, taking into account the reliability and relevance of the associated toxicological data.

The molecular structure of each chemical was codified in terms of its simplified molecular input line entry specification (SMILES) string (Weininger, 1988) for subsequent processing by the QSAR tools. The SMILES notation allows the encoding of stereochemical features such as the configuration around double bonds and at chiral centres. Further information on SMILES can be found at: http://en.wikipedia.org/wiki/Simplified_molecular_input_line_entry_specification

The QSAR tools used in this study are described in Sections 2.5-2.11. None of the tools used based their predictions on stereochemical features, even though these were encoded in the SMILES strings.

1.1. Developmental toxicity datasets

Two test sets were compiled:

a) an **original test set** of 76 pesticides selected by EFSA (37 positives; 39 negatives), provided in Appendix A

The chemicals included in the original EFSA test set were identified from the reports of EFSA-funded studies (AGES; CRD, 2010), EFSA conclusions, Draft Assessment Reports (DARs) and evaluations of the Joint Meeting on Pesticide Residues (JMPR).

The following selection criteria were used to identify positive developmental toxicants:

- Malformations or other specific early onset developmental effects in rat or rabbit or in both species at maternally non-toxic doses
- Effects observed in both rat and rabbit are of particular concern
- Specific malformations, such as cleft palates, irrespective of maternal toxicity
- Substances classified with EU risk phrases R61 or R63
- Chemicals for which an ARfD was based on developmental effects

The list of negatives was based on the following selection criteria:

- No adverse effects seen in valid development tests with rat and rabbit at doses associated with maternal toxicity (most usual approach)
- Effects on offspring are “unspecific” (for instance slightly reduced birth weights) at doses with clear maternal toxicity

The distribution of chemicals across different pesticide classes is given in Table 1.1 and illustrated in Figure 1.1 The assignment of chemicals to pesticide classes was carried out (by EFSA) following the taxonomy given here: <http://www.pesticideinfo.org/>

b) an **extended test set** of 135 chemicals including the original 76 chemicals + 59 chemicals classified for developmental toxicity and provided by the RIVM. The RIVM dataset (given in Appendix B) is chemically diverse, including both pesticides and industrial chemicals. The additional chemicals were added to increase the ratio of negatives to positives, thereby improving the accuracy of the negative predictivity estimation (Steinberg et al, 2008).

In addition, a third dataset derived from the US EPA's ToxRefDB database (<http://actor.epa.gov/toxrefdb/faces/BasicInfo.jsp>) was used as a QSAR test set and as a reference dataset for read-across. This dataset had already been compiled by the US EPA and made publicly available in Excel format. The ToxRefDB developmental toxicity dataset contains data for 384 pesticide actives, but 18 were removed for QSAR analysis because they were inorganics, mixtures or structurally undefined. This resulted in a dataset containing 366 pesticides.

The *in vivo* data in the ToxRefDB were interpreted in three different ways:

- according to interpretation A, a “positive” was interpreted as any adverse effect observed in a developmental rat or rabbit study, without taking maternal toxicity into account. This dataset consists of 366 chemicals (246 positives; 120 negatives)
- according to interpretation B, a “positive” was interpreted as any adverse effect observed in a developmental rat or rabbit study (dLEL), provided that this occurred at a dose lower than that causing maternal toxicity (mLEL). If the lowest dose causing developmental toxicity (dLEL) was equal to the lowest dose causing maternal toxicity (mLEL), the overall conclusion was “undefined” (ND). In other words, the call was “positive” if $dLEL < mLEL$ and negative if $dLEL > mLEL$. This dataset consists of 317 chemicals (59 positives; 258 negatives)
- according to interpretation C, a “positive” was interpreted as any adverse effect observed in a developmental rat or rabbit study (dLEL), provided that this occurred at a dose lower than or equal to that causing maternal toxicity (mLEL). In other words, the call was “positive” if $dLEL \leq mLEL$ and negative if $dLEL > mLEL$. This dataset consists of 366 chemicals (193 positives; 173 negatives)

Although not identical, interpretation B was considered to most closely match the one used in establishing the EFSA dataset.

1.2. Neurotoxicity dataset

The test set of chemicals for neurotoxicity is given in Appendix C. The distribution of chemicals across different pesticide classes is given in Table 1.2 and illustrated in Figure 1.2. This includes the classes of carbamates (3 substances), neonicotinoids (3 substances), pyrethroids (13 substances) and triazoles (2 substances), which are expected to show neurotoxic effects based on their chemical class. Indeed, all of these substances are positives, except for asulam, which is both a carbamate and a sulphonamide.

Two of the substances are duplicated, since they are two-component mixtures in which each component has a different molecular weight (MW): avermectin (1a and 1b in Appendix C); and milbemectin (29a and 29b in Appendix C). There were 42 positives in total, including the duplicated substance. One of the negative substances, guazatine, was replicated in triplicate,

since it is an oligomer with a variable number of repeating units (0,1 and 2 repeats were considered). There were 23 negatives in total, including the replicated substance.

Table 1.1. Distribution of the EFSA developmental toxicity dataset across pesticide classes in original test set of 76 substances

Pesticide class	Number of substances
Aliphatic nitrogen fungicide	1
Amide	2
Anilide	3
Aromatic	2
Aryl oxy phenoxy propionate	1
Aryl phenyl ketone	1
Benzi imidazole	1
Benzoic acid	1
Benzonitrile	1
Benzoylphenylurea	3
Chloroacetanilide	1
Dicarboximide	1
Dicarboximide. Oxazole	1
Dinitroaniline herbicide	1
Dinitrophenol	1
Fermentation product S. Avermitilis	1
Growth inhibitor	1
Growth retardants	1
Imidazole	1
Imidazolinone herbicide	1
Morpholine	3
Nitrophenyl ether herbicide	1
N-phenyl phtalamides	1
Organophosphate	1
Organothiophosphate	1
Oxazole	1
Oxime carbamate insecticide	1
Phenylpyridinamine	1
Phosphic acid	1
Pyridine	1
Pyrimidine	1
Pyrimidinyol carbinol	1
Pyrimidinyloxybenzoic acid herbicide	1
Pyrimidinylsulfonylurea	3
Pyrrole	1
Strobilurin	1
Tetrazine	1
Tetronic acid	1
Thiazolidine acaricide	1
Thiocarbamate	1
Triazinone herbicide	1

Pesticide class	Number of substances
Triazole	15
Triazolinone	1
Unclassified acaricide	1
Unclassified insecticide	1
Uracil herbicide	1
Urea	2
Urea fungicide	1

Table 1.2. Distribution of the EFSA neurotoxicity positives across pesticide classes

Pesticide class	Number of substances
Acetaldehyde	1
Acetamide, chloroacetanilide	1
Amidine	1
Antibiotic, avermectin, milbemectin	5
Carbamate, organochlorine	1
Chlorinated nitroaniline	1
Dithiocarbamate	2
Ethylene generator	1
Morpholine	1
Neonicotinoid	3
Organochlorine	3
Oxadiazine	1
Oxyacetamide	1
Piperidine, quaternary ammonium	1
Pyrethroid	13
Quaternary ammonium	1
Tetronic and tetramic acid derivative	2
Triazine	1
Triazole, conazole	2

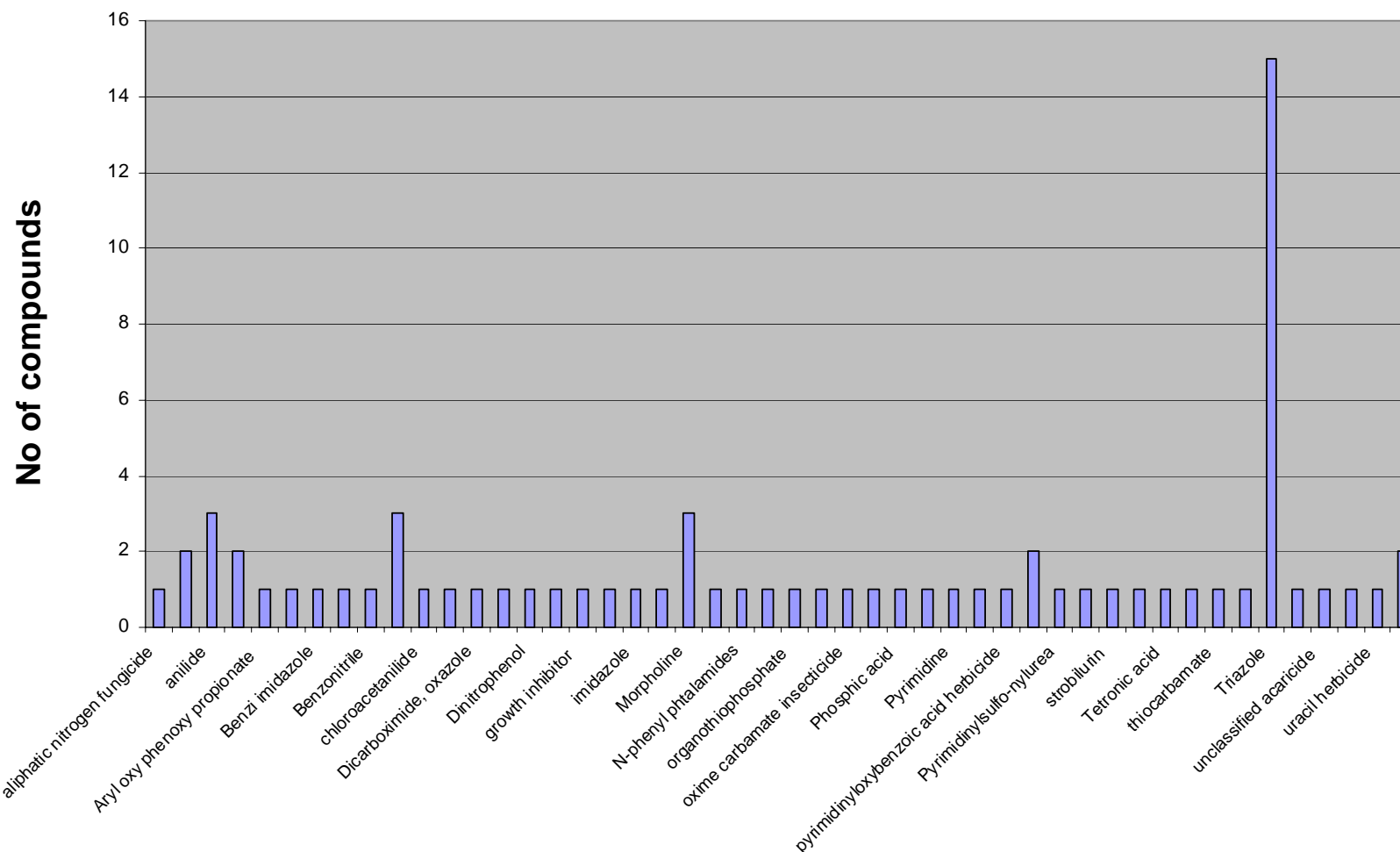


Figure 1.1. Structural diversity in the EFSA developmental toxicity dataset

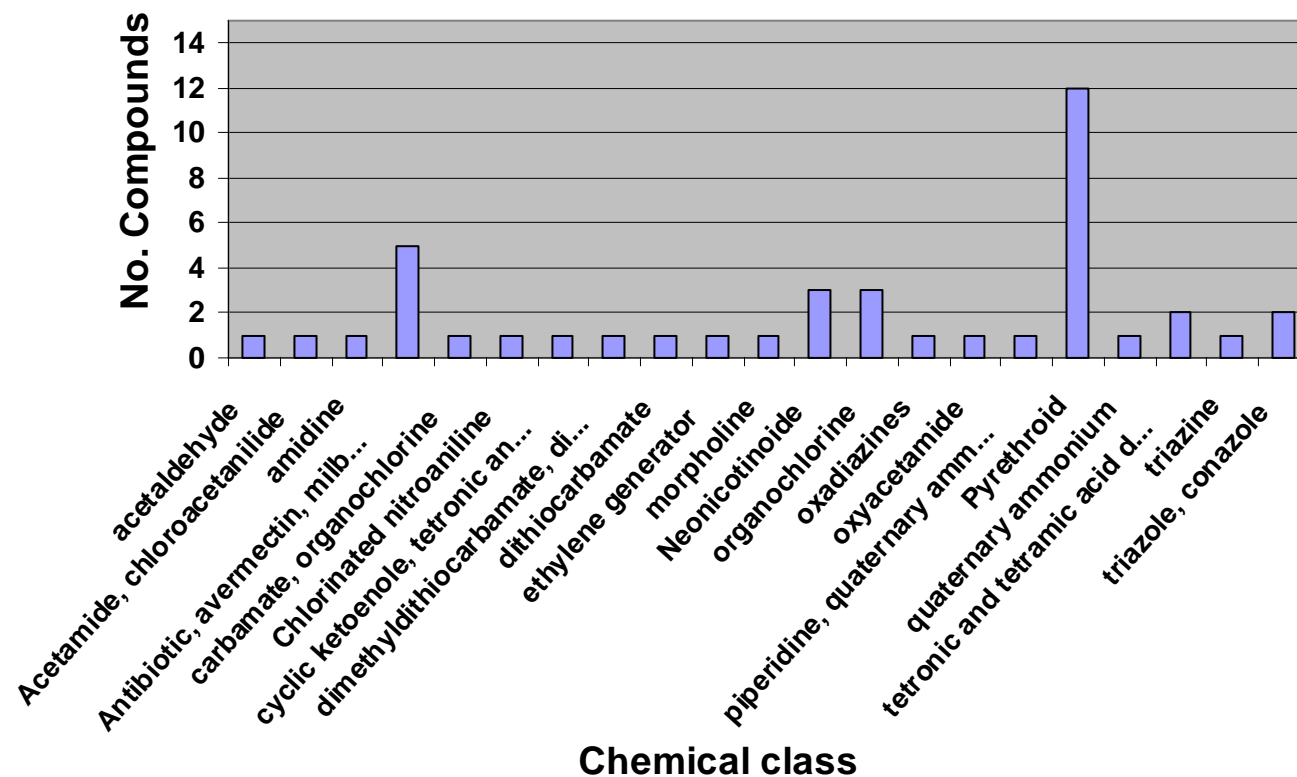


Figure 1.2. Structural diversity in the EFSA neurotoxicity dataset (42 positives)

2. QSAR analysis

2.1. Introduction to QSAR analysis

The term “QSAR analysis” is taken to include the development and use of Structure-Activity Relationships (SARs), Quantitative Structure Activity Relationships (QSARs), and computer-based tools (including expert systems) based on the use of one or more of these types of models.

Structure-Activity Relationships (SARs) and Quantitative Structure Activity Relationships (QSARs), collectively referred to as (Q)SARs, are theoretical models that relate the structure of chemicals to their biologic activities. (Q)SARs are used to predict the physicochemical, biological (e.g., toxicological) and fate properties of molecules from knowledge of chemical structure (Cronin, 2010).

More specifically, a SAR is a qualitative relationship between a molecular (sub)structure and the presence or absence of a given biological activity, or the capacity to modulate a biological activity imparted by another substructure. The term substructure refers to an atom, or group of adjacently connected atoms, in a molecule. A substructure associated with the presence of a biological activity is also called a structural alert. A SAR can also be based on the ensemble of steric and electronic features considered necessary to ensure the intermolecular interaction with a specific biological target molecule, which results in the manifestation of a specific biological effect. In this case, the SAR is sometimes called a 3D SAR or pharmacophore.

A QSAR is a quantitative relationship between a biological activity (e.g., toxicity), which may be categorical or quantitative, and one or more molecular descriptors that are used to predict the activity. A molecular descriptor is a structural or physicochemical property of a molecule, or part of a molecule, which specifies a particular characteristic of the molecule and is used as an independent variable in a QSAR. A comprehensive review of molecular descriptors has been published by Todeschini (Todeschini & Consonni, 2009).

An expert system is a formalised, computer based system that can be used to make predictions on the basis of prior information. Expert systems are based on three main modelling approaches referred to rule-based, statistically-based, or hybrid methods.

- Rule-based systems contain “if-then-else” rules that combine toxicological knowledge, expert judgment and fuzzy logic. Commonly used software tools based on this approach include Derek for Windows and HazardExpert. Derek and HazardExpert can be used in conjunction with their sister programs Meteor and Metabolexpert to predict the toxicity and carcinogenicity potential of metabolites as well as parent substances. In addition to these commercial tools, models included in the freely available Toxtree software are rule-based.
- Statistically-based systems use a variety of statistical, rule-induction, artificial intelligence, and pattern recognition techniques to build models, usually from non-congeneric databases. Statistically based systems are included in the commercial tools MultiCASE and TOPKAT, and the publicly available Lazar and CAESAR models. In addition, many models published in the literature and not implemented in software are statistically based.

- Hybrid models are based on a combination of knowledge-based rules and statistically-derived models. These are based on the general idea that, within the structural space of a single structural alert (considered to represent a single interaction mechanism), statistically derived models can quantitatively predict the variation in the reactivity of the alert conditioned by the rest of the molecular structure. Examples of the hybrid approach include models implemented in OASIS TIMES.

The advantages and disadvantages of the three main approaches are summarised in Table 2.1.

Table 2.1 Comparison of three main approaches in expert systems

Approach	Advantages	Disadvantages
Rule-based	<ul style="list-style-type: none"> • mechanistically connected to the predicted endpoint • provide reasoning for the predictions • in many cases support the prediction with literature references or expert knowledge 	<ul style="list-style-type: none"> • often restricted and/or ill-defined applicability domain • usually cannot explain differences of the activity within a chemical class • usually have lower accuracy of the prediction than statistical models
Statistical	<ul style="list-style-type: none"> • usually have high accuracy of the predictions • can be use for preliminary research when mechanism of action is unknown 	<ul style="list-style-type: none"> • usually difficult to interpret the model predictions • often do not provide mechanistically reasoning of the predictions • often non-transparent to the end-user
Hybrid	<ul style="list-style-type: none"> • combines advantages of rule-based and statistical approaches, including mechanistic interpretability and accuracy 	<ul style="list-style-type: none"> • likely to have restricted applicability domain

The QSAR tools used in this study are described in Sections 2.5-2.11.

2.2. Statistical measures of the performance of classification models

By comparing the predictions of the selected models against the known data (toxicological calls) in the various test sets, the predictive performance of each model was characterised in terms of the following statistics, which are commonly used to describe binary classification models:

- number of true positive (TP) and false positive (FP) predictions; sensitivity (percentage of positives correctly identified as positive), positive predictivity (average probability that a positive prediction is correct), and false positive rate (percentage of negatives falsely predicted as positive)
- number of true negative (TN) and false negative (FN) predictions; specificity (percentage of negatives correctly identified as negative), negative predictivity (average probability that a negative prediction is correct), and false negative rate (percentage of positives falsely predicted as negative)

- overall concordance (percentage of positives and negatives correctly identified)

It should be noted that the positive and negative predictivities are dependent on the ratio of positives to negatives in the test set. These statistics are summarised in Table 2.1.

Table 2.1. Statistics for classification models

Statistic	Definition	Meaning (proportion/percentage) of ...
Sensitivity	= TP / (TP+FN)	known positives that are correctly predicted
Specificity	= TN / (TN+FP)	known negatives that are correctly predicted
Concordance	= (TN+TP) / (TN+ TP+ FN+ FP)	known positives and negatives that are correctly predicted
False positive rate	= 1-specificity = FP/(TN+FP)	known negatives that are incorrectly predicted as positive
False negative rate	= 1-sensitivity = FN/ (TP+FN)	known positives that are incorrectly predicted as negative
Positive predictivity	= TP / (TP+FP)	positive predictions that are true positives (probability of a positive prediction being correct)
Negative predictivity	= TN / (TN+FN)	negative predictions that are true negatives (probability of a negative prediction being correct)

2.3. QSAR models for ADME prediction

A wide range of models have been published in the literature and implemented in software tools for the prediction of ADME and related physicochemical properties. Recent reviews are provided elsewhere (e.g. Mostrag-Szlichtyng & Worth 2010). This section describes the models and tools that were used in this study.

2.3.1. Literature model for placental barrier passage

A QSAR developed by Hewitt et al (2007) expresses transfer as a clearance index (CI) and was derived from a heterogeneous dataset of 78 substances. Placental transfer is expressed as a ratio between the clearance of a test substance (from the maternal circulation) and that of a reference substance (antipyrine), a small lipophilic substance known to be transferred across the placenta via passive diffusion. Thus, the higher the clearance index, the more readily the substance is predicted to cross the placental barrier. The QSAR is given by the following equation:

$$CI = -0.00246TPSA + 0.244\Sigma C2H5 + 0.139\Sigma Hal + 0.569$$

$$N = 78 \quad r^2 = 0.64 \quad q^2 = 0.58 \quad s = 0.192 \quad F = 45.6$$

where TPSA is the topological polar surface area, using polar contributions from N, O, S and P atoms; $\Sigma C2H5$ is the sum of ethyl groups; ΣHal is the sum of halogen atoms.

According to Hewitt et al, information on placental transfer can be used as a modulator of potential developmental toxicity. While placental transfer is not in itself sufficient to indicate toxicity, the authors argue that there is a general trend in which higher placental transfer is associated with greater toxic potential. In their test set of 57 substances (42 positives, 15

negatives) taken from Enoch et al (2009), 15% of toxicants had a placental transfer ratio below 0.4, compared with 62% of the non-toxicants 0.4.

2.3.2. ADMET Predictor

ADMET Predictor is developed by Simulations Plus (<http://www.simulations-plus.com/>) for the predictive modelling of ADMET (Absorption, Distribution, Metabolism, Elimination, and Toxicity) properties.

In this study, we used ADMET Predictor 5.0, which includes a qualitative model for predicting blood-brain barrier passage (probability of barrier penetration as low or high). We interpreted this outcome as a negative and positive prediction, respectively.

The software documentation offers no further information on the model, except for a classification accuracy of 97%, as well as a citation to Crivori et al (2000).

2.3.3. Accelrys ADME (add-in for Excel)

The Accelrys Accord for Excel program includes calculation of blood-brain (BB) barrier penetration. This property calculation functionality is based on a quantitative linear regression model for the prediction of blood-brain barrier penetration after oral administration, derived from over 800 substances that are known to enter the CNS after oral administration. According to the software documentation, the model was tested against a collection of 124 substances with known logBB values yielding a goodness-of-fit (R^2 value) of 0.889 and a standard deviation (SD) of = 0.306.

The model assigns brain-blood penetration levels as follows:

0. Very High: Brain-blood ratio greater than 5:1
1. High: Brain-blood ratio between 1:1 and 5:1
2. Medium: Brain-blood ratio between 0.3:1 and 1:1
3. Low: Brain-blood ratio less than 0.3:1
4. Undefined: Outside 99% confidence ellipse

We interpreted levels 0 and 1 as a positive prediction; levels 2 and 3 as a negative prediction; and level 4 as undefined (out of domain).

2.4. QSAR models for toxicity prediction

Various models have been published in the literature and implemented in software tools for the prediction of toxicological endpoints, including those relevant to the assessment of developmental toxicity and neurotoxicity. Recent reviews are provided elsewhere (e.g. Lapenna et al, 2010; Lo Piparo & Worth, 2010). This section describes the models and tools that were used in this study. The rationale for selecting these tools is provided in Section 5.

2.4.1. CAESAR

CAESAR comprises a series of statistically-based models developed within EU-funded CAESAR project (<http://www.caesar-project.eu>). The models have been implemented into open-source software and made publicly available, either as an online webservice or as a downloadable application. Predictions can be made for five endpoints: mutagenicity (Ames), carcinogenicity, developmental toxicity, skin sensitisation, and the bioconcentration factor. In the current version (v2.0) of the software, the mutagenicity and developmental toxicity models are available as downloadable tools.

The CAESAR developmental toxicity model was built using 292 substances from the dataset of Arena et al (2004). In fact, two classification models were developed in the CAESAR project: a random forest model based on 13 descriptors developed by using WEKA (Waikato Environment for Knowledge Analysis); and a model based on 6 descriptors developed by using Adaptive Fuzzy Partition (AFP). The CAESAR software implements the first of the models, the random forest. Methodological details are given in Cassano et al (2010).

CAESAR has its own in-built applicability domain assessment tool to help identify whether a chemical is out of domain on the basis of structural features or molecular descriptor values. A chemical is considered out of domain if it has a low degree (less than 0.5) of similarity to the training set, or if the probability of being active is close to 0.5. CAESAR gives the six substances in the training set that are most similar to the chemical of interest. If the similarity of the most similar one is below 0.5, the target chemical is out of the applicability domain. If the experimental values of the two most similar substances are in disagreement with the predicted value, the prediction is also considered unreliable.

In this study, the reliability assessments provided by the CAESAR developmental toxicity model were not taken into account, since they were found to be overly conservative (very few test chemicals were found to be in domain).

2.4.2. Derek

Derek is an expert system based on multiple structure alerts (2D SARs). It is developed by Lhasa Ltd, a non-profit company and educational charity (<https://www.lhasalimited.org/>). Derek contains hundreds of alerts covering a wide range of toxicological endpoints in humans, other mammals and bacteria. An alert consists of a toxicophore (a substructure known or thought to be responsible for the toxicity) and is associated with literature references, comments and examples. A key feature of Derek is the transparent reporting of the reasoning underlying each prediction. In principle, Derek should be used only for identifying positives, since the alerts are not designed to identify the absence of effects.

Derek can perceive stereochemistry when processing a query chemical. If an alert is deemed specific to a particular stereoisomer, this will be captured in the knowledge base. Otherwise, if there is no evidence to suggest that the toxicity is specific to a particular stereoisomer, the alert will consider all stereoisomeric forms.

All the rules in Derek are based either on hypotheses relating to mechanisms of action of a chemical class or on observed empirical relationships. Information used in the development of rules includes published data and suggestions from toxicological experts in industry, regulatory bodies and academia. The toxicity predictions are the result of two processes. The

program first checks whether any alerts in the knowledge base match toxicophores in the query structure. The reasoning engine then assesses the likelihood of a structure being toxic. There are nine levels of confidence: certain, probable, plausible, equivocal, doubted, improbably, impossible, open, contradicted. Derek can be integrated with Lhasa's Meteor software, which makes predictions of fate, thereby providing predictions of toxicity for both parent substances and their metabolites. Meteor is also capable of perceiving and handling stereochemistry.

The applicability domain of each alert within Derek is defined, in general terms, by the substructure and accompanying inclusion or exclusion rules. However, the concept of applicability domain for expert systems is not very well developed, and since there is no defined training set for the alerts and their combined use, it is not possible to use the applicability domain analysis applied to statistically based models. Research is in progress to define a suitable methodology for such systems (Ellison et al, 2009, 2011).

In this project, we used Derek Nexus 2.0, which includes the following alerts:

- 8 alerts for neurotoxicity: gamma-diketone or precursor; acrylamide or glycidamide; nitroimidazole; carbon disulphide or precursor; pyrethroid; 1-methyl-1,2,3,6-tetrahydropyridine; lead or lead compound; organophosphorus ester;
- 5 alerts for developmental toxicity: monothioglycol or glycol monoalkyl ether; alkoxy- or alkylthio-carboxylic acid or precursors; polyalkyl urea; epoxide; benzidine-based bisazo compound;
- 43 alerts for teratogenicity: hydroxamic acid or hydroxyurea derivate; 2-aminoquinazoline or analogue; retinoid or analogue; vinca alkaloid; phenol; hydantoin, phenyl barbiturate or analogue; 5-benzyl- or 5-phenyl-2,4-diaminopyrimidine, nitrogen or sulphur mustard; alkylthiourea; phenothiazine; cytidine analogue; alkyl sulphomate; short chain alkyl amide; triazole antifungal analogue; phthalate mono- or di-ester; 1,4-benzodiazepine or derivative; banzhdrylpiperazine or analogue; sartan; 4-hydroxycoumarin; N1-aryl- or N1-heteroaryl-4-aminophenylsulphonamide or analogue; vitamin D or analogue; N-acyl-pyrrolidine, -piperidine or analogue; 4-hydroxyquinazoline, nitrile, benzomorphan derivate; N-nitro or N-nitroso compounds; gamma-glutamyl derivative; benzimidazole carbamate or 2-aryl benzimidazole; aryl mono- or dialkyltriazene; hydrazine; pirroline ester, pirroline N-oxide ester, pirrole ester or pirrole alcohol; short chain carboxylic acid or precursor; arylsulphonylurea; 5-fluoropyrimidine; alkyl alcohol; thiuram disulphide or dicarbamate; aziridine; aryl sulphonamide; 5-halogenated pyrimidine; pyridine or 2-amino-1,3,4-thiadiazole derivative; phthalamide derivative; xanthine.

Among these alerts, the following teratogenicity alerts incorporate stereochemistry: aryl sulphonamide; vitamin D or analogue; cytidine analogue; and gamma-glutamyl derivative.

2.4.3. HazardExpert

HazardExpert is a module of the Pallas software developed by CompuDrug (<http://compudrug.com/>). It predicts the toxicity of organic compounds based on toxic fragments, and it also calculates bioavailability parameters (logP and pKa). It is a rule-based

system with an open knowledge base, allowing the user to expand or modify the data on which the toxicity estimation relies. It covers the following endpoints relevant to dietary toxicity assessment: carcinogenicity, mutagenicity, teratogenicity, membrane irritation, immunotoxicity and neurotoxicity

The results generated by HazardExpert (Pallas v 3.6.2.1) are provided as relative percentage toxicity values. On the basis of the ranges of the results, the software developers propose the classification of chemicals as “highly probable”, “probable”, “uncertain” and “not probable” to express activity. In order to compare the HazardExpert predictions with the results of other software tools we treated “highly probable” and “probable” chemicals as positive, “uncertain” chemicals as equivocal, and “not probable” ones as negative, as in Table 2.2.

Table 2.2. Interpretation of HazardExpert predictions

The range of relative percentage toxicity [%]	Toxic Class	Classification	Interpretation
100-60	1	Highly probable	Positive
59-48	2A	Probable	Positive
47-36	2B	Probable	Positive
35-3	3	Uncertain	Equivocal
2-0	4	Not probable	Negative

2.4.4. Leadscope Model Applier

Leadscope (<http://www.leadscope.com/>) includes models built from a pre-defined library of 27,000 hierarchically organised fragments (including functional groups, heterocycles and pharmacophores) that are typically found in small drug molecules. In addition, eight calculated molecular descriptors are available for use. The in-built applicability domain assessment is performed by means of a global analysis of the similarity in structural features between the test chemical and the training set chemicals.

In this study, we used Leadscope Model Applier Version 1.3.3, which includes the following classification models for developmental toxicity in the rat and rabbit foetus: foetal survival (foetal death, post-implantation loss and preimplantation loss), structural dysmorphogenesis and visceral organ toxicity.

2.4.5. PASS

PASS (Prediction of Activity Spectra for Substances) is a tool developed by the Institute of Biomedical Chemistry of the Russian Academy of Medical Sciences, Moscow. It predicts various toxicological effects including mutagenicity, carcinogenicity, teratogenicity and embryotoxicity, and also a range of mechanisms of action and pharmacological effects. The system uses a Bayesian algorithm to predict the biological activities in terms of the probabilities of presence (Pa) and absence (Pi) of each particular activity, by estimating the similarity/dissimilarity of the new substance to substances with well known biological activities present in the software training set (Poroikov et al, 2007). A freely accessible version is available online and a downloadable trial version is also available (<http://195.178.207.233/PASS/index.html>). Compared with the commercial version of PASS,

the freely available version contains a more restricted range of functionalities and predicted endpoints, and individual model predictions are generally based on smaller training sets.

In this study, we used PASS v 10.1, which includes:

- an embryotoxicity model for predicting the probability that a substance crosses the placental membrane and causes any toxic effect (e.g. foetal bradycardia, low birth weight) or death of an embryo.
- a teratogenicity model predicting the probability that a substance crosses the placental membrane and causes abnormal development of one or more body systems in the embryo.
- a neurotoxicity model for predicting the probability of neurotoxicity.

A positive prediction was considered to be a prediction with $P_a > P_i$, and negative otherwise ($P_a \leq P_i$).

2.4.6. TOPKAT

TOPKAT is a QSAR-based system, developed by Accelrys Inc. (<http://accelrys.com/>), makes predictions of a range of toxicological endpoints, including mutagenicity, developmental toxicity, rodent carcinogenicity, rat chronic LOAEL, rat Maximum Tolerated Dose (MTD) and rat oral LD50. The QSARs are developed by regression analysis for continuous endpoints and by discriminant analysis for categorical endpoints. TOPKAT models are derived by using a range of two-dimensional molecular, electronic and spatial descriptors. TOPKAT estimates the confidence in the prediction by applying the patented Optimal Predictive Space (OPS) validation method. The OPS is TOPKAT's formulation of the model applicability domain - a unique multivariate descriptor space in which a given model is considered to be applicable. Any prediction generated for a query structure outside of the OPS space is considered unreliable.

As suggested by the vendor, probability values were converted into binomial ones (positives or negatives) according to the following rules:

- if computed probability is greater than 0.7, then the prediction is positive (toxic);
- if computed probability is smaller than 0.3, then the prediction is negative (non-toxic);
- if computed probability is between 0.3 and 0.7, then the prediction is equivocal.

3. Grouping and read-across

3.1. Introduction to grouping and read-across

In addition to the formalised approach of QSAR analysis, it is possible to estimate chemical properties and endpoints by using a less formalised approach based on the grouping and comparison of chemicals. The grouping approach can be used, for example, to support the results of QSAR analysis or to generate estimated data (and fill data gaps) in the absence of suitable QSARs. The most comprehensive guidance currently available for applying the grouping approach has been published by the OECD (OECD, 2007) and by ECHA (ECHA, 2008). The ECHA and OECD guidance documents are scientifically equivalent, except that

the ECHA guidance makes additional references to REACH criteria and procedures. The concepts of grouping and read-across are further explained and illustrated by Enoch (2010).

The use of endpoint information for one chemical, called a “source chemical”, to make a prediction of the same endpoint for another chemical, called a “target chemical”, is termed “read-across”. The source and target chemicals are considered to be similar in some way, usually on the basis of structural similarity. It is assumed that, in general, similar substances will exhibit similar biological activity. In principle, read-across can be applied to characterise physicochemical properties, fate, human health effects and ecotoxicity, and it may be performed in a qualitative or quantitative manner, depending on whether the data being used is categorical or numerical in nature. To estimate the properties of a given substance, read-across can be performed in a one-to-one manner (one analogue used to make the estimate) or in a many-to-one manner (two or more analogues used). Read-across is equally well suited to the identification of positive and negative chemicals.

The reliability of read-across depends on the selection of appropriate analogues associated with the availability of reliable experimental data. In some cases, it is only possible to identify a limited number of suitable analogues, whereas in other cases, it is possible to build up a larger and more robust chemical group, called a chemical category. A chemical category is a group of chemicals whose physicochemical and human health and/or environmental toxicological and/or environmental fate properties are likely to be similar or follow a regular pattern as a result of structural similarity (or other similarity characteristic). The presence of common behaviour or coherent trends in the chemical category is generally associated with a common underlying mechanism of action. In general, the application of read-across between analogues in a chemical category is considered to be more reliable than the application of read-across in a smaller group of analogues (in which trends are not apparent).

In contrast with the use of (Q)SAR tools, the grouping and read-across approach is a more *ad hoc* approach involving a range of subjective choices in terms of categorisation tools, similarity metrics, datasets for the retrieval of analogue, and criteria for analogue selection. A broad chemical and toxicological expertise is needed to apply this approach. Consequently, the approach is unlikely to be reproducible, unless all of the expert choices are clearly documented.

Various tools can be used to assist grouping and read-across, including Toxmatch (Jeliaskova et al, 2010) and the OECD QSAR Toolbox (Diderich, 2010).

3.2. OECD QSAR Toolbox

In this project, we used the OECD QSAR Toolbox (v. 2.1.0.721) in order to group chemicals and perform read-across for developmental toxicity (Section 8). This freely available standalone software application (<http://www.qsartoolbox.org/>) is being developed to support the filling of data gaps needed for the hazard assessment of chemicals. Data gaps are filled by following a flexible workflow in which chemical categories are built and missing data are estimated by read-across or by applying local QSARs (trends within the category).

The Toolbox includes a range of profilers to quickly evaluate chemicals for common mechanisms or modes of action. There are four types of profilers: predefined (e.g. US EPA chemical categories); general mechanistic (e.g. DNA binding, protein binding); endpoint-

specific (e.g. classification of aquatic mode-of-action); and empiric (e.g. organic functional groups, structural similarity). It is also possible for the user to build customised profilers.

In order to support read-across and trend analysis, the Toolbox also contains numerous databases with results from experimental studies. The current version of the Toolbox does not contain any databases on neurotoxicity or developmental toxicity. However, it is possible for the user to import a database, if available.

Since the current version of the Toolbox does not contain any mechanistic or endpoint-specific profilers related to developmental toxicity, the organic functional groups (nested) profiler (developed by LMC, Bourgas University, Bulgaria) was used in order to characterise each query chemical and identify similar chemicals (in the ToxRef database).

4. Chemical space analysis

To rationalise the reliability of prediction for a query chemical, it is important to compare the chemical with the applicability domain of the model. Some software tools provide their own assessment of reliability based on applicability domain considerations, but other tools do not. Even in cases where the software applies its own applicability domain, the user may wish to consider an alternative assessment. It is important to realise that there is no absolute and unique definition of applicability domain – different definitions may be useful for different purposes. In general, the broader the chemical space covered, the lower the overall reliability of prediction. Conversely, the narrower the applicability domain, the higher the overall reliability of prediction. Furthermore, the relationship of the chemical to the model domain is indicative rather than conclusive. The fact that a chemical belongs to the domain of a model, does not guarantee that it will be predicted accurately; conversely, a chemical located outside the domain is not necessarily predicted incorrectly.

In the case of statistically-based models, an understanding of the applicability domain can be developed if the training set is provided, by constructing the domain in terms of a specific set of structural features and/or molecular descriptors. Test chemicals can then be compared in terms of their similarity in this chemical space. This exercise cannot be carried out for expert (knowledge-based) systems, such as Derek, since these models are not derived from explicit training sets.

In order to apply a chemical space analysis in this study, it was necessary to identify which of the statistically based models were transparent in terms of their underlying training sets (including availability of chemical structures and biological data).

In addition, in order to assess the representativeness / coverage of each test sets, the chemical space of each test set was compared with the chemical space of the Plant Protection Products (PPP) inventory. The chemical space of the PPP inventory was derived from a list of 821 pesticide actives for which structures were generated during the PESTISAR project (JRC, 2010). Following further structural processing and descriptor generation in the Dragon software, this list contained structures for 794 chemicals.

Although multiple considerations can be employed in the definition of chemical space, the two main approaches are the characterisation of structural (fragment) space, and the characteristics of molecular descriptor space, which provide complementary views.

4.1. Structural space analysis

Structural space was analysed by applying the systematic substructural analysis, based on 27,000 predefined structural fragments, in the Leadscape Enterprise (v.3.0.5-4) software (<http://www.leadscope.com/>). Major structural classes identified by Leadscape include: amino acids, bases and nucleosides, benzenes, carbocycles, carbohydrates, elements, functional groups, heterocycles, naphthalenes, natural products, peptidomimetics, pharmacophores, protective groups and spacer groups. The frequency distribution of fragments was identified in each dataset (PPP list, model training sets, and the test sets).

4.2. Molecular descriptor space analysis

Molecular descriptor space was analysed by using three interpretable Dragon descriptors, reflecting the three main types of descriptor typically used in QSAR modelling. These are descriptors of distribution/partitioning behaviour, molecular bulk and reactivity:

- 1) a measure of partitioning behaviour, the Moriguchi octanol-water partition coefficient (MLOGP)
- 2) a measure of the molecular volume in which a chemical can interact with biomolecules, the sum of atomic van der Waals volumes (Sv)
- 3) a measure of reactivity – the mean first ionisation potential (Mi). The higher the first ionisation potential of an atom, the more energy is required to remove an initial electron, and thus the more stable / less reactive the atomic site. Mi is a mean value taken over all atoms in the molecule. Thus, the reactivity (or electron-donating potential) of a molecule increases as Mi decreases.

Results

5. Available QSAR tools and databases

5.1. Software tools for developmental toxicity prediction

A list of software tools designed to predict developmental toxicity, along with an evaluation of their relevance for the purposes of this study, is given in Table 5.1.

Table 5.1. Software tools for predicting developmental toxicity

Software	Availability	Applicability
CAESAR 2.0 http://www.caesar-project.eu/	Freely available online and downloadable application	Two classification models for developmental toxicity based on the dataset of Arena et al. (2004) including 292 substances.
T.E.S.T http://www.epa.gov/nrmrl/std/cppb/qsar/index.html#TEST	Freely available	DevTox module same as CAESAR
Derek Nexus 2.0 http://www.lhasalimited.org/	Commercial	5 alerts for developmental toxicity and 43 alerts for teratogenicity
PALLAS HazardExpert v3.6.2.1 http://compudrug.com/	Commercial	Structural-rule based approach to toxicity prediction of organic compounds
Leadscope Model Applier Version 1.3.3, Developmental toxicity http://www.leadscope.com/	Commercial	Classification models for developmental toxicity in the rodent foetus: dysmorphogenesis (structural and visceral birth defects), developmental toxicity (foetal growth retardation and weight decrease), and foetal survival (foetal death, post-implantation loss, and preimplantation loss). For the purposes of this project, the models for structural and visceral dysmorphogenesis and foetal survival are relevant as endpoints whereas foetal growth is not considered relevant. Regarding species selection the mouse should be dismissed since it is not an adequate species for detecting developmental effects.
OSIRIS property explorer http://www.organic-chemistry.org/prog/peo/	Freely available	Classification model which predicts “undesirable” effects (mutagenicity, tumorigenicity, irritating effects and reproductive effects), mainly based on the RTECS database of >3500 substances. Not relevant for the purposes of this project, since it does not specifically make predictions for developmental toxicity.
PASS Institute of Biomedical Chemistry of the Russian	Freely available online and	Classification models giving probability of adverse effects. The embryotoxicity model

Software	Availability	Applicability
Academy of Medical Sciences, Moscow http://ibmc.p450.ru/PASS/ http://195.178.207.233/PASS/index.html	downloadable demo Commercial version also available, with additional functionalities and predicted endpoints	predicts the probability that a substance crosses the placental membrane and causes any toxic effect (e.g. fetal bradycardia, low birth weight) or death of an embryo. The teratogenicity model predicts the probability that a substance crosses the placental membrane and causes abnormal development of one or more body systems in the embryo.
TOPKAT (Accelrys) http://www.accelrys.com	Commercial	Classification model for developmental toxicity of pesticides, industrial chemicals.

5.2. Databases on developmental toxicity

A list of data sources potentially useful for reading across developmental toxicity and effects is given in Table 5.2.

Table 5.2. Databases containing information on developmental toxicity

Database	Availability	Information and remarks
ToxRefDB http://www.epa.gov/NCCT/toxrefdb/	Freely available	Standard toxicity test results for pesticides including developmental toxicity (387 chemicals) and multigeneration reproductive toxicity (316 chemicals). Considered suitable for the purposes of this study.
Toxicology Data Network (TOXNET) Developmental and Reproductive Toxicology Database (DART) http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?DARTETIC	Freely available	Bibliographic database containing over 200,000 references to literature published since 1965. It covers teratology and other aspects of developmental and reproductive toxicology. Users can search by subject terms, title words, chemical name, Chemical Abstracts Service Registry Number (RN), and author.
ICSAS Reprotox Database (US FDA) http://www.fda.gov/AboutFDA/CentersOffices/CDER/ucm092217.htm	Freely available	The ICSAS reproductive and developmental toxicity database contains data records from FDA segment I (reproductive toxicity in male and female animals), segment II (teratology, organ toxicity, and non-specific toxicity to the fetus), and segment III (behavioural toxicity in newborn pups) studies in Glires (primarily rats, mice, rabbits, and hamsters) and other animals. The data were acquired from publicly available sources, such as Shepard's Catalog of Teratogenic Agents, TERIS, REPROTOX, and RTECS, as well as studies reported in drug labeling, and other reproductive toxicity studies obtained from the EPA Toxdata-1g database. The combined developmental and reproductive toxicity database contains evaluated data for 2173 chemicals (most of them pharmaceuticals, plus limited numbers of industrial chemicals). Not ideally suited for the purposes of this study since it is mostly focussed on pharmaceuticals and the basis for the positive calls is not clear.
National Toxicology Program Bioassay On-line (NTPBSI) Database	Freely available	Developmental toxicity dataset containing data on 70 substances, without toxicological evaluation. Chemical list available from US EPA DSSTox website: http://www.epa.gov/ncct/dsstox/sdf_ntpbsi.html Searchable online at the National Toxicology Program website:

Database	Availability	Information and remarks
		http://ntp-apps.niehs.nih.gov/ntp_tox/index.cfm Not ideally suited for the purposes of this study, since it is a mixed dataset with different toxicological evaluations.
ILSI Developmental Toxicity database	Not yet available Under development	Expected to be available in downloadable format from the ILSI website (http://www.ilsil.org/Lists/Activities/AllItems.aspx) and via the US EPA DSSTox website (http://www.epa.gov/ncct/dsstox/)

5.3. Software tools for neurotoxicity prediction

A list of software tools designed to predict neurotoxicity, as well as blood-brain barrier passage, along with an evaluation of their relevance for the purposes of this study, is given in Table 5.3.

Table 5.3. Software tools for predicting neurotoxicity

Software	Availability	Applicability
ADMET Predictor 5.0 http://www.simulations-plus.com/	Commercial	Predicts the probability of blood-brain barrier penetration as low or high. Not directly relevant for the purposes of this study, since it does not predict the apical endpoint. Might provide useful supporting information.
Accelrys ADME add-in	Commercial	Blood-brain barrier passage A quantitative linear regression model for the prediction of blood-brain barrier penetration after oral administration. Not directly relevant for the purposes of this study, since it does not predict the apical endpoint. Might provide useful supporting information.
Derek Nexus 2.0 http://www.lhasalimited.org/	Commercial	Includes 8 alerts for neurotoxicity
PALLAS HazardExpert v3.6.2.1 http://compudrug.com/	Commercial	Structural-rule based approach to toxicity prediction of organic compounds
Leadscope Model Applier Version: 1.3.3, Neurotoxicity Suite http://www.leadscope.com/	Commercial	Developed by a CRADA with US FDA. The suite comprises three Rodent Newborn Behaviour models, which are: pup behaviour mouse (training set of 173 substances) pup behaviour rat (training set of 628 substances) pup behaviour rodent (training set of 672 substances) Not relevant for this study, since it is based on results obtained in developmental neurotoxicity tests
PASS Institute of Biomedical Chemistry of the Russian Academy of Medical Sciences, Moscow http://ibmc.p450.ru/PASS/ http://195.178.207.233/PASS/index.html	Freely available online and downloadable demo Commercial version also available, with additional functionalities and predicted endpoints	Neurotoxicity module + numerous receptor-mediated activity modules not considered directly relevant for the purposes of this study

5.4. Databases on neurotoxicity

The only data source initially identified as potentially useful for reading across neurotoxicity effects is the US EPA's ToxRefDB (<http://www.epa.gov/NCCT/toxrefdb/>). This includes

standard toxicity test results for pesticides and other environmental chemicals. However, on further examination, it was not considered suitable for use in this study. This was partly due to the technical difficulty in extracting and reconstructing the data, but also because the data were expected to have limited value as a basis for read-across for acute neurotoxicity, since the type of neurotoxic effects included seemed mostly related to effects occurring after repeated exposure rather than neurotoxic effects occurring after just a single exposure.

5.5. Conclusions on the availability of predictive tools and databases

For the purposes of this study, more QSAR tools and databases were identified as potentially useful for developmental toxicity prediction than for neurotoxicity prediction. In particular, there is a lack of freely available software tools for neurotoxicity prediction, as well as a lack of a suitable public reference database for the development of new QSARs and for the application of grouping and read-across.

6. Performance of selected QSAR tools

6.1. Assessment of a literature-based QSAR for placental transfer

Although it was not expected to be directly predictive of developmental toxicity, the QSAR for placental transfer developed by Hewitt et al (2008) was applied to the original and extended test sets. It turned out that the predicted clearance (expressed in relation to antipyrine as a passively-diffusing reference substance) was unable to discriminate between positive and negative chemicals, as illustrated in Figure 6.1 (for the extended test set). In the extended test set of 134 substances (one chemical, warfarin, could not be predicted), the average clearance for 95 positive substances was essentially the same as the average clearance for 39 negative substances, as shown in Table 6.1

Table 6.1. Distribution of predicted placental clearance values for developmental toxicants and non-developmental toxicants

Clearance statistics	Original test set		Extended test set	
	Positives (37)	Negatives (39)	Positives (95)	Negatives (39)
Average	0.66	0.70	0.66	0.70
Maximum	1.39	3.30	2.56	3.30
Minimum	0.15	0.14	-0.43	0.14
Standard deviation	0.25	0.54	0.39	0.54

6.2. Assessment of models for developmental toxicity prediction

6.2.1. Predictive performance of developmental toxicity models (used alone)

Table 6.2 shows the predictive performances of seven QSAR models when tested against the original and extended test sets. The key statistics are specificity, negative predictivity and

false negative rate. The statistics obtained with the extended test set are illustrated in a Receiver Operating Curve (ROC) in Figure 6.2.

A breakdown of the best-predicted chemicals (all models giving correct predictions) to the worst-predicted chemicals (all models giving incorrect predictions) is given in Table 6.3.

The best performing models, based on the statistics obtained with the extended test set, are the two PASS models, with a specificity of 59-62% (depending on the model), a negative predictivity of 44-45%, and a false negative rate of 29-31%.

6.2.2. Predictive performance of batteries of developmental toxicity models

Table 6.4 shows the predictive performances of the five best two-model combinations, based on the analysis of the extended test set. The following rule was used to combine the model predictions:

The overall prediction is negative if both models predict negative; positive if both models predict positive; and not determined (ND) otherwise.

This rule is intended to maximise the ability to correctly identify negatives, without also generating false negatives. In case the two models contradict each other, it is suggested not to rely on the QSAR prediction (and possibly proceed to a read-across evaluation).

The best performing combination of models, based on the statistics obtained with the extended test set, was TOPKAT combined with PASS (teratogenicity), with a negative predictivity of 48%, a false negative rate of 14%, and a specificity of 32%. This was followed by Derek combined with PASS (embryotoxicity), with a negative predictivity of 46%, a false negative rate of 27%, and a specificity of 56%.

6.2.3. Assessment of the developmental toxicity models against the ToxRefDB dataset

The models were also tested for their ability to predict the developmental toxicity potential of the 366 pesticide actives (246 positives; 120 negatives) in the US EPA's ToxRef Database. The results of this analysis are given in Table 6.5.

6.3. Assessment of models for neurotoxicity prediction

6.3.1. Predictive performance of neurotoxicity models (used alone)

Table 6.6 shows the predictive performances of five QSAR models (3 neurotoxicity and 2 BBB passage) when tested against the EFSA test set. The key statistics are specificity (percentage of negatives correctly identified as negative), negative predictivity (average probability that a negative prediction is correct), and false negative rate (percentage of positives falsely predicted as negative). The statistics are illustrated in a Receiver Operating Curve (ROC) in Figure 6.3.

The best performing model is Derek, with a negative predictivity of 43%, a false negative rate of 74%, and a specificity of 100%.

The statistics for the BBB passage models are reported for completeness, even though they are not considered directly relevant (since they predict distribution to the central nervous system, but not the peripheral nervous system).

A breakdown of the best-predicted chemicals (all models give correct predictions) to the worst-predicted chemicals (all models give incorrect predictions) is given in Table 6.7.

Table 6.6 also shows the statistics obtained with the reduced test in which chemicals expected to be positive, based simply on their chemical class, have been removed. These substances comprise 3 carbamates, 3 neonicotinoids, 13 pyrethroids and 2 triazoles.

The only model that appears to perform better on this narrower range of pesticides is PASS (increase of sensitivity from 43 to 62%; marginal increase of specificity from 65 to 68%; and increase in negative predictivity from 38 to 65%).

The predictions obtained for these “positive classes” are given in Table 6.8. Such results could be used by the model developers to improve the performances of their models. For example, the Derek knowledgebase could be extended to refine the alert for pyrethroids (to avoid false negatives) and to include alerts for neonicotinoids and triazoles.

6.3.2. Predictive performance of batteries of neurotoxicity models

Table 6.9 shows the predictive performances of the three two-model combinations of the neurotoxicity QSARs. The same rule was used to combine the model predictions.

The use of any two-model combination increases the negative predictivity to 48%. However, this is accompanied by an increased false negative rate (to 84-94%, depending on which two models are combined).

6.4. Conclusions on the performance of QSAR models and tools

6.4.1. Placental transfer and developmental toxicity

The literature-based model for placental transfer shows no tendency to distinguish between developmental toxicants and non-toxicants. This was not unexpected since the passage of a chemical across the fully formed placental barrier is not a sufficient condition for embryotoxicity. Furthermore, even if a chemical does not pass the barrier, this does not necessarily mean that it will not be a developmental toxicant, since adverse developmental effects may arise prenatally before implantation, during pregnancy while the placental barrier is being formed, or even postnatally. By definition, a developmental toxicant is a substance that can affect development from the time of conception until sexual maturity (UN, 2009). Thus, the conclusions of this study are not in agreement with the conclusions reported by Hewitt et al (2010) based on the application of the same QSAR for placental transfer developed earlier by Hewitt et al (2009). However, this difference can be explained on the grounds that the placental model was based on a heterogeneous dataset (including many drugs and industrial chemicals) and the subsequent comparison of its predictions with developmental toxicity was based on toxicity data taken from Briggs et al (2002) and interpreted according to the FDA classification scheme for the risk of teratogenic effects of drugs (Friedman, 1993). Thus, the placental transfer model was not specifically developed for pesticides and the toxicological data were interpreted according to different criteria.

To predict the absence of developmental toxicity, the best performing stand-alone models, based on the statistics obtained with the extended test set, are the two PASS models, with a specificity of 59-62% (depending on the model), a negative predictivity of 44-45%, and a false negative rate of 29-31%. The best performing combination of models, based on the statistics obtained with the extended test set, appears to be TOPKAT combined with PASS (teratogenicity), with a negative predictivity of 48%, a false negative rate of 14%, and a specificity of 32%. This was followed by Derek combined with PASS (embryotoxicity), with a negative predictivity of 46%, a false negative rate of 27%, and a specificity of 56%. Since the negative predictivities of these two-model combinations are less than 50%, none is expected to be adequate for use.

When assessed against the ToxRefDB dataset, the performances of the developmental models depended on how the Lowest Effect Levels (LELs) for developmental effects were compared with the LELs for maternal effects. When a “positive” is interpreted as any adverse effect (in rat or rabbit), provided that the dose in the developmental study is less than that in the maternal study ($dLEL < mLEL$), the best model could now be regarded as Leadscape, with a negative predictivity of 87%, a false negative rate of 37%, and a specificity of 47%. These statistics are different to those obtained with the extended test set, and are considered less relevant since the chemicals in ToxRefDB were neither selected by EFSA nor based on EU classification criteria.

In fact, it appears that the interpretation of developmental toxicity is more conservative for chemicals in the EFSA dataset than in ToxRefDB. A total of 35 substances were found to be common to the two test sets (Table 6.10). Of these 35 substances, 18 were considered positive in the EFSA test set, and 17 were considered negative; conversely, 11 were considered positive in ToxRefDB, and 24 were considered negative; 10 of the 18 EFSA positives were also ToxRefDB positives, and 16 of the 17 EFSA negatives were also ToxRefDB negatives.

In contrast to the identification of non-developmental toxicants, the statistics obtained with the EFSA Extended Test Set indicate that some tools, such as Derek, HazardExpert and PASS, might be useful for the identification of developmental toxicants (due to their high positive predictivities), but if used on their own, they would not serve the purpose of reducing the need to conduct short-term exposure assessments. However, they could be useful in the context of a stepwise assessment strategy in which the use of QSAR is followed by the use of grouping and read-across.

6.4.2. Neurotoxicity

To predict the absence of neurotoxic potential, no individual model for neurotoxicity, and no two-model combination, appears adequate for use (since the negative predictivities are less than 50%).

Conversely, the statistics indicate that some tools, such as Derek and HazardExpert, might be useful for the identification of neurotoxicants (due to their high positive predictivities between 90-100%), but if used on their own, they would not serve the purpose of reducing the need to conduct short-term exposure assessments. However, it is possible that they might be useful in the context of a stepwise assessment strategy in which the use of QSAR is followed by the use of grouping and read-across. This possibility could not be explored in this study,

due to the lack of a suitable reference database for the read-across exercise. To investigate the hypothesis, it would be necessary to develop such a database.

In view of the lack of available tools for predicting the absence of neurotoxic potential, the most pragmatic consideration in the risk assessment of pesticide metabolites/degradates is to apply the hypothesis that non-neurotoxic parent substances do not generate neurotoxic (bio)transformation products. Based on its experience of evaluating pesticide dossiers, EFSA could not identify any evidence that refutes this hypothesis. To explore this idea further, we carried out a web-based literature search using the powerful Scifinder tool (<http://cas.org/products/scifinder/index.html>). However, the only evidence we could find of non-neurotoxic parents (but not of pesticides) giving rise to neurotoxic products were a few papers describing *in vitro* / mechanistic findings (Brain et al, 1998; Dingemans et al, 2010), which are not necessarily relevant to the *in vivo* effects of pesticides.

Table 6.2. Predictive performance of developmental toxicity models (used alone against EFSA test set)

	Derek		Caesar		TOPKAT		Leadscope		HazardExpert		PASS (embryotoxicity)		PASS (teratogenicity)	
	A	B	A	B	A	B	A	B	A	B	A	B	A	B
% of chemicals														
sensitivity	14	27	73	66	51	53	33	45	61	49	35	70	32	69
specificity	97	97	24	24	57	57	48	48	73	73	59	59	62	62
concordance	57	47	48	54	54	54	40	46	66	53	47	67	47	67
negative predictivity	54	35	47	21	55	33	39	22	53	24	49	45	49	44
positive predictivity	83	96	48	68	53	75	42	74	79	89	45	81	44	81
false negative rate	86	73	27	34	49	47	67	55	39	51	65	29	68	31
false positive rate	3	3	76	76	43	43	52	52	27	27	41	41	38	38
No of chemicals (A 76, B 135 in total)														
TP	5	26	27	63	18	48	10	39	11	24	13	68	12	66
TN	38	38	9	9	21	21	13	13	8	8	23	23	24	24
FP	1	1	29	29	16	16	14	14	3	3	16	16	15	15
FN	32	70	10	33	17	43	20	47	7	25	24	28	25	30
ND	0	0	1	1	4	7	19	22	47	75	0	0	0	0

ND = Not determined; FN=false negative; FP=false positive; TN= true negative; TP= true positive;
A – original test set (76 substances); B – extended test set (135 substances)

Table 6.3. Best and worst predicted chemicals (developmental toxicity)

No of models making correct predictions	No of chemicals	Chemical names
6 out of 7 (best predicted chemicals)	1 (positive)	2-Methoxyethanol
6 out of 7	2 (negative)	Imazaquin Mepiquat
	9 (positive)	Carbendazim Propineb metabolite PTU (propylene thiourea) N-methylformamide Benomyl 2-Methoxypropanol 2-Methoxypropyl acetate N-methylacetamide 2-Ethoxyethanol C.I. Direct Blue 6
5 out of 7	7 (negative)	Benalaxyl-M Famoxadone Fenazaquin Fludioxonil Flufenacet Triazoxide Pencycuron
	14 (positive)	Diniconazole (M) Dinoseb Mancozeb and maneb common metabolite ETU (ethylene thiourea) Methoxyacetic acid 1-2-Dimethoxyethane Formamide N,N-Dimethylformamide 1-Methyl-2-pyrrolidone Fluazifop-butyl N,N-dimethylacetamide Bis(2-ethylhexyl) phthalate (DEHP) Cycloheximide 2-(2-Methoxyethoxy)ethanol C.I. Direct Red 28
4 out of 7	7 (negative)	Carboxin Diflubenzuron Dodine Metaflumizone Metrafenone Lenacil Lufenuron
	26 (positive)	Metabolite Desthio-prothiconazole (ethylene thiourea) Fluazinam Nitrofen (ISO) Octabromobiphenyl ether 2-(2-Aminoethylamino)ethanol 1,2-Diethoxyethane

No of models making correct predictions	No of chemicals	Chemical names
		Bis(2-Methoxyethyl)ether Dibutyl phthalate Dinocap 1,2-Bis(2-Methoxyethoxy)ethane 1,2-Benzenedicarboxylic acid; di-C6-8-branched alkylesters 1,2-benzenedicarboxylic acid, di-C7-11-branched and linear alkylesters Diisobutyl phthalate Benzo[a]pyrene Warfarin 6-(2-chloroethyl)-6-(2-methoxyethoxy)-2,5,7,10-tetraoxa-6-silaundecane 2-Ethylhexyl[[[3,5-bis(1,1-dimethylethyl)-4-Hydroxyphenyl]methyl]thio]acetate 4-Hydroxy-3,5-diiodobenzonitrile Bromoxynil phenol Nonylphenol Chlorotoluron 2-Ethylhexanoic acid 1,3,5-Trioxan Thiourea Mirex C.I. Direct Black 38
3 out of 7	15 (negative)	Amidosulfuron Azoxystrobin Chlormequat Chlorothalonil Dicamba Dicloran Dimethenamid-P Fosthiazate Nicosulfuron Rimsulfuron Tri-allate Teflubenzuron Hexythiazox Metamitron Phosmet
	19 (positive)	Bitertanol Bromuconazole Dichlobenil Dodemorph acetate Epoxiconazole Fenoxaprop-P Tridemorph Vinclozolin Prothioconazole Tetraconazole Ammonium pentadecafluorooctanoate Tridemorph N-Methylcaprolactam Butyl benzyl phthalate

No of models making correct predictions	No of chemicals	Chemical names
		Tetrachloroethylene Trichloroethylene Piperazine 1-Bromopropane Quinoclamine
2 out of 7	8 (negative)	Clofentezine Propisochlor Propyzamide Spiromesifen Aclonifen Benfluralin Bispyribac sodium Methomyl
	17 (positive)	Abamectin Azafenidin Fenarimol Flonicamid Flumioxazin Glufosinate-ammonium Trialkoxydim Flurprimidol Linuron Penconazole Triadimenol Triazole common metabolite 1,2,4-triazole Triazole common metabolite triazole alanine Imidazole, N, N'-1,2-ethenediyl-methanimidamide Di-n-pentylphthalate Methyl isocyanate Toluene
1 out of 7	0 (negative)	
	8 (positive)	Cyproconazole Fluzilazole Myclobutanil Tebuconazole Fenpropimorph Metconazole Isoxaflutole Amitrole (ISO)
0 out of 7 worst (predicted chemicals)	0 (negative)	
	2 (positive)	Hymexazole Oxadiargyl

The following software tools were applied: Derek, Caesar, TOPKAT, Leadscope, HazardExpert, PASS (embryotoxicity) and PASS (teratogenicity)

Table 6.4. Predictive performance of batteries of developmental toxicity models (used in combination for the identification of negatives)

%	Derek & TOPKAT	Derek & PASS (embryotoxicity)	Derek & PASS (teratogenicity)	TOPKAT & PASS (embryotoxicity)	TOPKAT & PASS (teratogenicity)
% of chemicals					
sensitivity	64	73	71	85	84
specificity	54	56	59	32	32
concordance	61	68	67	70	69
negative predictivity	38	46	45	46	44
positive predictivity	77	80	81	75	75
false negative rate	36	27	29	15	16
false positive rate	46	44	41	68	68
No of chemicals (135 in total)					
TP	58	70	68	78	76
TN	20	22	23	12	12
FP	17	17	16	25	25
FN	33	26	28	13	15
ND	7	0	0	7	7

Table 6.5. Predictive performance of developmental toxicity models (used alone against the ToxRefDB dataset)

	Derek			Caesar			TOPKAT			Leadscope			HazardExpert			PASS (embryotoxicity)			PASS (teratogenicity)		
% of chemicals	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C
sensitivity	6	2	6	75	81	78	40	42	41	51	63	49	36	32	35	62	53	61	58	47	57
specificity	91	93	92	27	27	29	59	60	61	40	47	42	59	58	60	30	33	32	33	37	35
concordance	33	76	46	59	37	55	46	56	50	47	49	45	44	53	47	52	37	47	49	39	47
negative predictivity	32	81	46	34	86	55	32	82	48	32	87	45	31	81	48	28	75	43	27	76	43
positive predictivity	56	5	44	68	20	55	67	19	55	59	18	46	65	13	47	64	15	50	64	15	50
false negative rate	94	98	94	25	19	22	60	58	59	49	37	51	64	68	65	38	48	39	42	53	43
false positive rate	9	7	8	73	73	71	41	40	39	60	53	58	41	42	40	70	67	68	68	63	65
No of chemicals	366	317	366	366	317	366	366	317	366	366	317	366	366	317	366	366	317	366	366	317	366
TP	14	1	11	185	48	151	89	22	73	91	24	70	33	6	24	152	31	118	141	28	110
TN	106	236	155	32	69	51	64	140	95	42	97	59	26	57	40	36	85	55	39	96	61
FP	11	18	14	88	189	122	44	95	60	62	110	83	18	42	27	84	173	118	81	162	112
FN	230	57	181	61	11	42	136	30	105	89	14	72	58	13	44	93	28	74	104	31	82
ND	5	5	5	0	0	0	33	30	33	82	72	82	231	199	231	1	0	1	1	0	1

A) The mLEL was neglected, and the toxicological data were evaluated as follows: If dLEL=NE then negative; if dLEL≠NE then positive.

B) The mLEL was taken into account, and the toxicological data were evaluated as follows: If mLEL=NE and dLEL=NE then negative (no maternal or developmental effects); if mLEL≠NE and dLEL=NE then negative (maternal effect, but no developmental effect); if mLEL>dLEL then positive (maternal effect at higher dose than developmental effect); if mLEL<dLEL then negative (maternal effect at lower dose than developmental effect); if mLEL=dLEL then undefined – ND (maternal and developmental effects at same dose).

C) The mLEL was taken into account, and the toxicological data were evaluated as follows: If mLEL=NE and dLEL=NE then negative (no maternal or developmental effects); If mLEL≠NE and dLEL=NE then negative (maternal effect, but no developmental effect); If mLEL>dLEL then positive (maternal effect at higher dose than developmental effect); if mLEL<dLEL then negative (maternal effect at lower dose than developmental effect); If mLEL=dLEL then positive (maternal and developmental effects at same dose).

Table 6.6. Predictive performance of neurotoxicity and blood-brain barrier models (used alone against the EFSA test sets)

	Derek		HazardExpert		PASS (neurotoxicity)		ADMET BBB		Accord BBB	
% of chemicals	A	B	A	B	A	B	A	B	A	B
sensitivity	26	0	21	14	43	62	40	38	67	54
specificity	100	100	96	95	65	68	70	68	50	47
concordance	52	51	48	56	51	65	51	53	61	50
negative predictivity	43	51	40	54	38	65	39	54	44	54
positive predictivity	100	-	90	75	69	65	71	53	71	47
false negative rate	74	100	79	86	57	38	60	62	33	46
false positive rate	0	0	4	5	35	32	30	32	50	53
No of chemicals	65	43	65	43	65	43	65	43	46	28
TP	11	0	9	3	18	13	17	8	20	7
TN	23	22	22	21	15	15	16	15	8	7
FP	0	0	1	1	8	7	7	7	8	8
FN	31	21	33	18	24	8	25	13	10	6
ND	0	0	0	0	0	0	0	0	19	15

BBB = Blood-Brain Barrier; ND = Not determined; FN=false negative; FP=false positive; TN= true negative;

TP= true positive

A) Original dataset including known classes of neurotoxicants (42 positives; 23 negatives)

B) Reduced dataset excluding known classes of neurotoxicants (21 positives; 22 negatives)

Table 6.7 Best and worst predicted chemicals (neurotoxicity)

No of models making correct predictions	No of chemicals	Chemical names
3 out of 3 (best predicted chemicals)	15 (negative)	6-Benzyladenine Azimsulfuron Azoxystrobin Bispyribac sodium Carboxin Hexythiazox Imazaquin Triflumuron Fludioxonil Amidosulfuron Bifenox Metazachlor Propaquizafop (1E,Z)-1,3-Dichloropropene Captane
	0 (positive)	
2 out of 3	7 (negative)	Bupirimate Dodine Guazatine n=0 Guazatine n=1 Guazatine n=2 Aclonifen Asulam
	6 (positive)	Acrinathrin Thiram Ziram Dicofol Gamma-cyhalothrin Lambda-cyhalothrin
1 out of 3	1 (negative)	2-Phenylphenol
	26 (positive)	Bifenthrin (1S,3S) Endosulfan Indoxacarb Tau-fluvalinate (R-cyano) Tefluthrin (Z-(1R,3R) Tri-allate a-Cypermethrin b-Cyfluthrin (1R,3R,αR) Cyfluthrin Cypermethrin Deltamethrin Zeta-cypermethrin Avermectin B1a Avermectin B1b Amitraz Chlormequat Dichloran Emamectin benzoate Esfenvalerate Eenpropimorth Imidacloprid

No of models making correct predictions	No of chemicals	Chemical names
		Lindane Mepiquat chloride Metaldehyde Milbemectin A3 (51596-10-2) Milbemectin A4 (51596-11-3)
0 (worst predicted chemicals)	0 (negative)	
	10 (positive)	Acetamiprid Acetochlor Ethephon Fufenacet Metribuzin Spiromesifen Spirotetramat Thiacloprid Triadimenol Triadimefon

The following software tools were applied: Derek, HazardExpert, PASS (neurotoxicity)

Table 6.8 Neurotoxicity predictions for “positive” classes of pesticides

No	Chemical name	Chemical class	PASS	Derek	HazardExpert
2	Acetamiprid	Neonicotinoid	0	0	0
4	Acrinathrin	Pyrethroid	1	1	0
5	a-Cypermethrin	Pyrethroid	0	1	0
7	b-Cyfluthrin (1R,3R,αR)	Pyrethroid	0	1	0
8	Bifenthrin (1S,3S)	Pyrethroid	0	0	1
10	Cyfluthrin	Pyrethroid	0	1	0
11	Cypermethrin	Pyrethroid	0	1	0
12	Deltamethrin	Pyrethroid	0	1	0
17	Esfenvalerate	Pyrethroid	1	0	0
21	Gamma-cyhalothrin	Pyrethroid	0	1	1
22	Imidacloprid	Neonicotinoid	1	0	0
24	Lambda-cyhalothrin	Pyrethroid	0	1	1
32	Tau-fluvalinate (R-cyano)	Pyrethroid	0	0	1
33	Tefluthrin (Z-(1R,3R)	Pyrethroid	0	0	1
34	Thiacloprid	Neonicotinoid	0	0	0
35	Thiram	Dithiocarbamate	1	1	0
36	Tri-allate	Thiocarbamate, Organochlorine	0	0	1
37	Triadimenol	Triazole	0	0	0
38	Triadimefon	Triazole	0	0	0
39	Zeta-cypermethrin	Pyrethroid	0	1	0
40	Ziram	Dithiocarbamate	1	1	0

Table 6.9 Predictive performance of batteries of neurotoxicity models (used in combination for the identification of negatives)

Percentage	Dere k	PASS	Derek & PASS
sensitivity	26	43	16
specificity	100	65	100
concordance	52	51	51
negative predictivity	43	38	48
positive predictivity	100	69	100
false negative rate	74	57	84
false positive rate	0	35	0
ND	0	0	46
	Dere k	HazardExper t	Derek & HazardExpert
sensitivity	26	21	8
specificity	100	96	100
concordance	52	48	50
negative predictivity	43	40	48
positive predictivity	100	90	100
false negative rate	74	79	92
false positive rate	0	4	0
ND	0	0	26
	PASS	HazardExper t	PASS & HazardExpert
sensitivity	43	21	6
specificity	65	96	94
concordance	51	48	48
negative predictivity	38	40	48
positive predictivity	69	90	50
false negative rate	57	79	94
false positive rate	35	4	6
ND	0	0	49

ND – number of non-determined chemicals (no predictions)

Table 6.10. Comparison of the EFSA Extended Test Set with the ToxRefDB data set

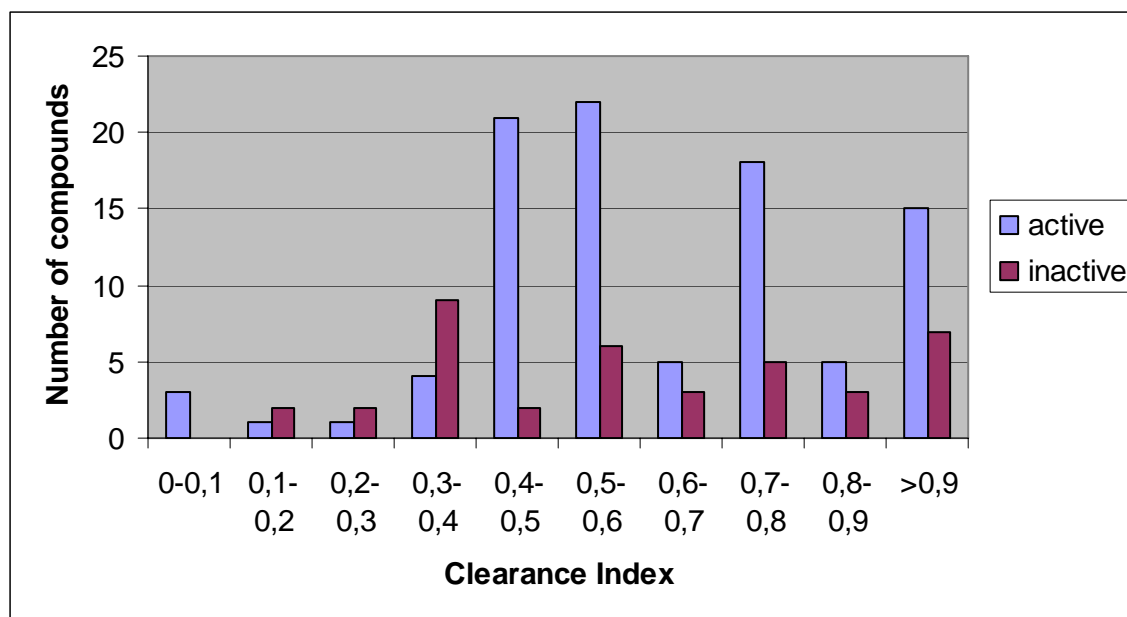
Name / CAS No	Rat study	Rabbit study	mLEL Rat	dLEL Rat	mLEL Rabbit	dLEL Rabbit	ToxRef call	EFSA call
Amitrole (ISO) (61-82-5)	X	X	NE	1000	80	80	1	1
Benomyl (17804-35-2)	X	X	NE	62.5	180	180	1	1
Carbendazim (10605-21-7)	X		90	20	NaN	NaN	1	1
Cyproconazole (94361-06-5)	X	X	12	12	50	10	1	1
Fluazinam (79622-59-6)	X	X	250	50	7	4	1	1
Fludioxonil (131341-86-1)	X	X	1000	100	100	NE	1	0
Flumioxazin (103361-09-7)	X	X	30	10	3000	NE	1	1
Fluzilazole (85509-19-9)	X	X	50	0.4	35	35	1	1
Isoxaflutole (141112-29-0)	X	X	500	100	100	5	1	1
Tralkoxydim (87820-88-0)	X	X	200	3	100	100	1	1
Triadimenol (55219-65-3)	X	X	15	5	125	NE	1	1
Azoxystrobin (131860-33-8)	X		25	NE	NaN	NaN	0	0
Benfluralin (1861-40-1)	X	X	475	NE	100	NE	0	0
Bispyribac sodium (125401-92-5)	X		1000	NE	NaN	NaN	0	0
Bromuconazole (116255-48-2)	X	X	70	70	50	200	0	1
Carboxin (5234-68-4)	X	X	90	NE	375	NE	0	0
Chlorothalonil (1897-45-6)	X	X	400	NE	20	NE	0	0
Clofentezine (74115-24-5)	X		320	NE	NaN	NaN	0	0
Dicamba (1918-00-9)	X	X	400	NE	150	300	0	0
Dimethenamid-P (163515-14-8)	X		25	150	NaN	NaN	0	0
Famoxadone (131807-57-3)	X	X	500	NE	1000	NE	0	0
Fenarimol (60168-88-9)		X	NaN	NaN	150	NE	0	1
Flonicamid (158062-67-0)	X		500	NE	NaN	NaN	0	1
Fosthiazate (98886-44-3)	X	X	10	NE	2	NE	0	0

Name / CAS No	Rat study	Rabbit study	mLEL Rat	dLEL Rat	mLEL Rabbit	dLEL Rabbit	ToxRef call	EFSA call
Glufosinate-ammonium (77182-82-2)	X	X	50	250	6.3	6.3	0	1
Hexythiazox (78587-05-0)	X	X	720	720	NE	NE	0	0
Imazaquin (81335-37-7)		X	NaN	NaN	500	NE	0	0
Linuron (330-55-2)	X	X	50	50	25	100	0	1
Methomyl (16752-77-5)	X	X	33.9	NE	16	NE	0	0
Ammonium pentadecafluorooctanoate (3825-26-1)	X	X	150	NE	50	50	0	1
Propyzamide (23950-58-5)		X	NaN	NaN	20	NE	0	0
Rimsulfuron (122931-48-0)	X	X	6000	NE	170	NE	0	0
Tetraconazole (112281-77-3)	X	X	22.5	100	30	NE	0	1
Tri-allate (2303-17-5)	X		30	90	NaN	NaN	0	0
Vinclozolin (50471-44-8)		X	NaN	NaN	200	400	0	1

mLEL Rat - overall maternal LEL in rat; dLEL Rat - overall developmental LEL in rat

mLEL Rabbit - overall maternal LEL in rabbit; dLEL Rabbit - overall developmental LEL in rabbit

NaN - LEL not assessed (study not available); NE - No Effect



Predicted placental transfer is expressed as a ratio between the clearance of a test substance and that of a reference substance (antipyrine), a small lipophilic substance known to be transferred across the placenta via passive diffusion. The QSAR used to predict placental transfer was developed by Hewitt et al (2007).

Figure 6.1 Distribution of developmental toxicants and non-developmental toxicants according to predicted placental transfer (135 substances)

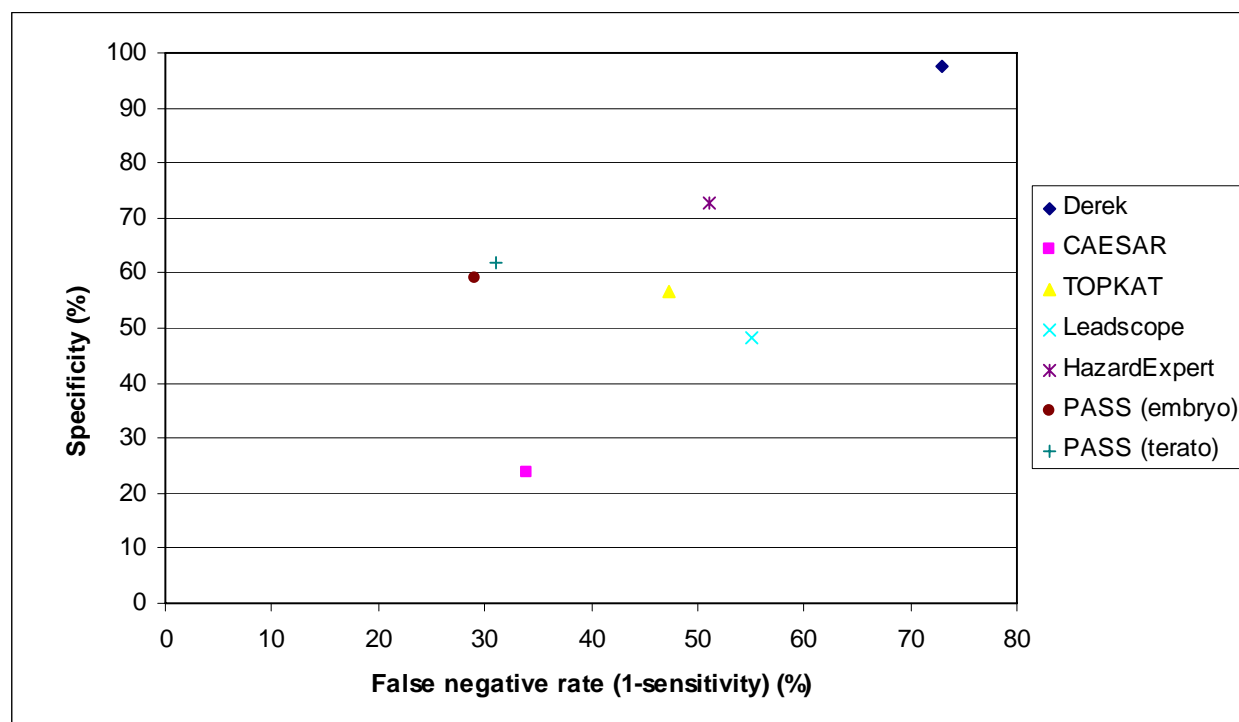


Figure 6.2 ROC curve for the developmental toxicity models (135 substances)

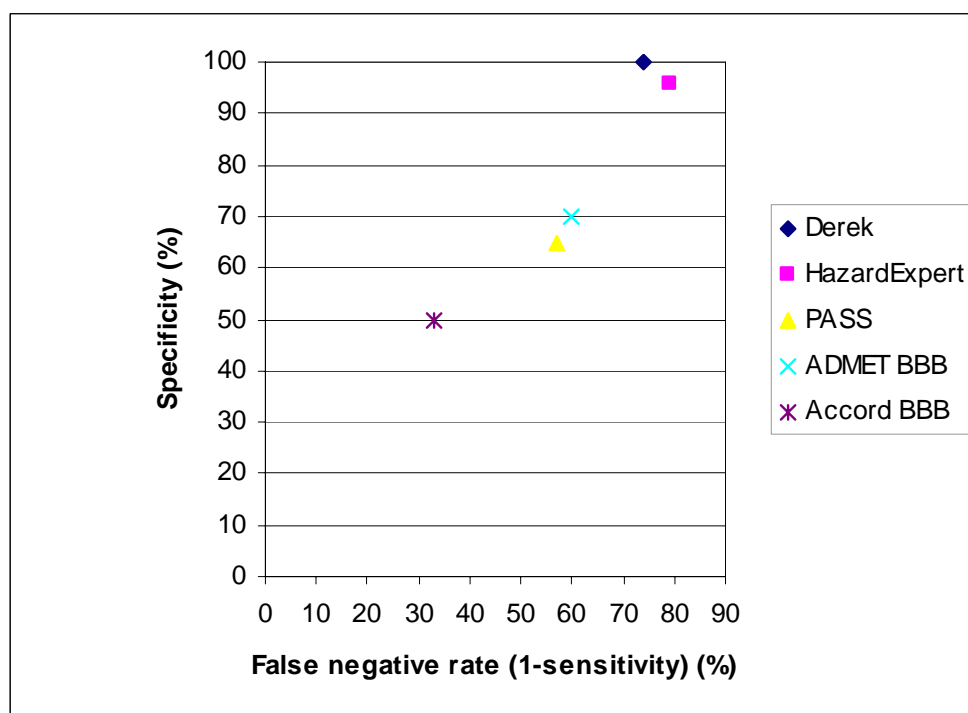


Figure 6.3 ROC curve for the neurotoxicity and blood brain barrier passage models (65 substances)

7. Chemical space analysis

As discussed in Section 3, the applicability domains of statistically-based models can be reconstructed if the training set is provided, by constructing the domain in terms of a specific set of structural features and/or molecular descriptors. Test chemicals can then be compared in terms of their similarity in this chemical space.

To carry out a chemical space analysis in this study, it was necessary to first identify which statistically-based models are transparent in terms of their underlying training sets (including availability of chemical structures and biological data).

For developmental toxicity, the only models that fulfil the transparency criterion are:

- CAESAR (training set of 292 substances taken from Arena et al (2004).
- Leadscope, which contains 27 QSAR models for structural dysmorphogenesis (4), visceral dysmorphogenesis (3), foetal survival (12) and foetal growth (8) based on data for more than 5700 substances. Among these, models for structural and visceral dysmorphogenesis and foetal survival were considered relevant, but not foetal growth. Furthermore, models for effects in the mouse were not considered relevant.

For neurotoxicity, the only models that fulfil the transparency criterion are the Leadscope models for pup behaviour, but these were not considered relevant, being based on results from developmental neurotoxicity tests.

7.1. Structural space analysis

The results of the Leadscope structural fragments analysis are provided in Appendix D. The analysis was applied to 6 datasets: the PPP list; the training sets for the CAESAR and Leadscope developmental toxicity model; the neurotoxicity test set, the extended developmental toxicity test set; and the ToxRefDB developmental toxicity dataset.

7.1.1. Comparison of the PPP inventory with the developmental test sets

In Appendix D, chemical classes that are not present in the PPP list or in one of the developmental test sets (EFSA, ToxRefDB) are highlighted in red. This can be used, for example, to identify chemical classes of pesticides that are not covered in the EFSA test set (e.g. carbohydrates).

7.1.2. Comparison of the EFSA developmental test set with the model training sets

It was found that 14 classes in the extended developmental test set are missing from the CAESAR training set (Table 7.1), and two are missing from the Leadscope training set (Table 7.2). False negative predictions for these chemical classes could indicate that they represent important structural features not included in the corresponding model algorithms. However, the interpretation is not straightforward, as illustrated in Table 7.1 for CAESAR and Table 7.2 for Leadscope.

7.1.3. Comparison of the PPP inventory with the neurotoxicity test set

In Appendix D, chemical classes that are not present in the neurotoxicity test set list are highlighted in red. These include, for example, bases, carbohydrates, and several types of organic functional groups.

7.1.4. Comparison of the EFSA neurotoxicity test set with the model training sets

It was not possible to compare the EFSA test set with the neurotoxicity model training sets (HazardExpert and PASS) since the latter were unavailable.

7.2. Molecular descriptor space analysis

7.2.1. Comparison of the PPP inventory with the developmental test sets

Figures 7.1a-d show the comparison in terms of molecular descriptor space. The figures indicate that the test sets are largely representative of the PPP inventory, although the latter is more densely populated and diffuse.

7.2.2. Comparison of the EFSA developmental test set with the model training sets

The comparison in terms of molecular descriptor space is given in Figures 7.2a-d for the Leadscape model and in Figures 7.3a-d for the CAESAR model. The large degree of overlap indicates that the EFSA test set gives a fair coverage of the model training sets, except in the case of 6 test substances that do not overlap with the chemical space of the CAESAR model (Figure 7.3d). These six substances comprise abamectin and ammonium pentadecafluorooctanoate (both incorrectly predicted as negative), as well as octabromobiphenyl ether, C.I. direct blue 6, C.I. direct red 28 and C.I. direct black 38 (all correctly predicted as positive). These results reflect the fact that when a chemical is outside the applicability domain of a model, it does not necessarily mean that its predicted toxicity is wrong, but simply that the prediction cannot be made with as much confidence.

7.2.3. Comparison of the PPP inventory with the neurotoxicity test set

Figures 7.4a-d show the comparison in terms of molecular descriptor space. The figures indicate that the EFSA test set covers a fair range of the PPP inventory in terms of the partitioning (MLOGP) vs molecular size (Sv) space although there are some areas of the PPP where no test chemicals were selected (Figure 7.4b). In contrast, in terms of the reactivity (Mi) space, the EFSA test set is biased towards more reactive chemicals (lower Mi values), as indicated in Figures 7.4c and 7.4d.

7.2.4. Comparison of the EFSA neurotoxicity test set with the model training sets

It was not possible to compare the EFSA test set with the neurotoxicity model training sets (HazardExpert and PASS) since the latter were unavailable.

7.3. Conclusions on the chemical space analysis

Chemical space analysis can be used to explore and define the applicability domains of statistically-based models if their training sets (including structures and biological data) are available. It can also be used to inform model development by identifying areas of chemistry that are not adequately covered in existing models.

There are different approaches to building applicability domains (Netzeva et al, 2005), but two of the most commonly used approaches are based on structural fragments and molecular descriptors. Model applicability domains can be used to rationalise the predictions made for test set chemicals (for which the toxicological effects are known) and to help determine the reliability of prediction for untested chemicals. However, the interpretation is not straightforward. If a chemical is outside the applicability domain of a model, it does not necessarily mean that its predicted toxicity is wrong, but simply that the prediction cannot be made with as much confidence. Conversely, when a chemical is within the applicability domain, it does not necessarily follow that the predicted toxicity will be accurate, but simply that the prediction can be made with a defined level of confidence. Furthermore, there is no absolute definition of a model applicability domain – different interpretations may be useful for different purposes. Some software tools provide their own assessment of prediction reliability based on applicability domain considerations, whereas other software tools do not. In practice, the definition and interpretation of applicability domains is not a trivial exercise. In this study, only the CAESAR and Leadscape models were amenable to chemical space analysis. The results were not particularly informative in terms of understanding the reliability of prediction, which reinforces the view that such analyses should be regarded as indicative rather than conclusive.

In addition to exploring the applicability domains of QSAR models, chemical space analysis can be used to compare the test sets with the “universe” of pesticides, as represented by the Plant Protection Products (PPP) inventory. In this study, the developmental toxicity test sets were found to largely cover the space of the PPP inventory, while the (smaller) neurotoxicity test set was less diffuse and chemically diverse. In any future efforts to build a more extended neurotoxicity test set as a means of providing a more comprehensive challenge to available models, it would be useful to search for chemicals in these areas of the PPP. However, there is no guarantee that reference chemicals with adequate data will be found.

Table 7.1 Chemical classes in the EFSA test set not present in the CAESAR training set

Chemical class	Predictions	Substances
isocyanate	1 positive correctly predicted	Methyl isocyanate
nitrile	5 positive correctly predicted 4 negative wrongly predicted as positive	Benzonitrile, chlorothalonil, Myclobutanil, ioxynil (ISO) and its salts, bromoxynil (ISO) and its salts Fludioxonil, azoxystrobin, Metaflumizone, flonicamid
1,2,4-triazolidine	2 positives correctly predicted	Azafenidin, prothioconazole
1,2,4-triazine(H)	1 negative wrongly predicted as positive	Metamitron
1,2,4-triazine	1 negative wrongly predicted as positive	Triazoxide
1,2,4,5-tetrazine	1 negative wrongly predicted as positive	Clofentezine
1,3,4-oxadiazole	1 positive wrongly predicted as negative	Oxadiargyl
rings size 4-7 O+S*	1 negative wrongly predicted as positive	Carboxin
oxazole	1 positive correctly predicted	Fenoxaprop-P
isoindole, 1,3-dioxo	1 negative wrongly predicted as positive	Phosmet
1,3-benzoxazole	1 positive correctly predicted	Fenoxaprop-P
1,4-benzoxazine	1 positive correctly predicted	Flumioxazin
benzimidazole	1 positive correctly predicted 1 positive wrongly predicted as negative	Benomyl Carbendazim
sulfone	1 positive correctly predicted 1 negative wrongly predicted as positive	Isoxaflutole Rimsulfuron

*This means that a cyclic structure is present, containing 4-7 atoms in the ring, one of which is O and another S.

Table 7.2 Chemical classes in the EFSA test set not present in the Leadscape training set

Chemical class	Predictions	CAS Numbers
1,2,4,5-tetrazine	1 negative correctly predicted	Clofentezine
1,3-benzoxazole	1 positive wrongly predicted as negative	Fenoxaprop-P

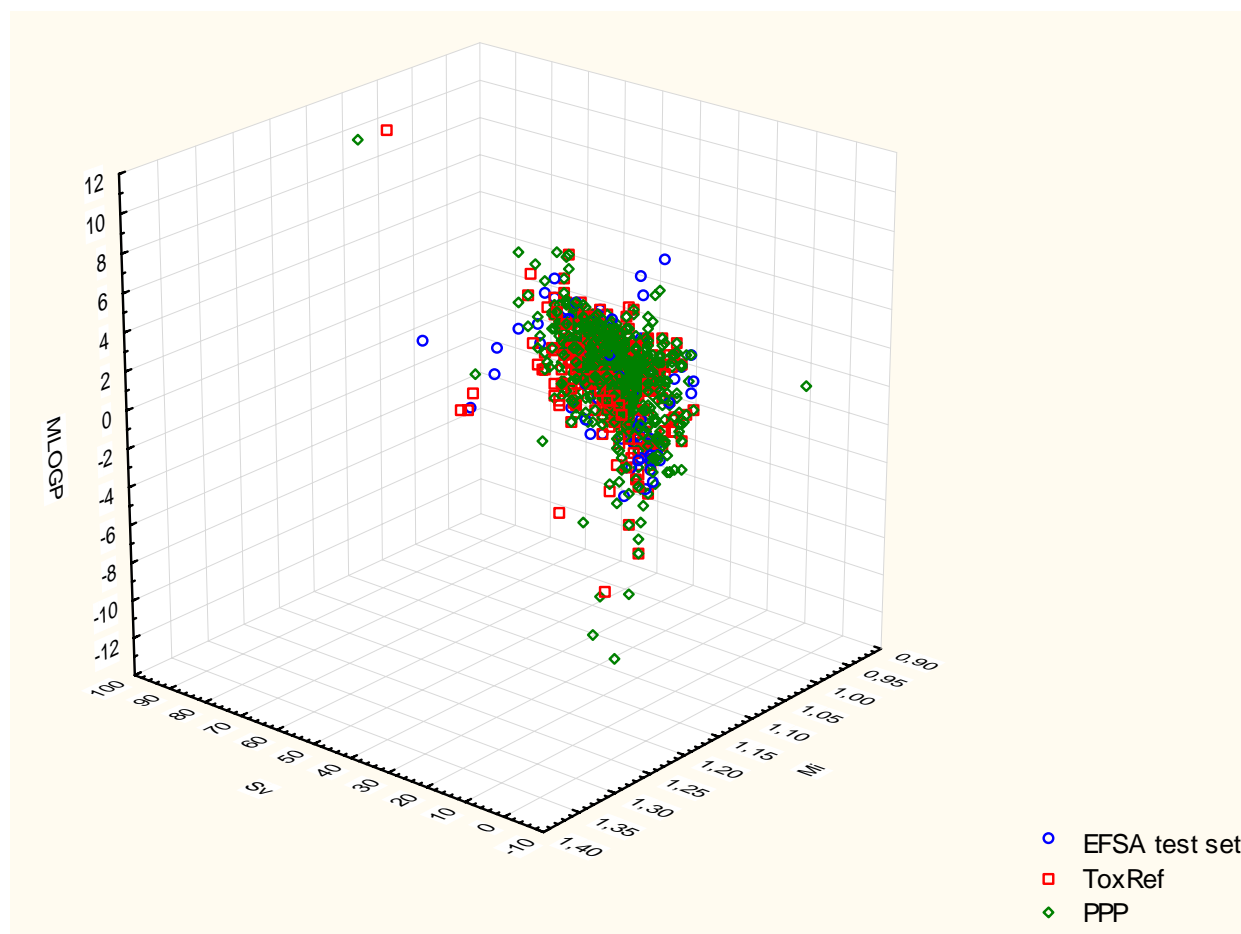


Figure 7.1a Comparison of the PPP inventory, EFSA developmental and ToxRef datasets in terms of molecular descriptor space

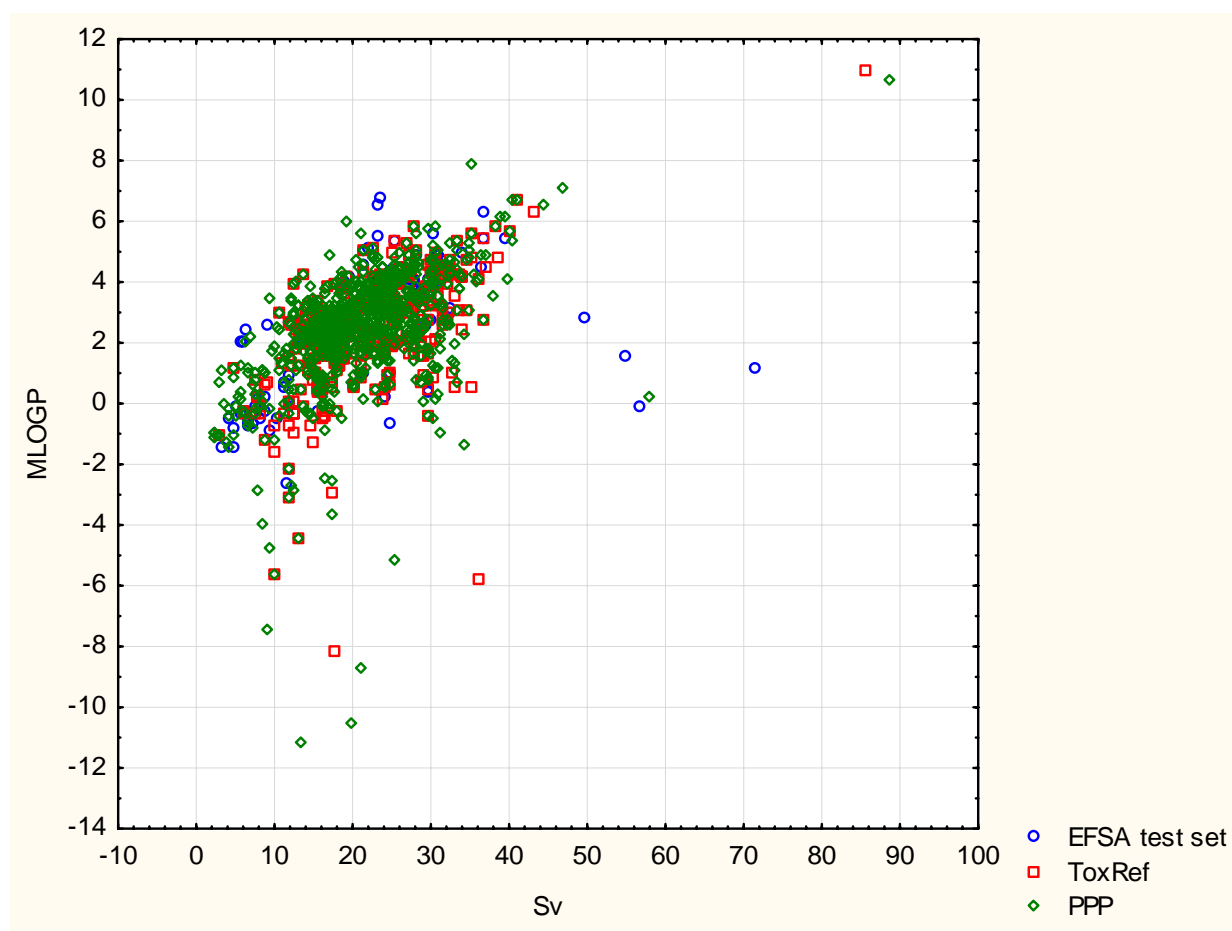


Figure 7.1b Comparison of the PPP inventory, EFSA developmental and ToxRef datasets in terms of molecular descriptor space

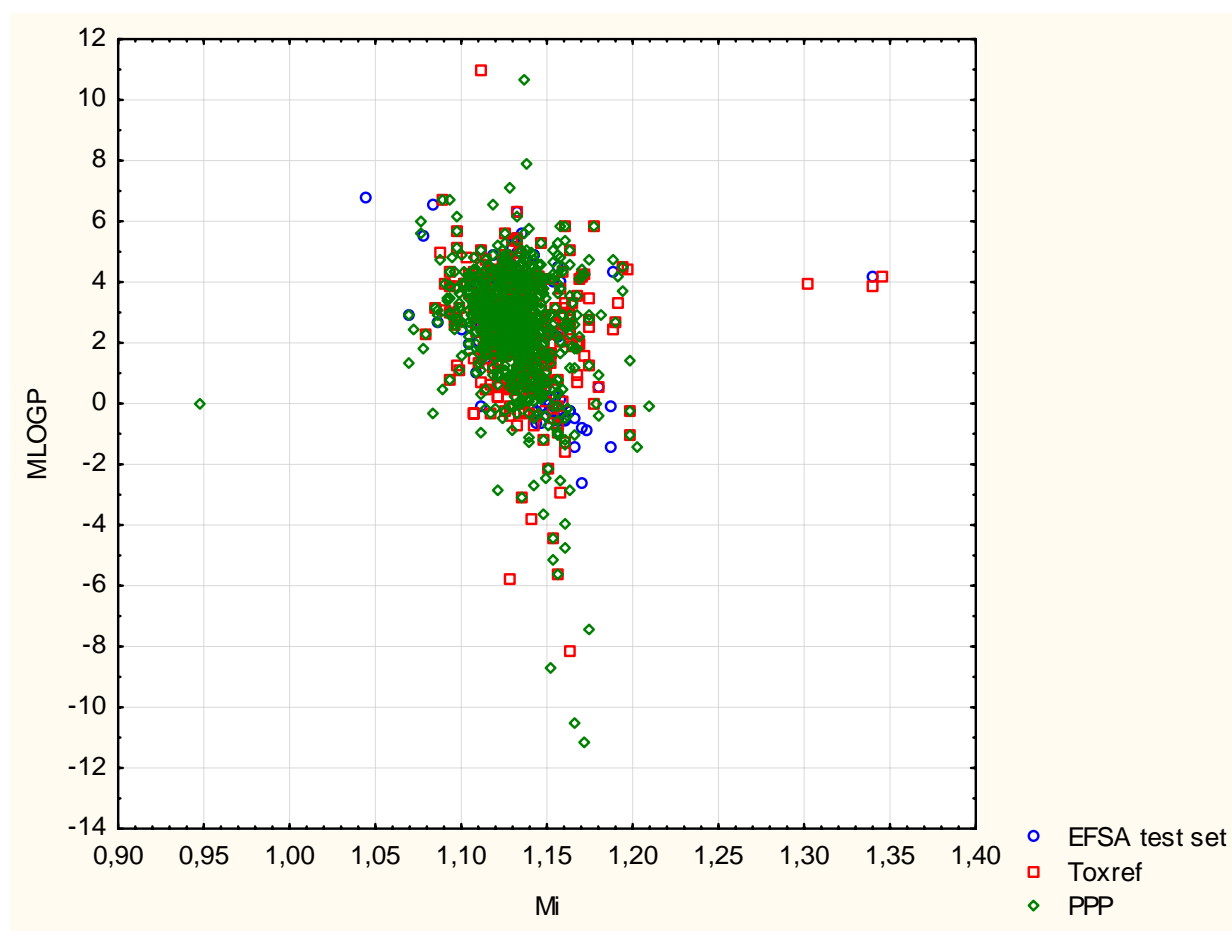


Figure 7.1c Comparison of the PPP inventory, EFSA developmental and ToxRef datasets in terms of molecular descriptor space

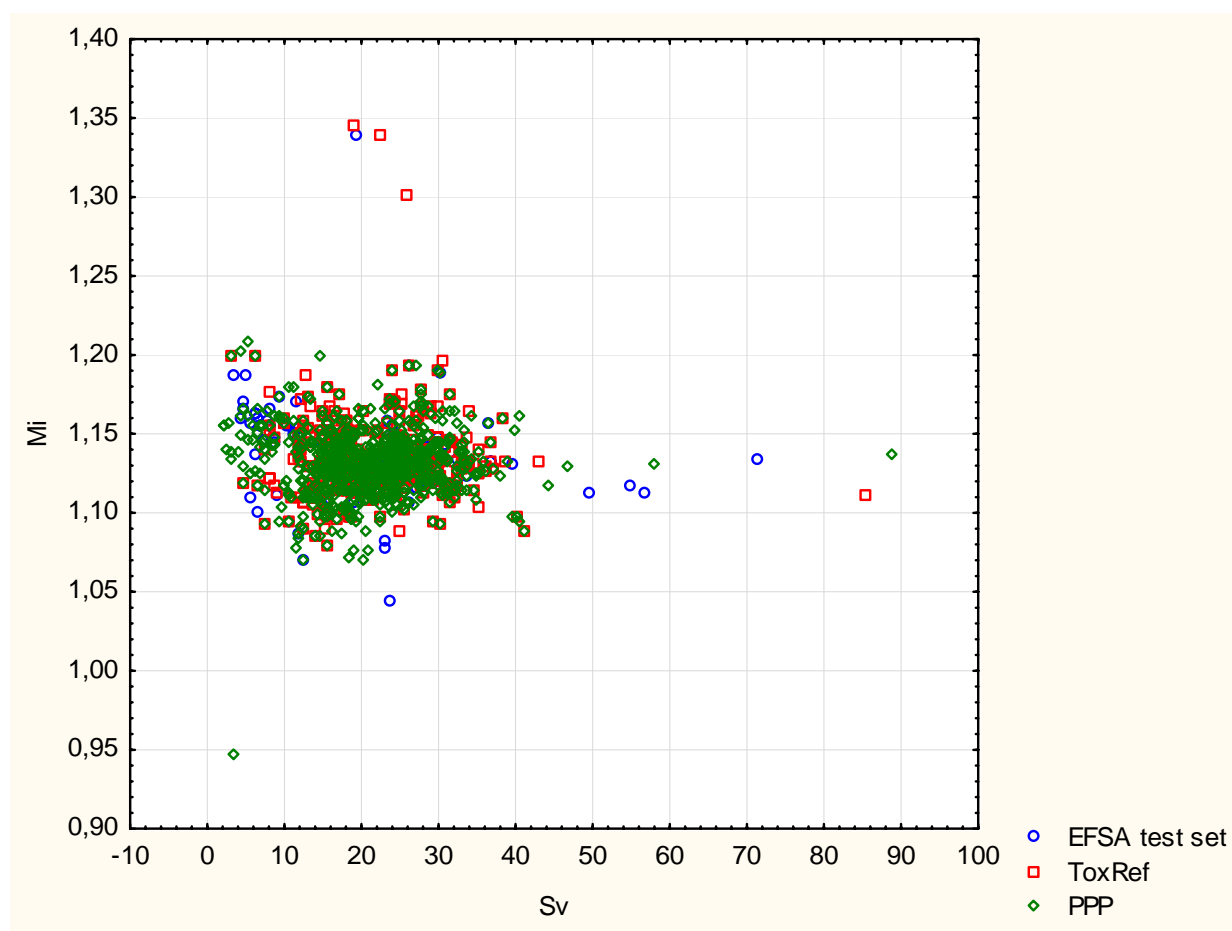


Figure 7.1d Comparison of the PPP inventory, EFSA developmental and ToxRef datasets in terms of molecular descriptor space

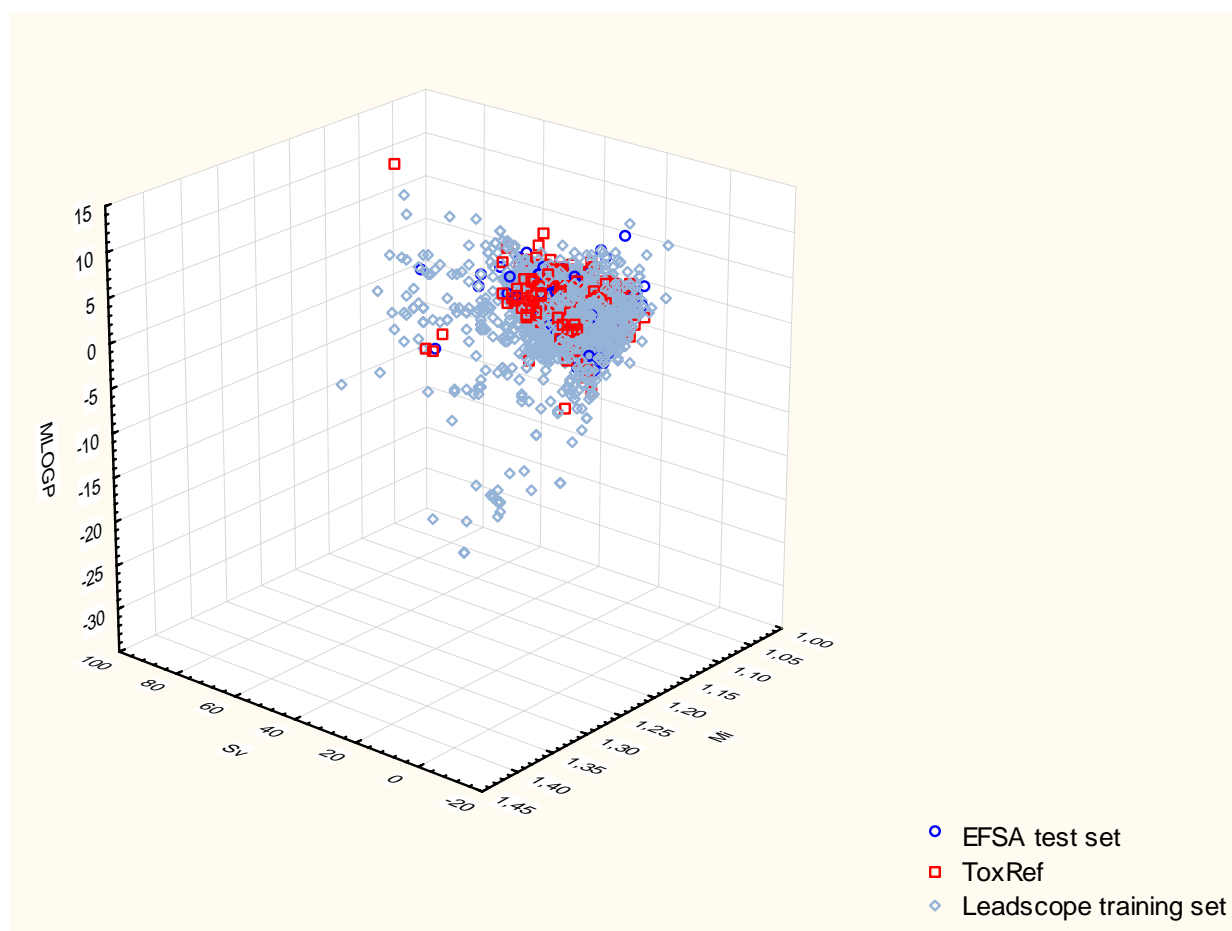


Figure 7.2a Comparison of the EFSA developmental and ToxRef test sets with the Leadscope training set

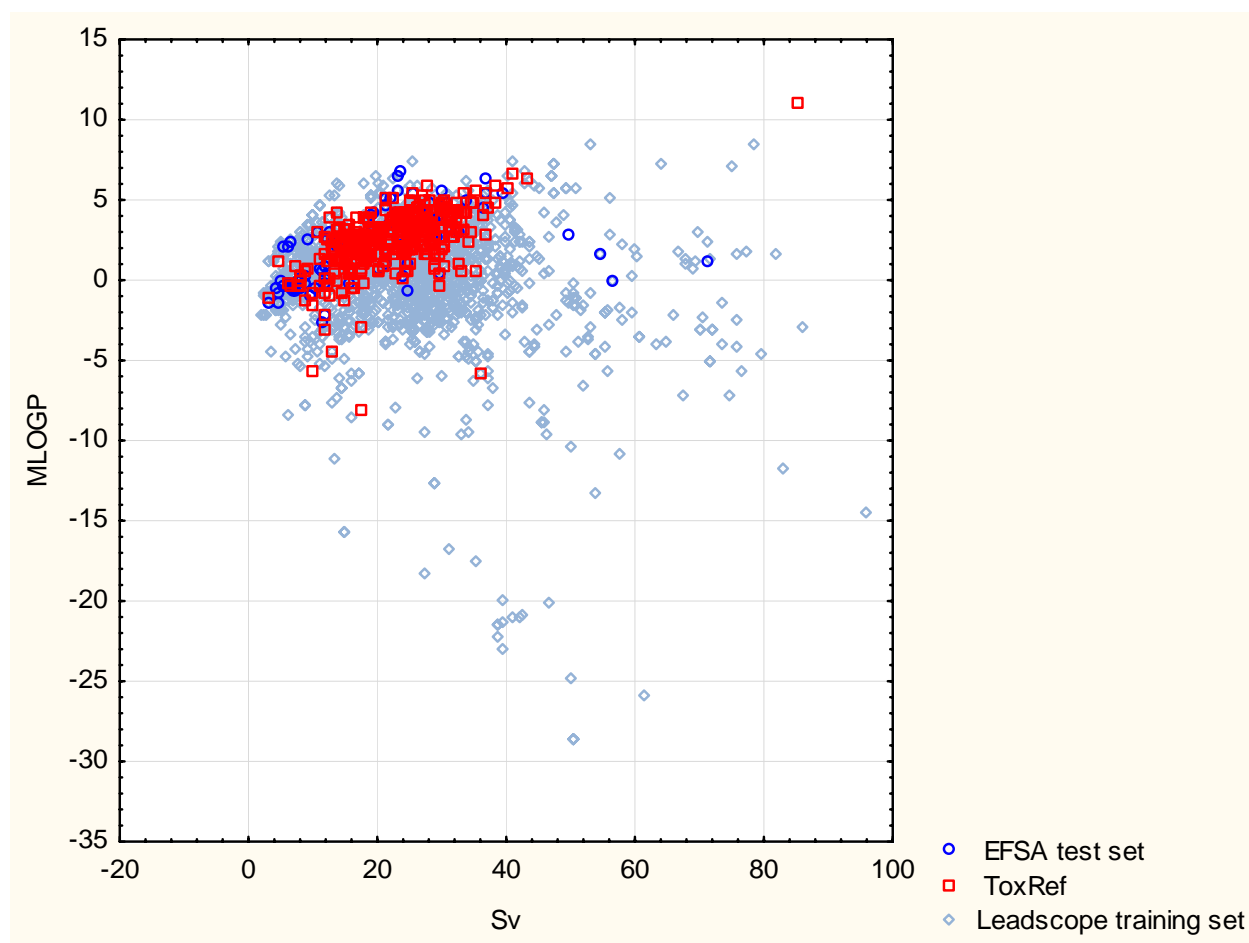


Figure 7.2b Comparison of the EFSA developmental and ToxRef test sets with the Leadscape training set

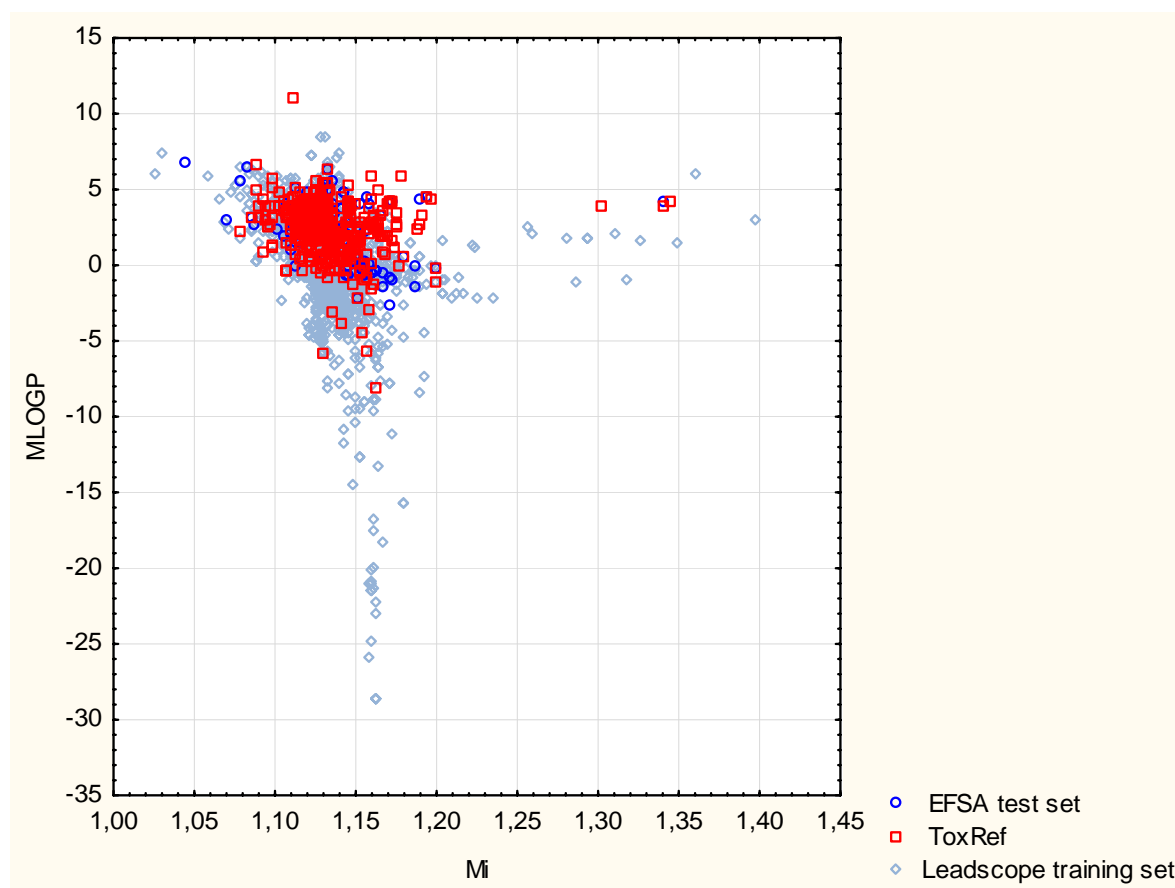


Figure 7.2c Comparison of the EFSA developmental and ToxRef test sets with the Leadscope training set

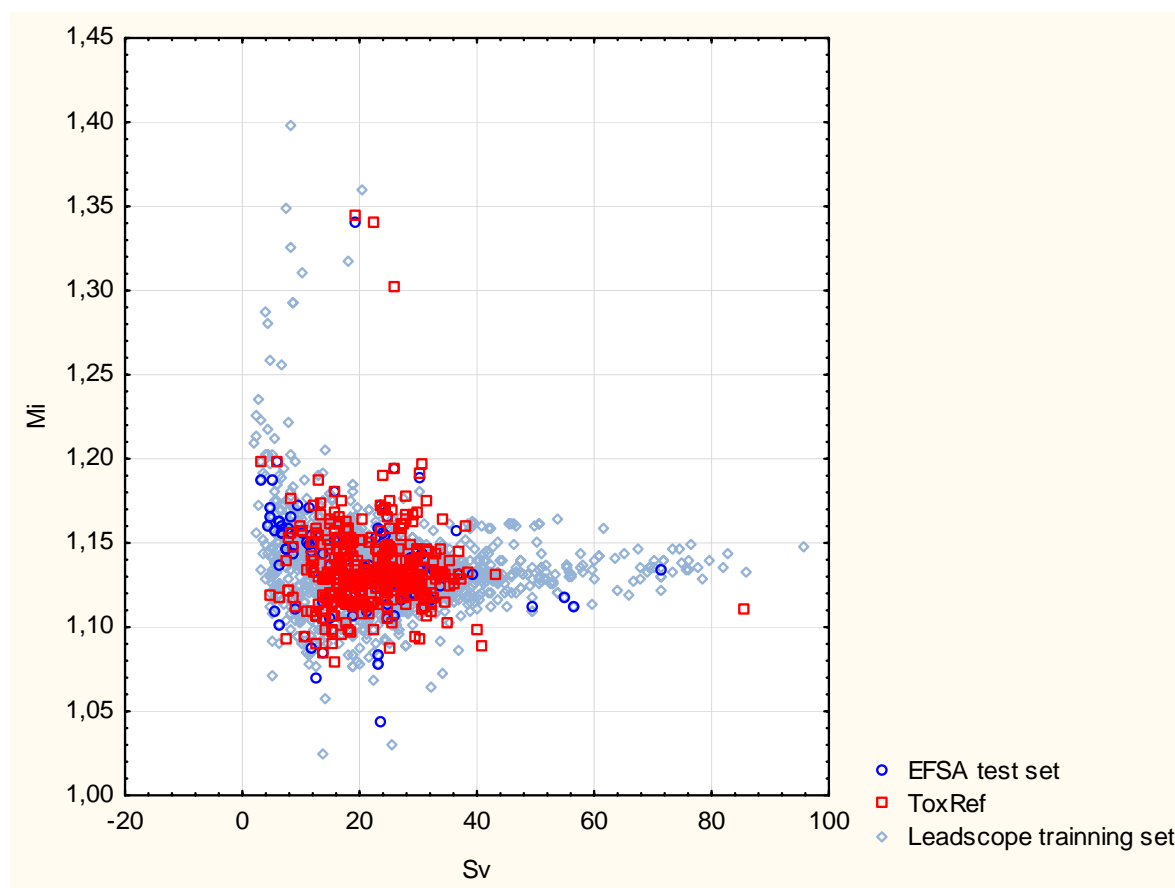


Figure 7.2d Comparison of the EFSA developmental and ToxRef test sets with the Leadscape training set

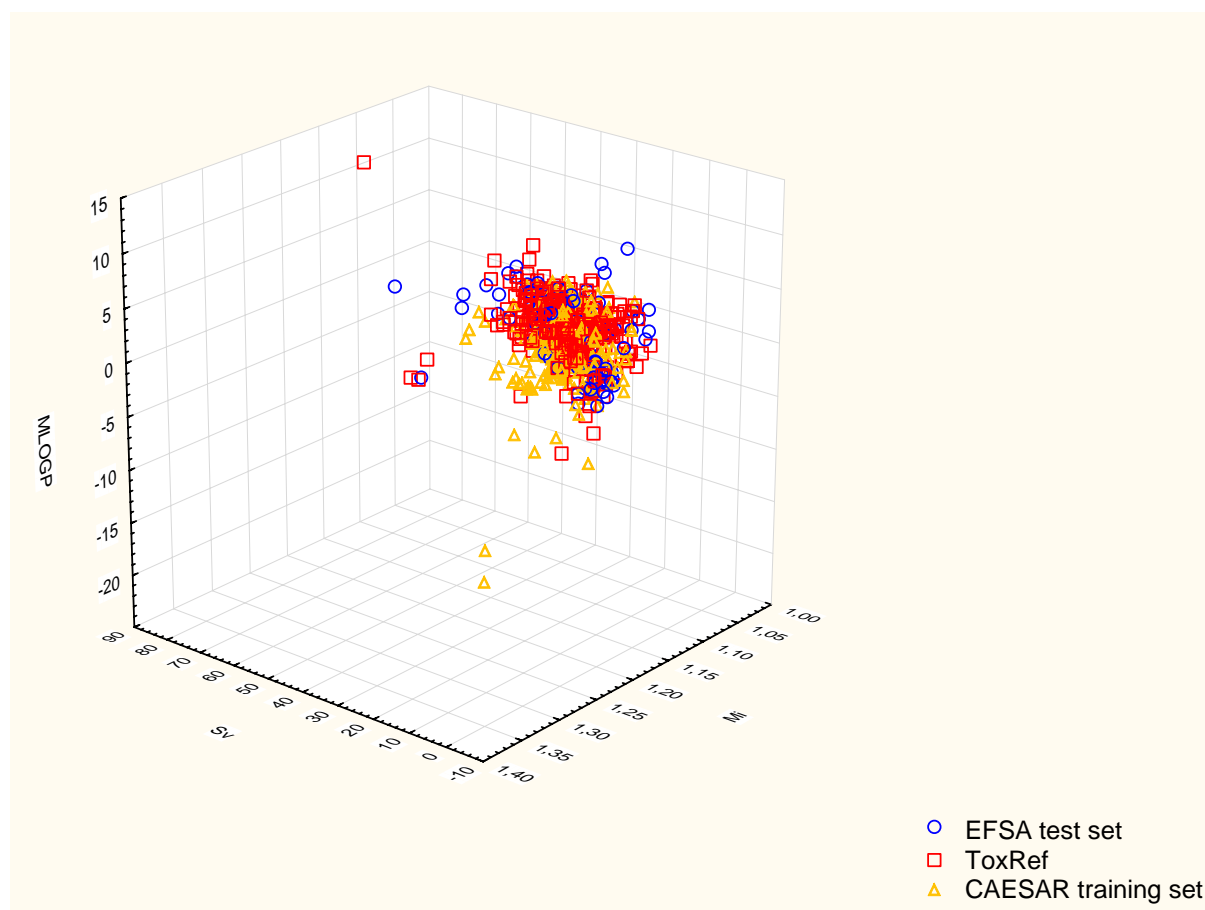


Figure 7.3a Comparison of the EFSA developmental and ToxRef test sets with the CAESAR training set

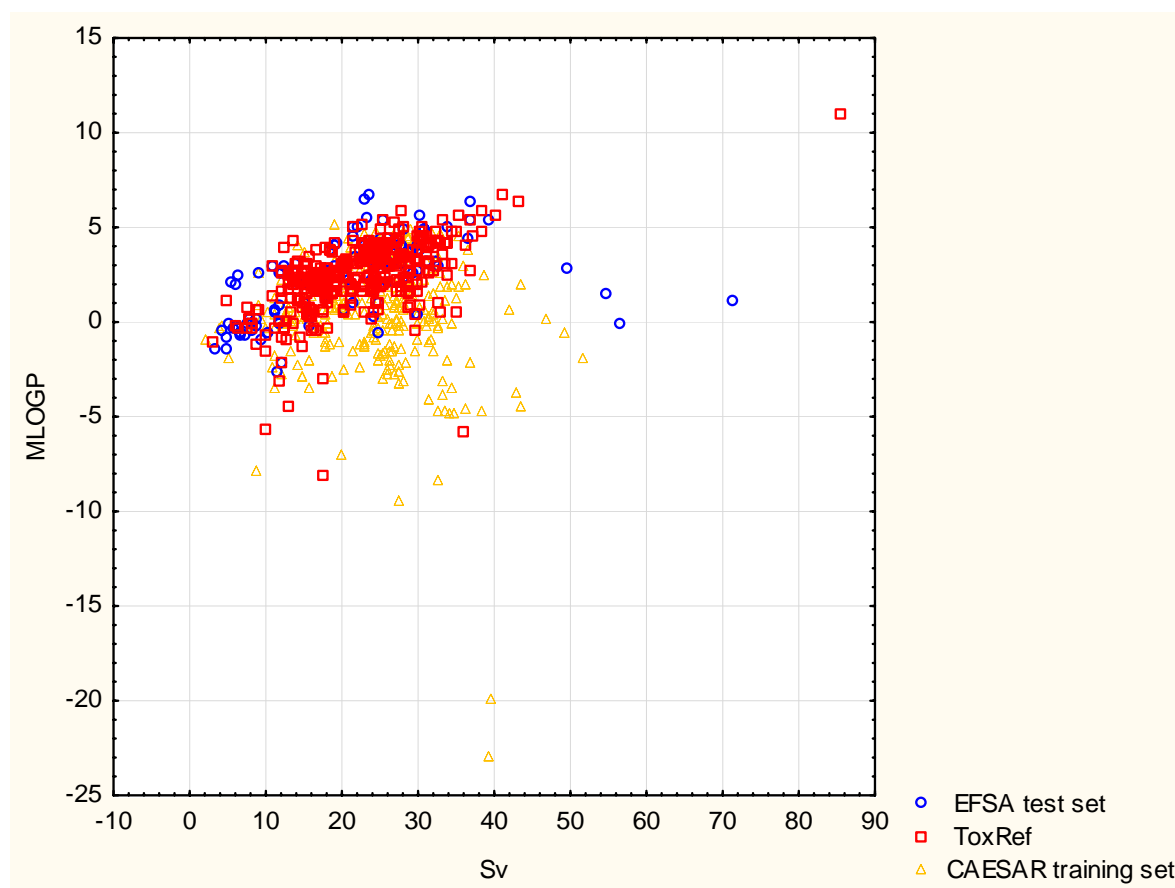


Figure 7.3b Comparison of the EFSA developmental and ToxRef test sets with the CAESAR training set

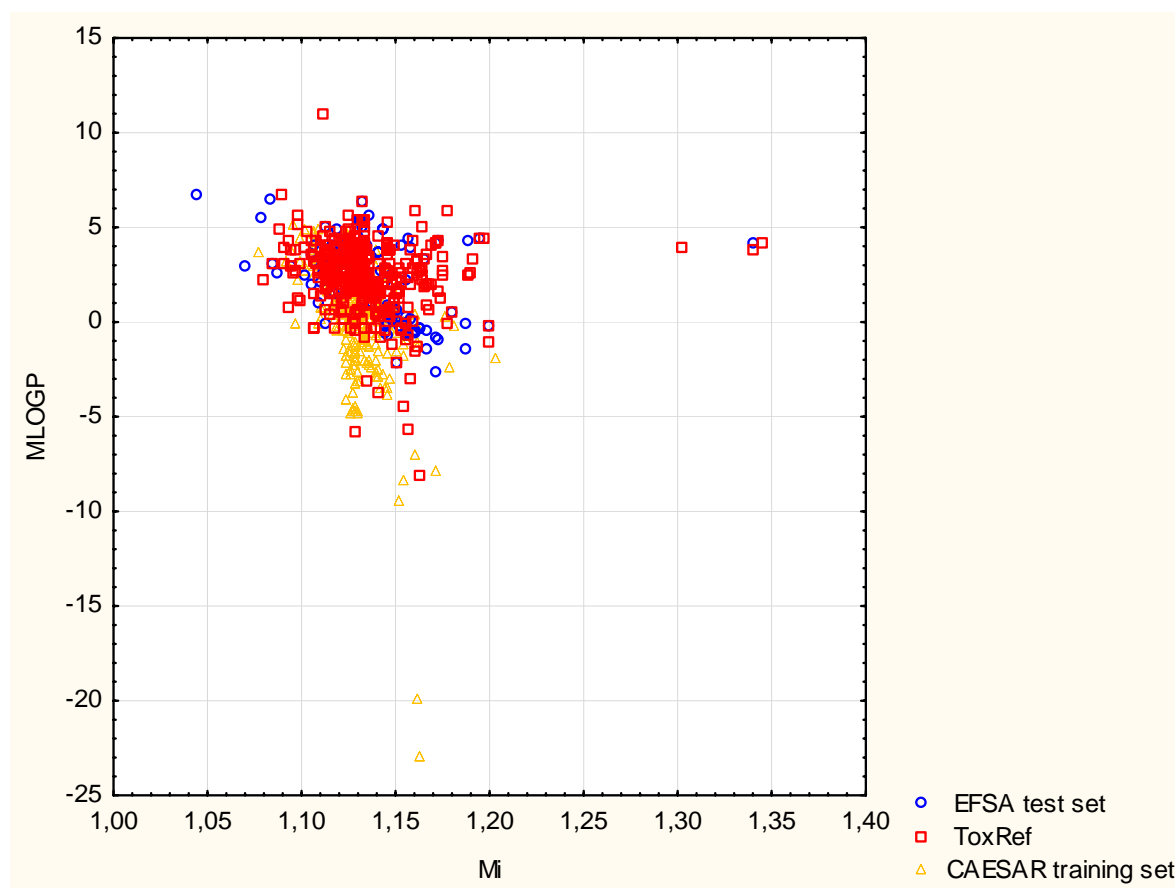
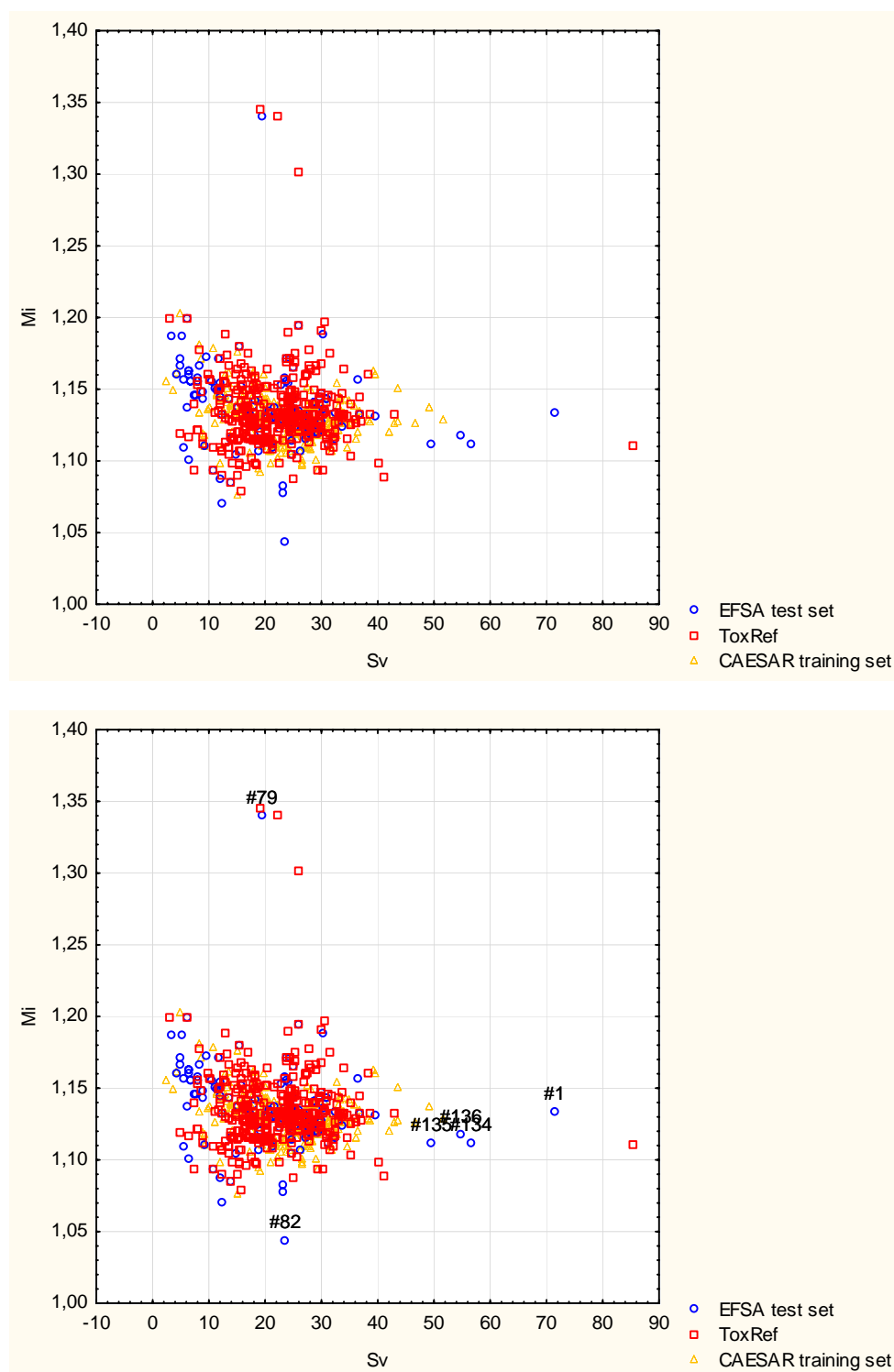


Figure 7.3c Comparison of the EFSA developmental and ToxRef test sets with the CAESAR training set



#1 Abamectin; #79 Ammonium pentadecafluorooctanoate; #82 octabromobiphenyl ether; #134 C.I. direct blue 6; #135 C.I. direct red 28; #136 C.I. direct black 38

Figure 7.3d Comparison of the EFSA developmental and ToxRef test sets with the CAESAR training set

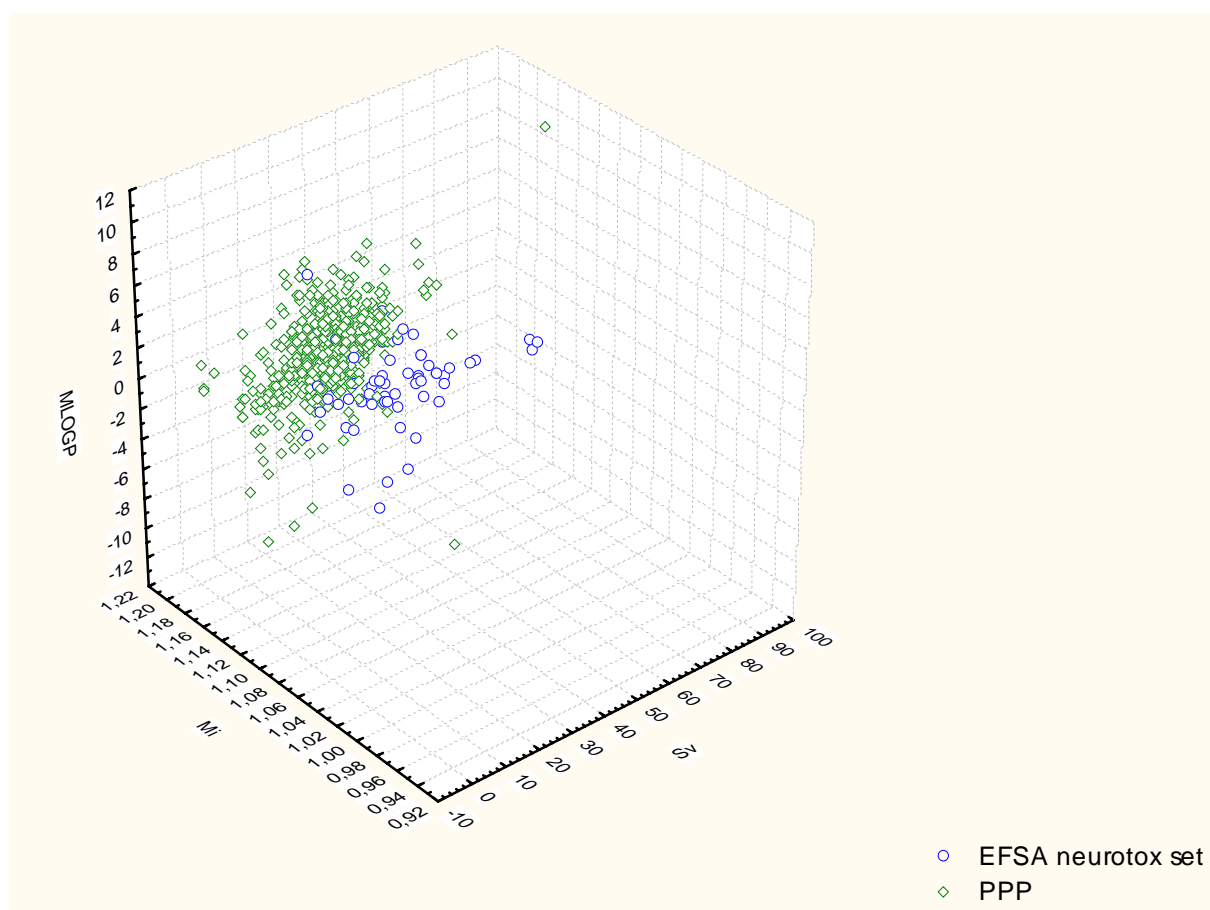


Figure 7.4a Comparison of the PPP inventory and EFSA neurotoxicity dataset in terms of molecular descriptor space

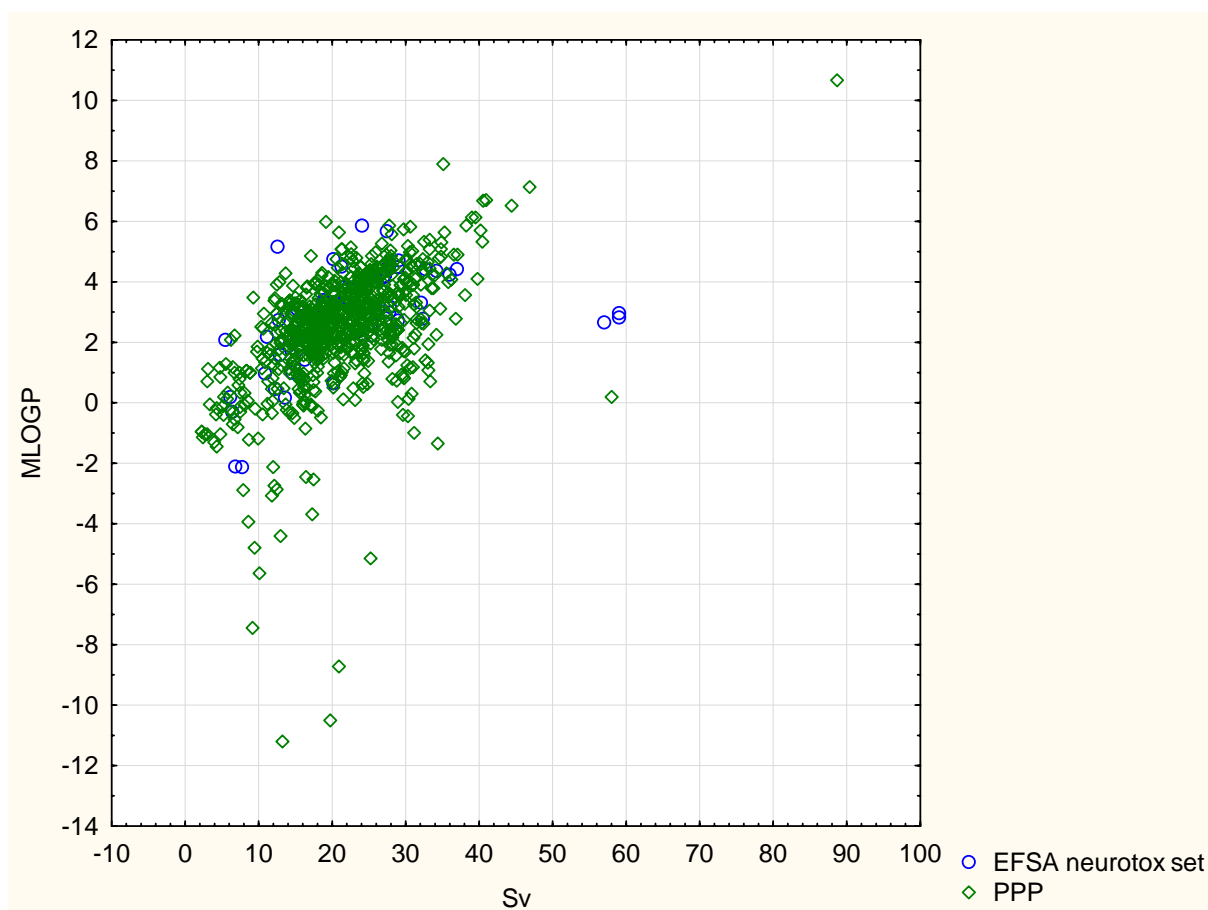


Figure 7.4b Comparison of the PPP inventory and EFSA neurotoxicity dataset in terms of molecular descriptor space

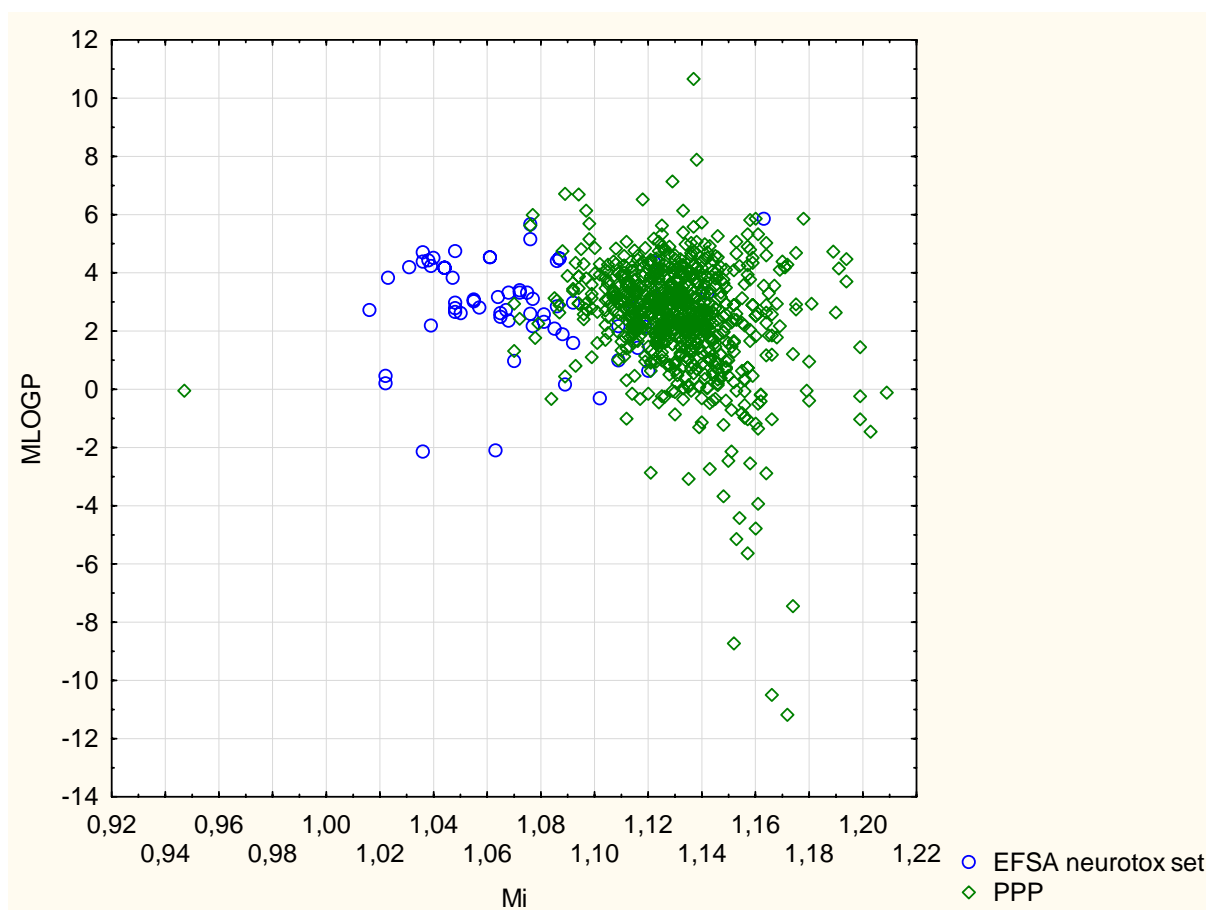


Figure 7.4c Comparison of the PPP inventory and EFSA neurotoxicity dataset in terms of molecular descriptor space

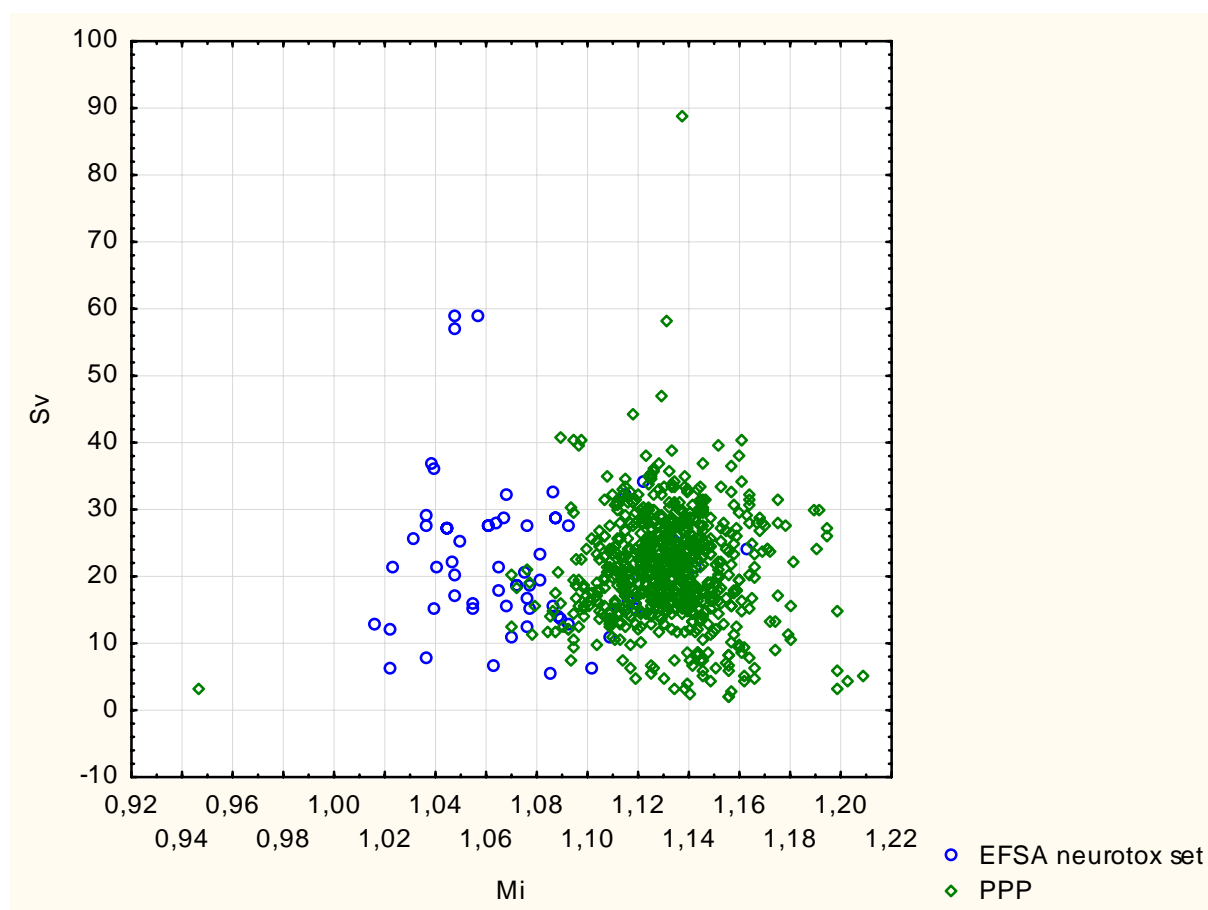


Figure 7.4d Comparison of the PPP inventory and EFSA neurotoxicity dataset in terms of molecular descriptor space

8. Case studies on the read-across assessment of developmental toxicity

The grouping of chemicals and the application of read-across with groups can have three outcomes: the chemical can be predicted to be positive or negative, or the read-across can be inconclusive, either because there are sufficient high quality data on analogues of the chemical of interest, or because the analogues have contradictory results (some are positive, while others are negative). The following sections illustrate each possible outcome, when read-across was carried out in the OECD QSAR Toolbox (v. 2.1.0.721) using the organic functional group (nested) profiler, developed by the Bourgas University (Bulgaria), in order to identify similar chemicals from the ToxRefDB.

In the absence of endpoint-specific similarity metrics and profilers based on a mechanistic understanding of the toxicological endpoint, it is rational to use an organic functional group profiler since functional groups are specific groups of atoms within molecules that are responsible for their chemical reactions. The same functional group will undergo the same or similar chemical reaction(s) regardless of the size of the molecule. However, its relative reactivity can be modified by nearby functional groups. The OECD/Bourgas University profiler includes 227 organic functional groups.

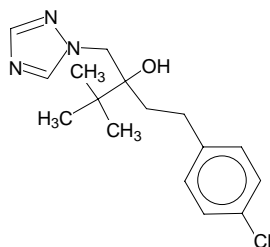
It should be remembered that grouping and read-across is not an automatic procedure like QSAR prediction, since the user needs to make a number of expert choices, and differences in such choices may lead to different conclusions based on the analogue data obtained. This exercise is therefore intended to be illustrative, since other tools could also be applied for grouping, and databases other ToxRefDB might also be available and suitable to find data-rich analogues. For example, Enoch et al (2009) used Toxmatch (<http://ecb.jrc.ec.europa.eu/qsar/qsar-tools/>) to read-across teratogenicity classifications from the dataset of Briggs et al (2002). The choice of database is important – it should be structurally close to the substances of interest and have high quality experimental data. For the purpose of this exercise, the US EPA's Toxicity Reference Database (ToxRefDB) was used. This contains information on chronic developmental and reproductive toxicity, and on carcinogenicity, in the rat and mouse, for more than 300 pesticides. The data are considered relevant and reliable since they are based EPA guideline toxicity studies (information on route and exposure time are not available).

8.1. Read-across prediction of positive (development toxicant)

Substance: Tebuconazole, CAS 107534-96-3

Conclusion based on experimental data: developmental toxicant

Structure:



Application of OECD Toolbox organic functional groups profiler: the substance contains the following organic functional groups: alcohol, alkane branched with quaternary carbon, aryl halide, triazole (substituted).

Identification of analogues from ToxRefDB: Using the above-mentioned functional groups to define the similarity criteria, four similar substances were selected (Figure 8.1).

CAS 55219-65-3	CAS 83657-24-3	CAS 76738-62-0	CAS 83657-17-4
Triadimenol	Diniconazole	Paclobutrazol	Uniconazole-P
Positive	Positive	Positive	Negative

Figure 8.1 Analogues of Tebuconazole identified from ToxRefDB

Three of the analogues are developmental toxicants whereas one, uniconazole-P (CAS 83657-17-4), is negative. Examination of the ToxRef data reveal that the three positive substances do not cause developmental toxicity in the rabbit, but cause developmental toxicity in the rat at concentrations equal to the maternal toxicity (mLEL=dLEL=25mg/kg/day).

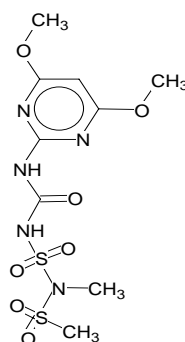
Conclusion based on read-across: tebuconazole is likely to be positive (developmental toxicant), which is in agreement with the toxicological evaluation based on experimental data for this substance.

8.2. Read-across prediction of negative (non-development toxicant)

Substance: Amidosulfuron, CAS 120923-37-7

Conclusion based on experimental data: non-developmental toxicant

Structure:



Application of OECD Toolbox organic functional groups profiler: the substance contains the following organic functional groups: ether, sulfonamide, sulfonyl urea.

Identification of analogues from ToxRefDB: Using the above-mentioned functional groups to define the similarity criteria resulted in the selection of 15 similar substances. To reduce this to a manageable number of analogues (more closely related to the query substance), all analogues containing elements not present in the query substance were excluded. The resulting chemical category contains 6 substances all are which are negative (Figure 8.2).

CAS 74223-64-6	CAS 97780-06-8	CAS 122931-48-0
Metsulfuron-methyl	Ethametsulfuron methyl	Rimsulfuron
Negative	Negative	Negative
CAS 173159-57-4	CAS 208465-21-8	CAS 141776-32-1
Foramsulfuron	Mesosulfuron-methyl	Sulfosulfuron
Negative	Negative	Negative

Figure 8.2 Analogues of amidosulfuron identified from ToxRefDB

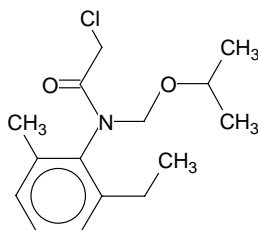
Conclusion based on read-across: amidosulfuron is likely to be negative (non-developmental toxicant), which is in agreement with the toxicological evaluation based on experimental data for this substance.

8.3. Read-across inconclusive due to contradictory evidence

Substance: Propisochlor, CAS 86763-47-5

Conclusion based on experimental data: non-developmental toxicant

Structure:



Application of OECD Toolbox organic functional groups profiler: the substance contains the following organic functional groups: alkyl halide, ether, haloacetamide.

Identification of analogues from ToxRefDB: Using the above-mentioned functional groups to define the similarity criteria, 6 similar substances were selected (Figure 8.3). Five of the analogues are negative but the substance most similar (acetochlor) to the query substance is positive. This analogue is most similar in the sense that it differs by a single methyl group (on the ethoxy side chain).

CAS 34256-82-1	CAS 15972-60-8	CAS 23184-66-9
Acetochlor	Alachlor	Butachlor
Positive	Negative	Negative
CAS 51218-45-2	CAS 87392-12-9	CAS 163515-14-8
Metolachlor	S-Metolachlor	Dimethenamide-P
Negative	Negative	Negative

Figure 8.3 Analogues of propisochlor identified from ToxRefDB

Conclusion: based on the weight-of-evidence (majority of analogues), propisochlor could be regarded as negative. More conservatively, in view of the fact that the most similar analogue is positive, the read-across could be regarded as inconclusive.

It is interesting to note that the only positive analogue identified in the ToxRefDB, acetochlor, is not considered positive for developmental toxicity according to the EFSA conclusion on this pesticide active substance (EFSA, 2011), which suggests a difference in the underlying toxicity data and/or interpretation.

8.4. Conclusions on grouping and read-across

Compared with QSAR models, the predictive performance of the grouping and read-across approach cannot be generalised as easily, since this is an ad hoc approach in which a number of subjective choices are made. For example, decisions have to be made concerning the choice of reference database(s) for analogue searching, the similarity criteria used to identify close analogues, and the assessment of the relevance and reliability of the analogue data, and the interpretation of positive, negative and inconclusive outcomes. Nevertheless, the read-across approach is equally well-suited to the identification of positive and negative chemicals, and for this reason, it is proposed in this study as a means of clarifying the negative predictions resulting from the application of QSAR.

Developmental toxicity and neurotoxicity are complex endpoints, which are only partially understood in mechanistic terms. In the absence of endpoint-specific profilers for analogue identification, analogues can be identified based on the presence of organic functional groups.

In this study, we have illustrated the usefulness of read-across in identifying developmental toxicants, non-developmental toxicants, and chemicals for which there is insufficient evidence to draw a conclusion on developmental toxicity. Only a few other studies have been published on the read-across of developmental toxicity, and these also support the applicability of grouping and read-across, especially when used in a weight-of-evidence approach along with QSAR (Enoch et al, 2009; Hewitt et al, 2010).

9. Use of QSAR and read-across in the assessment of toxicological relevance

In this section, the possible use of QSAR and read-across in the assessment of the toxicological relevance of metabolites and degradation products is discussed.

9.1. Stepwise assessment scheme based on combined use of QSAR and read-across

To illustrate a possible outcome of applying the stepwise assessment strategy (Figure 9.1), the following scenario was simulated:

- a) test set: the extended EFSA dataset of 135 chemicals (96 positives, 39 negatives)
- b) reference database for analogue searching: ToxRefDB
- c) QSAR: PASS teratogenicity model, since this has a good compromise between sensitivity (69%) and positive predictivity value (81%) and false positive rate (38%).
- d) Grouping and read-across: application of organic functional groups profiler in OECD Toolbox

The outcome, in terms of how the 135 chemicals proceed through the various steps, is summarised in Table 9.1. The predictive performance of the QSAR (alone), read-across (alone) and the combined (sequential) use of the two approaches is summarised in Table 9.2.

In step 1, all 135 chemicals are evaluated. However, 43 of these are removed from further evaluation since they are considered to have adequate toxicological data (in ToxRefDB) in order to decide on the appropriate exposure assessment strategy and complete the risk assessment. This step is included as a generally useful consideration, although it is recognised that, in practice, there will be adequate toxicological data for relatively few pesticide metabolites/degradates.

Thus, 92 chemicals proceed to step 2 in which the PASS teratogenicity model is used to identify positive chemicals. Of these 92 chemicals, 63 are predicted as positive and 29 as negative. Thus, only the 29 chemicals predicted as negative proceed to step 3, whereas the remaining 63 chemicals would need to be considered for further toxicity testing and exposure assessment, as appropriate.

In step 3, read-across is used to further evaluate the 29 chemicals predicted as negative by PASS. It turns out that predictions can only be made for 14 of chemicals (6 positives, 8 negatives; see Table 9.3 for results) since the outcome is inconclusive for 15 chemicals (due to lack of analogue data). This means that 15 of chemicals entering step 3 would need to be considered for further toxicity testing and exposure assessment, in accordance with EFSA guidance.

The overall effect of applying the stepwise strategy to 135 chemicals, according to the above-mentioned scenario can be summarised as follows:

- a) predictions are made for 77 chemicals, of which 68 are predicted positive and 9 negative;
- b) no predictions can be made for 58 chemicals, which either need to be tested, or evaluated using a different computational approach (different QSARs and/or different approach to read-across)

- c) 62 of the 68 chemicals predicted as positive are true positives, corresponding to a positive predictivity of 91% (compared with 81% when PASS is applied on its own).
- d) 8 of the 9 chemicals predicted as negative are true negatives, corresponding to a negative predictivity of 89% (compared with 44% when PASS is applied on its own).

9.2. Conclusions on the use of QSAR and read-across

The sequential use of QSAR for identifying positive chemicals (developmental toxicants) followed by the use of read-across for distinguishing between the true and false negative predictions generated by QSAR leads to an overall improvement in both positive and negative predictivity, which means there is a greater confidence in the ability to identify both developmental toxicants and non-developmental toxicants.

The scenario described above is intended purely for illustrative purposes. Other QSAR models, or model batteries, could be applied in Step 2, and an alternative approach to read-across could be applied in Step 3. It is also recognised that the ratio of positives to negatives entering Step 1 is unrealistically high – in practice, a much smaller percentage would be expected.

Further research is needed to further explore and optimise this strategy by: a) investigating the outcomes with different prevalences of positives (e.g. 1-10% of the test set); b) maximising the positive predictivity of the QSAR step (one or more models could be used in parallel) and; c) extending the applicability of the read-across step by developing a more extensive reference database, thereby minimising the percentage of inconclusive results.

Furthermore, while this exercise has focussed on the prediction of developmental toxicity, it is expected that the same general strategy would be effective for the prediction of neurotoxicity (since two models, Derek and HazardExpert had high positive predictivities of 96-100% for this endpoint). In order to evaluate this possibility, it would be necessary to develop a sufficiently large reference database for the read-across step.

It is expected that the use of non-testing methods will be used in combination with the Threshold of Toxicological Concern (TTC) approach (Barlow et al, 2005). EFSA's Scientific Committee is currently finalising a draft Opinion in which the relevance and reliability of the TTC approach is evaluated as a means of providing scientific advice about possible human health risks across the various areas of risk assessment under EFSA's remit. A key component of current TTC assessment strategies is the application of the Cramer classification tree (Cramer et al, 1978) that categorises chemicals into three structural groups according to the level of concern for oral systemic toxicity – Cramer classes I (low concern), II (intermediate concern) and III (high concern). The robustness of the Cramer scheme, and possibilities for improving it by using QSAR approaches, has been investigated in an EFSA-funded project (Bassan et al, 2011), as a contribution to the Scientific Committee's Opinion.

In parallel, EFSA's PPR Panel is drafting an opinion on approaches for the evaluation of the toxicological relevance of metabolites and degradates of pesticide active substances for Dietary Risk Assessment. In this context, a decision tree is being developed for the evaluation of the toxicological relevance of pesticide metabolites. This decision tree includes steps for short-term and long-term exposure assessment, as appropriate, including the hazard-based triggering of the short-term exposure assessment. In this context, an important question or the

practicality of the decision tree is whether it is also possible to waive the short-term exposure assessment, based on the predicted absence of short-term toxicological effects, such as developmental toxicity and neurotoxicity, in particular. The stepwise non-testing strategy proposed in this report is intended to feed into the decision tree being developed by the PPR Panel.

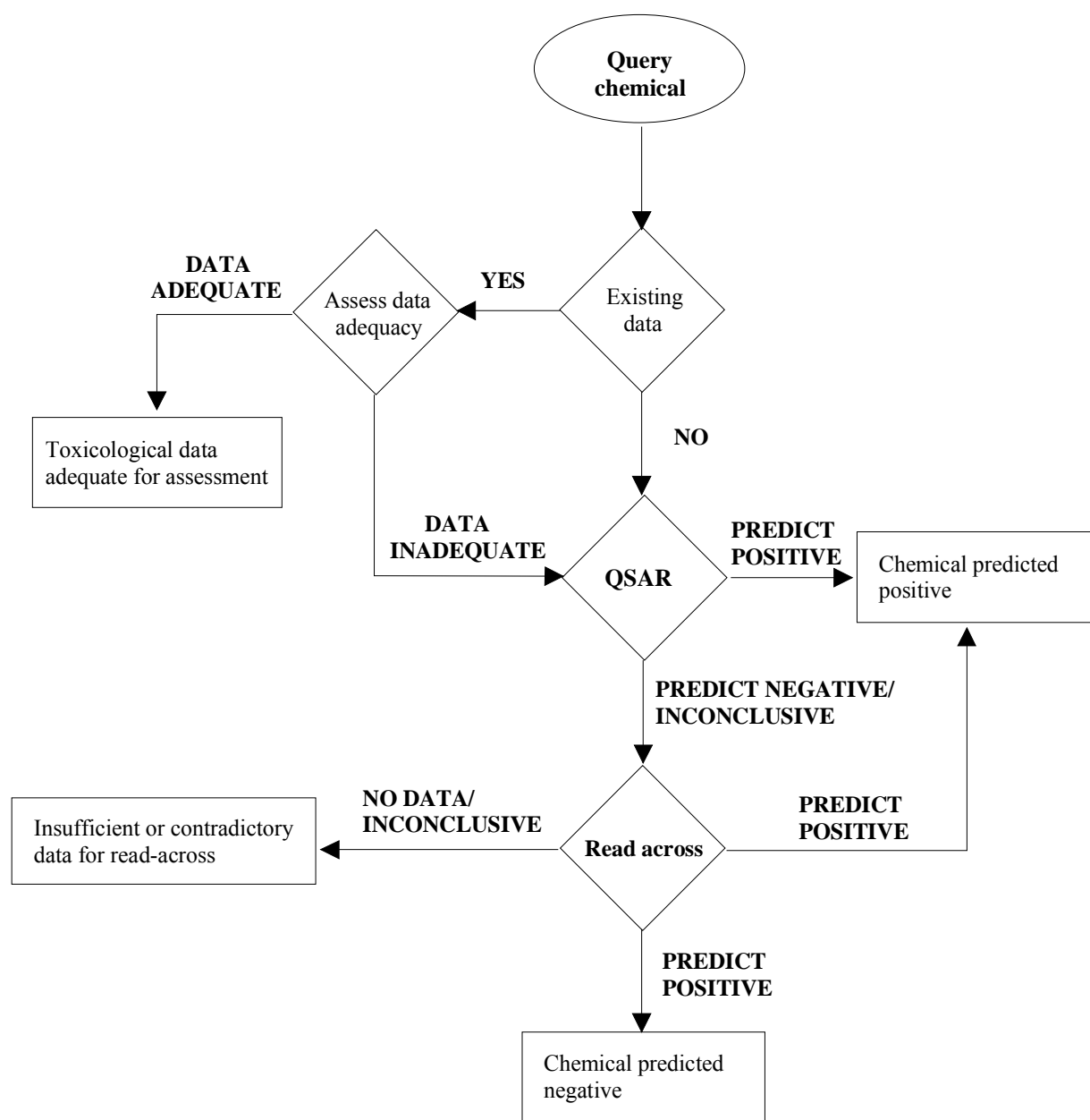


Figure 9.1 General stepwise assessment scheme based on the use of existing data and non-testing methods

Table 9.1 Possible outcome of applying the stepwise assessment strategy

Step	Entering	Predicted positive	Predicted negative	Not predicted	Filtered out	Proceeding to next step
1. Existing data	135 96P,39N			43	43 (adequate data)	92 72P,20N
2. QSAR model (PASS teratogenicity)	92	63	29	0	63	29
3. Read-across (OECD Toolbox)	29	5	9	15	15 (no data)	
Totals		68 62TP,6FP	9 8TN, 1FN	58		

Table 9.2 Performance of the QSAR and read-across steps when applied on their own and in sequence

9.2a) Performance of the QSAR (PASS teratogenicity) model

Total number of substances 9 (72P, 20N)		Predicted (QSAR)	
		positive	negative
Experimental data	positive	57	15
	negative	6	14

9.2b) Performance of the read-across

Total number of substances 14 (6P, 8N)		Predicted (read-across)	
		positive	negative
Experimental data	positive	5	1
	negative	0	8

9.2c) Performance of the sequential use of QSAR and read-across

Total number of substances 63 (51P, 12N)		Predicted (QSAR+read-across)	
		positive	negative
Experimental data	positive	62	1
	negative	6	8

Table 9.3 Read-across predictions of developmental toxicity for 14 pesticides from the Extended EFSA Dataset

CAS	Name	Conclusion based on experimental data	Organic functional groups identified by OECD Toolbox profiler	Similar compounds	Conclusion based on read-across
55179-31-2	Bitertanol	positive	alcohol, alkene branched with quaternary carbon, triazole (substituted)	3 positive : Triadimenol, Diniconazole, Paclobutrazol 1 negative : Uniconazole-P (in ToxRef shows some activity)	positive
133855-98-8	Epoxiconazole	positive	aryl halide, cycloalkane, triazole (substituted)	4 negative : Propiconazole, Bromuconazole, Difenconazole, Triticonazole	negative
120983-64-4	Metabolite Desthio-prothiconazole	positive	alcohol, aryl halides, cycloalkane, triazole (substituted)	1 positive : Cyproconazole (very similar)	positive
107534-96-3	Tebuconazole	positive	alcohol, alkene branched with quaternary carbon, triazole (substituted), aryl halide	3 positive : Triadimenol, Diniconazole, Paclobutrazol 1 negative : Uniconazole-P (in ToxRef shows some activity)	positive
125116-23-6	Metconazole	positive	alcohol, aryl halide, cycloalkane, triazole (substituted)	1 positive : Cyproconazole	positive
66246-88-6	Penconazole	positive	aryl halides, triazole (substituted) AND structural similarity more than 80%	4 positive : Hexaconazole, Diniconazole, Cyproconazole, Paclobutrazol 2 negative : Bromuconazole, Uniconazole-P (in ToxRef show some activity)	positive
120923-37-7	Amidosulfuron	negative	ether, sulfonamide, sulfonylurea AND chemical elements	6 negative : Metsulfuron-methyl, Ethametsulfuron methyl, Rimsulfuron, Foramsulfuron, Mesosulfuron-methyl, Sulfosulfuron	negative
98243-83-5	Benalaxyl-M	negative	carboxamide, carboxyl acid ester	4 negative : Metalaxyl, D-Alanine, N-(2,6-	negative

CAS	Name	Conclusion based on experimental data	Organic functional groups identified by OECD Toolbox profiler	Similar compounds	Conclusion based on read-across
				dimethylphenyl)-N-(methoxyacetyl)-, methyl ester, Triazamate, Carfentrazone-ethyl	
120928-09-8	Fenazaquin	negative	arene, ether AND chemical elements	7 negative: Propoxur, Isoxaben, Ethoxyquin, Fenoxycarb, Ethofenprox, Pyriproxyfen, Azoxystrobin	negative
83121-18-0	Teflubenzuron	negative	arene, aryl halide, imide, urea (substituted)	4 negative: Flucycloxuron, Novaluron, Butafenacil, Noviflumuron	negative
35367-38-5	Diflubenzuron	negative	arene, aryl halide, imide, urea (substituted)	4 negative: Flucycloxuron, Novaluron, Butafenacil, Noviflumuron	negative
103055-07-8	Lufenuron	negative	alkyl halide, arene, aryl halide, ether, imide, urea (substituted)	2 negative: Novaluron, Noviflumuron	negative
66063-05-6	Pencycuron	negative	aryl halide, cycloalkane, urea (substituted)	2 negative: Flucycloxuron, Hexythiazox	negative
111991-09-4	Nicosulfuron	negative	Pyridine (substituted), sulfonamide, sulfonylurea AND chemical elements	2 negative: Rimsulfuron, Sulfosulfuron	negative

Conclusions and Recommendations

CONCLUSIONS

10. Availability of QSAR tools and databases

A limited range of software tools and databases were identified as potentially useful for developmental toxicity and neurotoxicity prediction. At present, there are more tools for predicting developmental toxicity than neurotoxicity. In the case of developmental toxicity, the US EPA's ToxRef Database is a potentially useful reference database for the development of new models and the application of grouping and read-across. In the case of neurotoxicity, there is a lack of freely available QSAR tools, as well as a suitable public reference database for the development of new models and the application of grouping and read-across. A general caveat, which needs to be considered irrespective of the endpoint being predicted and the QSAR or reference database used, is that different regulatory bodies may apply different criteria in the evaluation of raw data. In order to take such differences into account, it is important that the conclusions (positive or negative toxicity) are accompanied by a description of the underlying effects at the organ, tissue and cellular levels. The ideal situation to meet EFSA's needs in predictive toxicology would be to develop an in-house database of relevant evaluated data, in which the conclusions are based on guideline criteria, and are linked to the underlying findings in the original study reports.

When compiling and searching chemical databases, the stereochemistry of molecules can be encoded into their SMILES strings. This could, in principle, be important in the prediction of developmental toxicity and neurotoxicity. Such information can in principle be encoded into QSAR models and structural alert-based rulebases. For example, the Derek knowledgebase includes several alerts that are sensitive to stereochemistry.

11. Predictive performance of selected QSAR tools for developmental toxicity

On the basis of the analysis of the performances of selected QSAR tools for developmental toxicity, it is concluded that:

- The literature-based model for placental transfer shows no tendency to distinguish between developmental toxicants and non-toxicants. This is not unexpected since the passage of a chemical across the placental barrier is not a sufficient condition for embryotoxicity. Furthermore, even if a chemical does not pass the barrier, this does not necessarily mean that it will not be a developmental toxicant, since adverse effects may be caused before the placental barrier is fully formed.
- To predict the absence of developmental toxicity, the PASS models for embryotoxicity and teratogenicity appear to be the best stand-alone models in terms of their negative predictivities (44-45%) when assessed against the EFSA Extended Test Set. The combined use of two models led to a marginal increase in negative predictivity to 48%. With negative predictivities less than 50%, none of the models investigated, and no two-model combination, is expected to be adequate for use.

- Some QSAR tools, such as Derek, HazardExpert and PASS, might be useful for the identification of developmental toxicants (due to their high positive predictivities of 81-96% when assessed against the EFSA Extended Test Set). In particular, such models could be useful in the context of a stepwise assessment strategy in which the use of QSAR to identify positives is followed by the use of read-across to identify negatives.
- When evaluating the performances of QSAR models, care should be taken in the choice of test set, since different criteria for discriminating between positives and negatives may be used by different regulatory bodies or database providers. For example, when the developmental toxicity models were assessed against the US EPA's ToxRefDB dataset, their performances were strongly dependent on how the Lowest Effect Levels (LELs) for developmental effects were compared with the LELs for maternal effects.

12. Predictive performance of selected QSAR tools for neurotoxicity

On the basis of the analysis of the performances of selected QSAR tools for neurotoxicity, it is concluded that:

- To predict the absence of neurotoxic potential, no individual model, and no two-model combination, appears adequate for use (since the negative predictivities are less than 50%).
- Conversely, some tools, such as Derek and HazardExpert, might be useful for the identification of neurotoxicants (due to their high positive predictivities between 90-100% when assessed against the EFSA test set). In particular, such software tools might be useful in the context of a stepwise assessment strategy in which the use of QSAR to identify positives is followed by the use of read-across to identify negatives. This possibility could not be explored in this study, due to the lack of a suitable reference database for the read-across exercise. To investigate the applicability of this assessment strategy, it will be necessary to develop such a database.
- In view of the lack of available tools for predicting the absence of neurotoxic potential, the most pragmatic consideration in the risk assessment of pesticide metabolites/degradates is to apply the hypothesis that non-neurotoxic parent substances do not generate neurotoxic (bio)transformation products. Based on its experience of evaluating pesticide dossiers, EFSA could not identify any evidence that refutes this hypothesis. In addition, the only evidence we could find of non-neurotoxic parents (but not of pesticides) giving rise to products with neurotoxic effects were a few papers describing *in vitro* / mechanistic findings, which are not necessarily relevant to the *in vivo* effects of pesticides.

13. Chemical space analysis

Chemical space analysis can be used to explore and define the applicability domains of statistically-based models if their training sets (including structures and biological data) are available. It can also be used to inform model development by identifying areas of chemistry that are not adequately covered in existing models.

There are different approaches to building applicability domains, but two of the most commonly used approaches are based on structural fragments and molecular descriptors. Model applicability domains can be used to rationalise the predictions made for test set chemicals (for which the toxicological effects are known) and to help determine the reliability of prediction for untested chemicals. However, the interpretation is not straightforward. If a chemical is outside the applicability domain of a model, it does not necessarily mean that its predicted toxicity is wrong, but simply that the prediction cannot be made with as much confidence. Conversely, when a chemical is within the applicability domain, it does not necessarily follow that the predicted toxicity will be accurate, but simply that the prediction can be made with a defined level of confidence. Furthermore, there is no absolute definition of a model applicability domain – different interpretations may be useful for different purposes. Some software tools provide their own assessment of prediction reliability based on applicability domain considerations, whereas other software tools do not. In practice, the definition and interpretation of applicability domains is not a trivial exercise. In this study, only the CAESAR and Leadscape models were amenable to chemical space analysis. The results were not particularly informative in terms of understanding the reliability of prediction, which reinforces the view that such analyses should be regarded as indicative rather than conclusive.

In addition to exploring the applicability domains of QSAR models, chemical space analysis can be used to compare the test sets with the “universe” of pesticides, as represented by the Plant Protection Products (PPP) inventory. In this study, the developmental toxicity test sets were found to largely cover the space of the PPP inventory, while the (smaller) neurotoxicity test set was less diffuse and chemically diverse. In any future efforts to build a more extended neurotoxicity test set as a means of providing a more comprehensive challenge to available models, it would be useful to search for chemicals in these areas of the PPP. However, there is no guarantee that reference chemicals with adequate data will be found.

14. Predictive ability of read-across

Compared with QSAR models, the predictive performance of the grouping and read-across approach cannot be generalised as easily, since this is an *ad hoc* approach in which a number of subjective choices are made. For example, decisions have to be made concerning the choice of reference database(s) for analogue searching, the similarity criteria used to identify close analogues, and the assessment of the relevance and reliability of the analogue data, and the interpretation of positive, negative and inconclusive outcomes. Nevertheless, the read-across approach is equally well-suited to the identification of positive and negative chemicals, and for this reason, it is proposed in this study as a means of clarifying the negative predictions resulting from the application of QSAR.

Developmental toxicity and neurotoxicity are complex endpoints, which are only partially understood in mechanistic terms. In the absence of endpoint-specific profilers for analogue identification, analogues can be identified based on the presence of organic functional groups.

In this study, we have illustrated the usefulness of read-across in identifying developmental toxicants, non-developmental toxicants, and chemicals for which there is insufficient evidence to draw a conclusion on developmental toxicity. Only a few other studies have been published on the read-across of developmental toxicity, and these also support the

applicability of grouping and read-across, especially when used in a weight-of-evidence approach along with QSAR (Enoch et al, 2009; Hewitt et al, 2010).

15. Practical use of evaluated models

An evaluation of the practical usefulness of QSAR and read-across tools, for the purposes of pesticide metabolite assessment, is not a straightforward task. Such an evaluation needs to take into account the availability and cost of the software, the expertise required to use the software, as well as the validation characteristics and the regulatory context in which the models are being used. A comprehensive framework for characterising and assessing model validation characteristics has been proposed previously (JRC, 2010), but this does not go as far as including clear criteria.

When establishing the pesticide residue definition for risk assessment, any decision on which models / software tools are fit-for-purpose should be taken by EFSA, ideally on the basis of a transparent set of acceptance criteria (which could be developed, for example, by the PPR Panel). In particular, EFSA needs to decide on the acceptable false positives and false negatives in the use of the models, taking into account that models with different strengths and weaknesses can be combined in a stepwise strategy that optimises the overall predictive performance. To help EFSA experts make this decision, the main considerations and the findings of this project are summarised in Tables 15.1 and 15.2, for developmental toxicity and neurotoxicity, respectively.

The challenge of this project - to assess the ability of models to predict the absence of toxic potential - was quite unusual from a QSAR perspective, since QSAR models are generally designed to predict toxicity by identifying structural features associated with the molecular interactions that lead to toxicological outcomes. Structural features are rarely associated with the absence of toxicity, unless they are structural groups that have a mitigating effect on the properties of another group (e.g. steric hindrance, or alteration of chemical reactivity via electronic polarisation). To some extent, QSARs may also capture the absence of toxicity to the extent that they implicitly encode ADME characteristics, such as limitations in bioavailability due to molecular size or hydrophobicity. In contrast, the read-across approach is equally suited to the identification of toxicants and non-toxicants, provided that a sufficient number of analogues can be found with adequate experimental data. In this respect, in the absence of a more specific mechanistic understanding of toxicity, analogue searching by organic functional groups is particularly useful, since analogues that contain different (and potentially reactive) functional groups to the chemical of interest can be excluded.

In this study, the software models evaluated were found to have good abilities to identify positive chemicals (positive predictivities greater than 80%) but poor abilities to identify negative chemicals (negative predictivities less than 50%). The strengths of these QSAR models can be exploited in a stepwise strategy in which QSARs are used in a preliminary step only for the identification of positive chemicals (in other words, the positive predictions are trusted, but no confidence is attached to the negative predictions), whereas a subsequent step based on grouping and read-across is used to discriminate between the true and false negatives generated by QSAR. This was found to be an effective strategy for the prediction of developmental toxicity. The concept could not be tested for the prediction of neurotoxicity due to a lack of a suitable reference database for read-across. Nevertheless, one would expect

a similar stepwise approach to neurotoxicity prediction to be more effective than the use of QSAR models alone.

The stepwise non-testing approach could be used, along with computational methods for other toxicological endpoints (e.g. genotoxicity) and the TTC approach, into a decision tree for evaluating the toxicological relevance of metabolites and degradates.

16. Future developments in predictive toxicology

At present, the ability to predict complex toxicological endpoints such as neurotoxicity and developmental toxicity is limited by an incomplete understanding of the multiple and context-dependent mechanisms by which the adverse outcomes are triggered by chemicals. However, the increasing application of advanced data analysis methods to multi-parametric bioactivity datasets generated by high throughput *in vitro* methods promises to revolutionise predictive toxicology, and in the long-term change the way in which chemical risk assessments are performed. There is a growing recognition by the scientific and regulatory communities that predictive toxicology should move away from the modelling of apical endpoints, especially for long-term systemic effects, towards modelling the steps that link the initial exposure to a chemical to the final adverse outcome (at the individual or population level) via a series of so-called intermediate effects or “key events”. This is called the mode-of-action (MOA) or adverse outcome pathway (AOP) approach (Schultz, 2010).

Initial findings are encouraging. For example, a study carried out in the context of the OECD QSAR Toolbox project (Philip Judson, Lhasa Ltd, personal communication) has suggested that developmental toxicity can be understood in terms of the chemically-induced perturbations of a (limited) set of 17 molecular signalling pathways that are conserved across species. The approach was illustrated with reference to the Wnt/ β -catenin pathway. In relation to neurotoxicity, examples of the MOA/AOP approach have been provided for kainate receptor-mediated excitotoxicity (Watanabe et al, 2010) and for the interactions of pyrethroids with voltage-gated sodium channels (OECD, 2011).

This is a long-term vision. The implementation of the MOA/AOP approach in a regulatory context will require a considerable amount of research into the underlying biology of the toxicological endpoints, as well as the development of suitable software tools, investigations into the validity and applicability of the approach, and the establishment of consensus on how to report and interpret the results in a regulatory framework.

Table 15.1 Considerations to support decisions on the usefulness of approaches for predicting developmental toxicity

Consideration	QSAR models	Grouping and read-across
1. Availability of models / data	<ul style="list-style-type: none"> 7 models were identified as potentially suitable (Derek, CAESAR, TOPKAT, HazardExpert, PASS embryotoxicity and PASS teratogenicity) 	<ul style="list-style-type: none"> Several tools are available for grouping and read-across, including the OECD QSAR Toolbox and Toxmatch ToxRefDB was considered a suitable public domain database to support read-across assessments of developmental toxicity Important to check the criteria used to distinguish between positives and negatives
2. Ease-of-use	<ul style="list-style-type: none"> Most of the investigated tools require specialised expertise, either of a practical nature, or in terms of setting up model specifications and interpreting the model predictions Other tools are suitable for non-specialists (e.g. CAESAR) The use of some tools, in particular Derek, is facilitated by strong customer support from the developer 	<ul style="list-style-type: none"> Requires specific expertise and access to suitable reference database in electronic format The application of grouping read-across is not an automatic process, and requires the services of an expert user
3. Cost	<ul style="list-style-type: none"> Some tools for developmental toxicity prediction are freely available: PASS (limited version) and CAESAR Other tools require a license, possibly with periodic renewal: Derek, HazardExpert, TOPKAT, PASS (full version) 	<ul style="list-style-type: none"> OECD QSAR Toolbox and Toxmatch are freely downloadable from the internet Training materials freely available for the OECD Toolbox and Toxmatch
4. Validation characteristics	<ul style="list-style-type: none"> QSARs are often better suited for the identification of positives Standalone models found to have positive predictivities in the range 68-96% Standalone models found to have negative predictivities in the range 21-44% Consensus modelling based on the combined use of two models can produce a marginal increase in negative predictivity (48%) 	<ul style="list-style-type: none"> Equally well suited to the identification of positives and negatives Validation statistics are not meaningful, since this is an <i>ad hoc</i> approach based on expert choices and evaluation Since the QSAR models evaluated have high positive predictivities, read-across adds value when used to distinguish between the true and false negatives generated by QSAR

The comments in this table reflect the views and experience of the authors.

Table 15.2 Considerations to support decisions on the usefulness of approaches for predicting neurotoxicity

Consideration	QSAR models	Grouping and read-across
1. Availability of models / data	<ul style="list-style-type: none"> 2 models were identified as potentially suitable (Derek, HazardExpert) Several models are available for predicting blood-brain barrier penetration, although these were not considered directly relevant on their own 	<ul style="list-style-type: none"> Several tools are available for grouping and read-across, including the OECD QSAR Toolbox and Toxmatch. No suitable public domain database was found to support read-across assessments of neurotoxicity
2. Ease-of-use	<ul style="list-style-type: none"> The investigated require specialised expertise (Derek, HazardExpert), either of a practical nature, or in terms of setting up model specifications and interpreting the model predictions The use of some tools, in particular Derek, is facilitated by strong customer support from the developer 	<ul style="list-style-type: none"> Requires specific expertise and access to suitable reference database in electronic format The application of grouping read-across is not an automatic process, and requires the services of an expert user
3. Cost	<ul style="list-style-type: none"> At present, there are no freely available tools for neurotoxicity prediction. Other tools require a license, possibly with periodic renewal: Derek, HazardExpert, TOPKAT, PASS (full version), Leadscope 	<ul style="list-style-type: none"> OECD QSAR Toolbox and Toxmatch are freely downloadable from the internet Training materials freely available for the OECD Toolbox and Toxmatch
4. Validation characteristics	<ul style="list-style-type: none"> QSARs are often better suited for the identification of positives Standalone models have positive predictivities in the range 69-100% Standalone models have negative predictivities in the range 38-43% Consensus modelling based on the combined use of two models can produce a marginal increase in negative predictivity (48%) 	<ul style="list-style-type: none"> Equally well suited to the identification of positives and negatives Validation statistics are not meaningful, since this is an <i>ad hoc</i> approach based on expert choices and evaluation Since the QSAR models evaluated have high positive predictivities, read-across adds value when used to distinguish between the true and false negatives generated by QSAR

The comments in this table reflect the views and experience of the authors.

RECOMMENDATIONS

Based on the findings of this study, the following recommendations are made with a view to improving the use of computational methods to identify the presence and absence of short-term effects of high concern, such as neurotoxicity and developmental toxicity.

Short-term (< 2 years)

1. EFSA should define acceptance criteria for the use of QSAR models, taking into account the context in which they are expected to be used (for example, in tiered assessment approaches, and/or in a broader TTC assessment scheme).
2. The stepwise strategy based on the application of QSAR to identify positive chemicals, and the subsequent application of read-across to distinguish between true and false QSAR negatives should be further explored, using existing software tools and databases. Additional combinations of models should be explored in the context of a strategy for developmental toxicity. In addition, the applicability of this approach should be investigated for the assessment of neurotoxic potential, which will require the establishment of a suitable neurotoxicity reference database. Real-world scenarios based on realistic ratios of positive to negative substances should be explored.
3. The ability to categorise, and develop predictive models, for pesticides would be enhanced by the development of a structure-searchable database containing the high quality toxicological data available in EFSA dossiers. The planning of such a database development project, including the definition of technical specifications and user requirements, could be carried out in the short-term.

Mid-term (2-5 years)

4. The development of a searchable database, enabling both knowledge management and the development of predictive models, will require considerable effort in terms of identifying suitable data sources, digitising original study records, and checking the quality of the data. The US EPA carried out this type of exercise by digitising data on hundreds of regulated pesticides and by including the data, along with quality-checked chemical structures, in the ToxRef Database.
5. New and updated QSAR models and expert systems could be developed in parallel with the collation of relevant and reliable data.

Long-term (> 5 years)

6. It can be anticipated that new models for predicting complex toxicological endpoints such as neurotoxicity and developmental toxicity will move away from the modelling of apical effects, toward the modelling of key events in a mode-of-action (MOA) and adverse outcome pathway (AOP). Research on the MOA/AOP approach to hazard and risk assessment has already started. The implementation of this approach in a regulatory context will require a considerable amount of research into the underlying

biology of the toxicological endpoints, investigations into the validity and applicability of the approach, and consensus on how to report and interpret the results in a regulatory framework.

References

- AGES, 2010. Impact of metabolic and degradation processes on the toxicological properties of residues of pesticides in food commodities. Final report to EFSA in relation to EFSA Grant Agreement EFSA/PPR/2008/02. Austrian Agency for Health and Food Safety (AGES), Austria. Available at: <http://www.efsa.europa.eu/en/scdocs/scdoc/49e.htm>
- Arena VC, Sussman NB, Mazumdar S, Yu S and Macina OT, 2004. SAR QSAR Environ Res. 15, 1-18
- Barlow S, 2005. Threshold of toxicological concern (TTC) - A tool for assessing substances of unknown toxicity present at low levels in the diet. ILSI Europe's Concise Monographs and Report 34.
- Bassan A, Fioravanzo E, Pavan M & Stocchero M, 2011. Applicability of physicochemical data, QSARs and read-across in Threshold of Toxicological Concern assessment. Final report of a study carried out by Soluzioni Informatiche (S-In, Vicenza, Italy) for the European Food Safety Authority (EFSA).
- Brain EGC, Yu LJ, Gustafsson K, Drewes P and Waxman DJ, 1998. Modulation of P450-dependent ifosfamide pharmacokinetics: a better understanding of drug activation in vivo. Br J Cancer. 77, 1768-1776
- Briggs GC, Freeman RK and Yaffe SJ, 2002. Drugs in Pregnancy and Lactation. 6th edn. Lippincott Williams and Wilkins
- Cassano A, Manganaro A, Martin T, Young D, Piclin N, Pintore M, Bigoni D and Benfenati E, 2010. CAESAR models for developmental toxicity. Chemistry Central Journal. 4, S4. Open access article available at: <http://www.journal.chemistrycentral.com/content/4/S1/S4>
- Cramer GM, Ford RA and Hall RL, 1978. Estimation of toxic hazard--a decision tree approach. Food Cosmet Toxicol. 16, 255-276
- CRD, 2009. Applicability of thresholds of toxicological concern in the dietary risk assessment of metabolites, degradation and reaction products of pesticides. Final report to EFSA in relation to EFSA Grant Agreement EFSA/PPR/2008/01. Chemicals Regulation Directorate, UK. Available at: <http://www.efsa.europa.eu/en/scdocs/scdoc/44e.htm>
- Crivori P, Cruciani G, Carrupt PA and Testa B, 2010. Predicting Blood-Brain Barrier Permeation from Three-Dimensional Molecular Structure. J Med Chem. 43, 2204-2216
- Cronin M, 2010. Quantitative structure-activity relationships (QSARs) – applications and methodology. Chapter 1 in *Recent Advances in QSAR Studies: Methods and Applications*. T Puzyn, J Leszczynski & MTD Cronin (Eds). Springer, Heidelberg, Germany, pp. 3-11.
- Diderich R, 2010. Tools for Category Formation and Read-Across: Overview of the OECD (Q)SAR Application Toolbox. Chapter 16 in *In Silico Toxicology: Principles and Applications*. MTD Cronin & JC Madden (Eds). The Royal Society of Chemistry, Cambridge, UK, pp 385-407

- Dingemans MML, Heusinkveld HJ, Bergman Å, van den Berg M and Westerink RHS, 2009. Bromination Pattern of Hydroxylated Metabolites of BDE-47 Affects □ Their Potency to Release Calcium from Intracellular Stores in PC12 Cells. *Environ Health Perspect.* 118, 519-525
- EC, 1991. Council Directive 91/414/EEC of 15 July 1991 concerning the placing of plant protection products on the market. Official Journal of the European Union, L 230/1 of 19.08.1991. Office for Official Publications of the European Communities (OPOCE), Luxembourg.
- EFSA, 2011. Conclusion on the peer review of the pesticide risk assessment of the active substance acetochlor. *EFSA Journal* 9(5):2143. European Food Safety Authority. Available online: <http://www.efsa.europa.eu/en/efsajournal/doc/2143.pdf>
- Ellison C, Enoch S, Cronin M, Madden J and Judson P, 2009. Definition of the applicability domains of knowledge-based predictive toxicology expert systems by using a structural fragment-based approach. *Altern Lab Anim.* 37, 533-545
- Ellison CM, Sherhod R, Cronin MT, Enoch SJ, Madden JC and Judson PN, 2011. Assessment of Methods to Define the Applicability Domain of Structural Alert Models. *J Chem Inf Model.*
- Enoch S, 2010. Chemical category formation and read-across for the prediction of toxicity. Chapter 7 in *Recent Advances in QSAR Studies: Methods and Applications*. T Puzyn, J Leszczynski & MTD Cronin (Eds). Springer, Heidelberg, Germany, pp. 209-219
- Enoch SJ, Cronin MTD, Madden JC and Hewitt M, 2009. Formation of Structural Categories to Allow for Read-Across for Teratogenicity. *QSAR & Combinatorial Science.* 28, 696-708
- Friedman JM, 1993. Report of the Teratology Society Public Affairs Committee symposium on FDA classification of drugs. *Teratology.* 48, 5-6. Available at: <http://onlinelibrary.wiley.com/doi/10.1002/tera.1420480103/pdf>
- Hewitt M, Madden JC, Rowe PH and Cronin MTD, 2007. Structure-based modelling in reproductive toxicology: (Q)SARs for the placental barrier. *SAR QSAR Environ Res.* 18, 57-76
- Hewitt M, Ellison CM, Enoch SJ, Madden JC and Cronin MTD, 2010. Integrating (Q)SAR models, expert systems and read-across approaches for the prediction of developmental toxicity. *Reprod Toxicol.* 30, 147-160
- Enoch S, 2010. Chemical category formation and read-across for the prediction of toxicity. Chapter 7 in *Recent Advances in QSAR Studies: Methods and Applications*. T Puzyn, J Leszczynski & MTD Cronin (Eds). Springer, Heidelberg, Germany, pp. 209-219
- Jeliazkova N, Jaworska J and Worth A, 2010. Open source tools for read across and category formation. Chapter 17 in *In Silico Toxicology: Principles and Applications*. MTD Cronin & JC Madden (Eds). The Royal Society of Chemistry, Cambridge, UK, pp 408-445
- JRC, 2010. Applicability of QSAR analysis to the evaluation of the toxicological relevance of metabolites and degradates of pesticide active substances for dietary risk assessment. Final

- report to EFSA in relation to Service Level Agreement SLA/EFSA-JRC/PPR/2010/01. Available at: <http://www.efsa.europa.eu/en/scdocs/scdoc/50e.htm>
- Lapenna S, Fuat-Gatnik M and Worth A, 2010. Review of QSAR Models and Software Tools for predicting Acute and Chronic Systemic Toxicity. JRC Technical Report EUR 24639 EN. Publications Office of the European Union, Luxembourg. Available at: <http://publications.jrc.ec.europa.eu/repository/>
- Lo Piparo E and Worth A, 2010. Review of QSAR Models and Software Tools for predicting Developmental and Reproductive Toxicity. JRC Technical Report EUR 24522 EN. Publications Office of the European Union, Luxembourg. Available at: <http://publications.jrc.ec.europa.eu/repository/>
- Mostrag-Szlichtyng A and Worth A, 2010. Review of QSAR Models and Software Tools for predicting Biokinetic Properties. JRC Technical Report EUR 24377 EN. Publications Office of the European Union, Luxembourg. Available at: <http://publications.jrc.ec.europa.eu/repository/>
- Netzeva TI, Worth A, Aldenberg T, Benigni R, Cronin MT, Gramatica P, Jaworska JS, Kahn S, Klopman G, Marchant CA, Myatt G, Nikolova-Jeliazkova N, Patlewicz GY, Perkins R, Roberts D, Schultz T, Stanton DW, van de Sandt JJ, Tong W, Veith G and Yang C, 2005. Current status of methods for defining the applicability domain of (quantitative) structure-activity relationships. The report and recommendations of ECVAM Workshop 52. *Altern Lab Anim.* 33, 155-173
- OECD, 2006. Guidance Document on the Definition of Residue. Series on Testing and Assessment Number 63, Series on Pesticides Number 31. ENV/JM/MONO(2006)31. Organisation for Economic Cooperation and Development, Paris, France. Available at: <http://www.oecd.org/>
- OECD, 2007. Guidance on Grouping of Chemicals. Series on Testing and Assessment Number 80. ENV/JM/MONO(2007)28. Organisation for Economic Cooperation and Development, Paris, France. Available at: <http://www.oecd.org/>
- OECD 2011. Report of the Workshop on Using Mechanistic Information in Forming Chemical Categories. OECD Environment, Health and Safety Publications, Series on Testing and Assessment, No. 138. ENV/JM/MON(2011)/8. Available at: [http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO\(2011\)8&docLanguage=En](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO(2011)8&docLanguage=En)
- Poroikov V, Filimonov D, Lagunin A, Gloriozova T and Zakharov A, 2007. PASS: identification of probable targets and mechanisms of toxicity. *SAR QSAR Environ Res.* 18, 101-110
- Schultz T, 2010. Adverse Outcome Pathways: A Way of Linking Chemical Structure to In Vivo Toxicological Hazards. Chapter 14 in *In Silico Toxicology: Principles and Applications*. MTD Cronin & JC Madden (Eds). The Royal Society of Chemistry, Cambridge, UK, pp 346-371
- Solecki R, Davies L, Dellarco V, Dewhurst I, Raaij M and Tritscher A, 2005. Guidance on setting of acute reference dose (ARfD) for pesticides. *Food Chem Toxicol.* 43, 1569-1593

- Steinberg DM, Fine J and Chappell R, 2008. Sample size for positive and negative predictive value in diagnostic research using case-control designs. *Biostatistics* 10, 94-95. Open-access online: <http://biostatistics.oxfordjournals.org/content/10/1/94.full>
- Todeschini R and Consonni V (2009). Molecular Descriptors for Chemoinformatics. In *Series of Methods and Principles in Medicinal Chemistry* (Mannhold R., Kubinyi H, Timmerman H, Eds), Volume 41, Wiley-VCH.
- UN, 2009. Globally Harmonized System of Classification and Labelling of Chemicals (GHS). United Nations. Available at: http://live.unece.org/trans/danger/publi/ghs/ghs_rev03/03files_e.html
- Van Raaij MTM, Janssen PAH and Piersma AH, 2003. The relevance of developmental toxicity endpoints for acute limit setting. RIVM report 601900004. Available at: <http://www.rivm.nl/bibliotheek/rapporten/601900004.pdf>
- Watanabe KH, Andersen ME, Basu N, Carvan MJ, Crofton KM, King KA, Suñol C, Tiffany-Castiglioni E and Schultz IR 2010. Defining and modeling known adverse outcome pathways: Domoic acid and neuronal signaling as a case study. *Environmental Toxicology and Chemistry* 30, 9-21
- Weininger D, 1988. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *Journal of Chemical Information and Computer Sciences* 28, 31-36

Appendices

APPENDIX A	DEVELOPMENTAL TOXICITY DATASET (EFSA SET OF PESTICIDES)
APPENDIX B	RIVM DATASET OF CHEMICALS CLASSIFIED FOR DEVELOPMENTAL TOXICITY
APPENDIX C	NEUROTOXICITY DATASET (EFSA SET OF PESTICIDES)
APPENDIX D	RESULTS OF LEADSCOPE STRUCTURAL FRAGMENT ANALYSIS

Glossary / Abbreviations

AGES	Austrian Agency for Health and Food Safety
AOP	Adverse Outcome Pathway
ARfD	Acute Reference Dose
CAS	Chemical Abstract Service
CRD	UK Chemicals Regulations Directorate
DAR	Draft Assessment Report
EFSA	European Food Safety Authority
EC	European Commission
ECHA	European Chemicals Agency
EPA	Environmental Protection Agency
EU	European Union
FDA	Food and Drug Administration
FN	False Negative
FP	False Positive
GHS	Globally Harmonised System
JMPR	Joint Meeting on Pesticide Residues
JRC	Joint Research Centre
LEL	Lowest Effect Level
MOA	Mode of Action
MRL	Maximum Residue Level
OECD	Organisation for Economic Cooperation and Development
PPP	Plant Protection Product
PPR Panel	EFSA Panel on Protection Products and their Residues
(Q)SAR	(Quantitative) Structure-Activity Relationship
SMILES	Simplified Molecular Input Line Entry System
ToxRefDB	Toxicity Reference Database (US EPA)
TN	True Negative
TP	True Positive
TTC	Threshold of Toxicological Concern

SCIENTIFIC REPORT submitted to EFSA

Applicability of QSAR analysis in the evaluation of developmental and neurotoxicity effects for the assessment of the toxicological relevance of metabolites and degradates of pesticide active substances for dietary risk assessment¹

APPENDICES

Prepared by

Andrew Worth, Mojca Fuart-Gatnik, Silvia Lapenna and Rositsa Serafimova

**Systems Toxicology Unit, Institute for Health & Consumer Protection,
European Commission - Joint Research Centre, Ispra, Italy**

Final version of 31/05/11

¹ Question No: EFSA-Q-2010-01258. Accepted for publication on 10 June 2011.

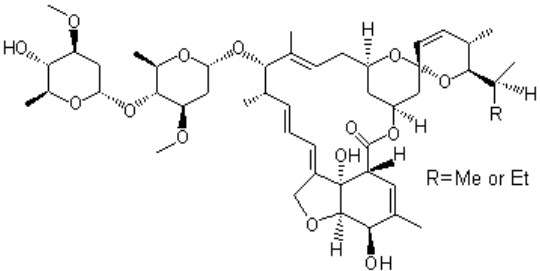
Appendices

A.	Developmental toxicity dataset (EFSA set of pesticides).....	3
B.	RIVM dataset of chemicals classified for developmental toxicity	36
C.	Neurotoxicity dataset (EFSA set of pesticides)	51
D.	Results of Leadscope structural fragment analysis	66

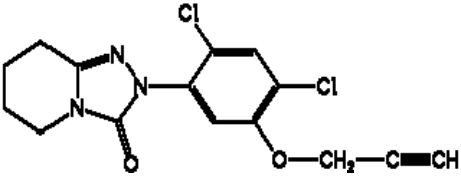
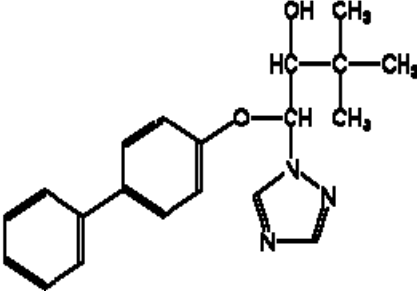
Appendices

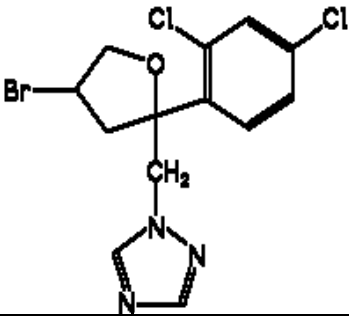
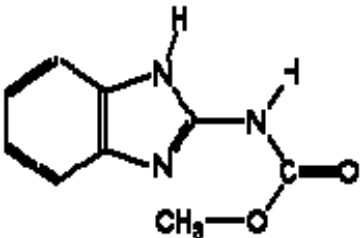
A. Developmental toxicity dataset (EFSA set of pesticides)

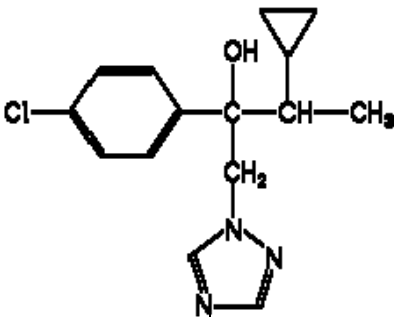
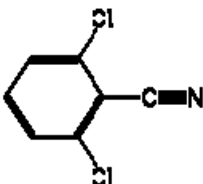
26 CLEAR DEVELOPMENTAL POSITIVES SELECTED BY EFSA

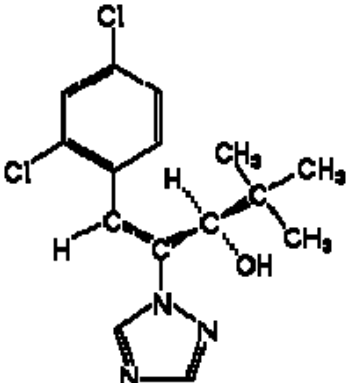
No	Compound	Pesticide class	CAS No	Classification Current EU classification	ARfD mg/kg bw	Basis for ARfD	Comments
1	<p>Abamectin</p>  <p>R=Me or Et</p>	Fermentation product <i>S. avermitilis</i>	71751-41-2	<p>Proposed R61 or 63</p> <p>Not classified</p> <p>R63 adopted by ECHA</p>	0.005	Acute neurotoxicity rat	<p>Rat: cleft palate, lumbar rib and lumbar count variation (in the absence of maternal toxicity)</p> <p>Rabbit: cleft palate, omphaloceles, clubbed fore-feet and delayed ossification (at maternally toxic dose)</p> <p>Maternal NOAELs: 1.6 mg/kg bw/day (rat) 1.0 mg/kg bw/day (rabbit)</p> <p>Developmental NOAELs: 0.8 mg/kg bw/day (rat) 1.0 mg/kg bw/day (rabbit)</p>

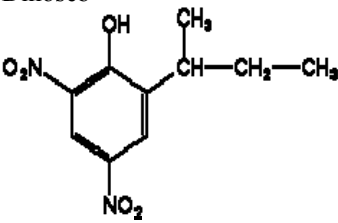
The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a service level agreement between the European Food Safety Authority and the European Community. The present document is published complying with the transparency principle to which the European Food Safety Authority is subject. It may not be considered as an output adopted by EFSA. EFSA reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors

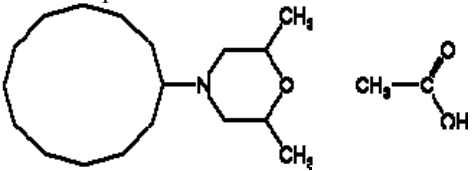
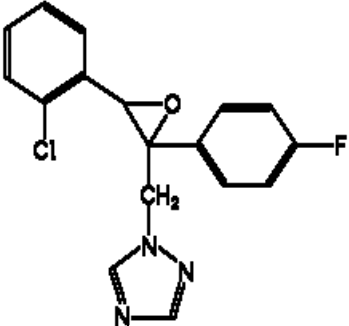
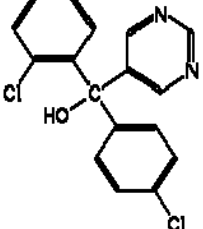
No	Compound	Pesticide class	CAS No	Classification Current EU classification	ARfD mg/kg bw	Basis for ARfD	Comments
2	<p>Azafenidin</p> 	Triazolinone	68049-83-2	R61 Category 1 (2)?? Repr. Cat. 2; R61	0.0004	13-week mechanistic study in dog (ECCO 123)	<p>Annex 1 not included (withdraw from Company) DAR available</p> <p>Although no data available. Classification suggests dev tox. NOAEL dev oral rat: 16 mg/kg bw/d (↑ fetal resorptions, ↓ litter size, ↓ live fetuses, ↓ mean fetal weight, stat ↑ stern variations including fusion, mis-alignment and mis-development) NOAEL mat tox oral rat: 24 mg/kg bw/d</p> <p>NOAEL dev dermal rat: 5 mg/kg bw/d (↑ post implantation loss, ↓ litter size, ↓ fetal weight, ↑ skeletal and visceral variations, skeletal malformations (bent ribs)) NOAEL mat tox dermal rat: 100 mg/kg bw/d</p>
3	<p>Bitertanol</p> 	Triazole	55179-31-2	R61?? Proposed in EFSA conclusions Not classified (not discussed in ISPRA or ECHA after PRAPeR)	0.01	90-day dog study, clinical signs (alterations of skin and hair loss), transient weight loss, small increases in AP and ALT activities and effects on the prostate	<p>Rat developmental –maternal and developmental NOAEL 10 mg/kg bw/d (↑ stunted fetuses, skeletal variations). Malformations in rat and rabbit, evidence of adverse effects in the absence of overt maternal toxicity in rats. NOAEL for cranio-facial malformations in rat 30 mg/kg bw/day</p>

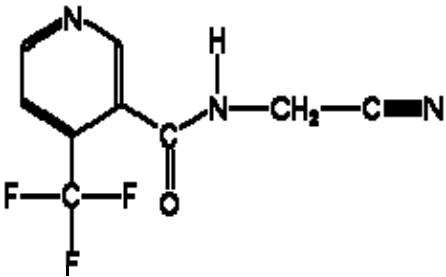
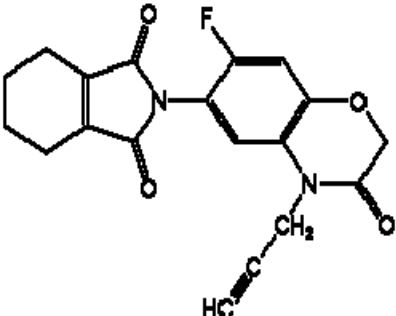
No	Compound	Pesticide class	CAS No	Classification Current EU classification	ARfD mg/kg bw	Basis for ARfD	Comments
4	Bromuconazole 	Triazole	116255-48-2	R63?? (EFSA concl) Not classified (not yet discussed at ECHA)	0.1	Developmental NOAEL 10 mg/kg bw/d (ossification delay, supplementary cervical ribs at 70 mg/kg, which is maternal NOAEL)	EFSA re-evaluation 2010 - usually supernumerary ribs are considered as variations rather than as malformations. But, they appeared at doses where no maternal toxicity was observed. The observed “domed head” (a hydrocephalus- like malformation) is a typical observation upon exposure to triazoles.
5	Carbendazim 	Benzi imidazole	10605-21-7	R61 Category 1 (2) ?? Repr. Cat. 2; R61	0.02 SF 500	Developmental, rat and rabbit NOAELs of 10 mg/kg bw/d, Rat: high resorption rate, foetal wt ↓, skeletal variations, malformations (e.g. hydrocephalus, anophthalmia) Rabbit: implantation ↓, resorptions ↑, live litter size ↓, skeletal malformations In maternal non toxic doses – maternal NOAELs Rat: 30 mg/kg bw/d Rabbit: 20 mg/kg bw/d	JMPR 2005 ARfD 0.1 mg/kg bw for women of childbearing age based on an overall NOAEL of 10 mg/kg bw per day for developmental toxicity from three studies in rats and one study in rabbits, SF 100. ARfD 0.5 mg/kg bw for the general population, including children based on the NOAEL of 50 mg/kg bw in the study of toxicity to the male reproductive system in rats and supported by the studies on micronucleus or aneuploidy induction in vivo, SF 100.

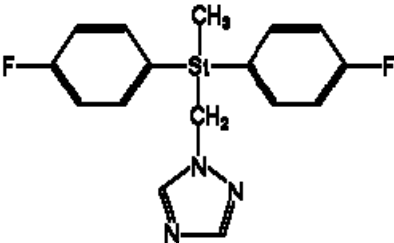
No	Compound	Pesticide class	CAS No	Classification Current EU classification	ARfD mg/kg bw	Basis for ARfD	Comments
6	Cyproconazole 	Triazole	94361-06-5	R63 Category 2 (3)?? Repr. Cat. 3; R63	0.02	NOAEL developmental, rabbit 2 mg/kg bw/d. Serious malformations at doses from 20 mg/kg bw/d Hydrocephalus and cleft palate in all rat studies	JMPR 2010 ARfD 0.06 mg/kg bw based on maternal NOAEL 6 mg/kg bw/d Use to be discussed. Maybe further info on maternal toxicity would help to decide Maternal (rabbit): ↓mean body weight NOAEL 10 mg/kg bw/d Developmental (rabbit): Increased post-implantation loss; Increased foetal malformations Maternal (rat) ↓mean body weight gain. Developmental: Reduced foetal body weight teratogenicity (cleft palate, hydrocephaly) in the rat at maternally toxic doses.
7	Dichlobenil 	Benzonitrile	1194-65-6	Proposed R63 Not classified (not yet discussed at ECHA)	0.45	Rabbit developmental NOAEL 45 mg/kg bw/d	Specific effects (grossly enlarged fontanelles, and other major cranial defects, open eyes; cleft palate and hydrocephaly, decreased body weight) at a dose (135 mg/kg bw/day) with limited maternal toxicity (rabbit only) Use to be discussed Developmental toxicity clear but only in rabbit Clear dev tox in rabbit not related to maternal tox

No	Compound	Pesticide class	CAS No	Classification Current EU classification	ARfD mg/kg bw	Basis for ARfD	Comments
8	Diniconazole (M) 	Triazole	83657-18-5	R61 Not classified	0.02 (proposed in DAR)	Based on NOAEL 5 mg/kg bw/day, SF 250	Annex 1 not included Developmental NOAEL – rat, oral: 5 mg/kg bw/day Rat, oral: embryo/foetotoxicity (lower implantation efficiency, early resorptions) and skeletal variations (cervical and 14th ribs, bifid centra of thoracic vertebrae) below maternal toxic dose; NOAEL for teratogenicity in rat 80 mg/kg bw/day – external (cleft palate and minor microcephaly) and skeletal anomalies (maxillo-mandibular synostosis) at maternally toxic dose.

No	Compound	Pesticide class	CAS No	Classification Current EU classification	ARfD mg/kg bw	Basis for ARfD	Comments
9	<p>Dinoseb</p> 	Dinitrophenol	88-85-7	<p>R61 Category 1 (2)?? Repr. Cat. 2; R61</p>			<p>Annex 1 not included DAR and EFSA conclusion not available EPA- Rabbit: Developmental Toxicity NOEL=3 mg/kg/day [based on biological and statistically significant increases in malformations and/or anomalies at the high dose (10 mg/kg/day) with external, internal and skeletal defects observed in 11/16 litters examined; brain/spinal cord defects accounted for majority of developmental toxicity and included dyscrania associated with hydrocephaly, hydrocephaly alone, scoliosis, malformed/fused caudal or sacral vertebrae and encephalocoele]; Maternal NOEL=10 mg/kg/day (based on lack of significant observable systemic toxicity). Rat: Developmental Toxicity NOEL=3 mg/kg/day [based on relative increase in reported incidence of absence of ossification for a number of skeletal sites (phalangeal) nuclei, cervical vertebrae, etc.) and supernumerary ribs (left or right sides of rib 14) at high dose]; Maternal Systemic NOEL=3 mg/kg/day (based on moderate mean body weight depression).</p>

No	Compound	Pesticide class	CAS No	Classification Current EU classification	ARfD mg/kg bw	Basis for ARfD	Comments
10	Dodemorph acetate 	Morpholine	31717-87-0	Not classified	0.4	NOAEL of 40 mg/kg bw/d for developmental effects seen in the rabbit teratogenicity study, SF 100	Rabbit: Malformations (open eye; variation (irregularly shaped sternebra), at maternal NOAEL 120 mg/kg bw/d
11	Epoxiconazole 	Triazole	133855-98-8 (formerly 106325-08-0)	R63 Category 2 (3)?? Repr. Cat. 3; R63 R61 proposed at ECHA	0.023	NOAEL 2.3 mg/kg bw/d (parental, reproductive, offspring)	Reproductive toxicity at parentally toxic doses: impaired fertility, prolonged gestation, dystocia, number of viable pups reduced, perinatal mortality increased evidence for aromatase inhibition in vitro and in vivo. Developmental toxicity NOAEL (developmental, maternal) rat 15 mg/kg bw/d Malformations at maternally toxic doses (cleft palates), at lower dose increased skeletal variations (additional cervical ribs) and higher placental weight
12	Fenarimol 	Pyrimidine	60168-88-9	R63 Category 2 (3)? Repr. Cat. 3; 63	0.02	Rat multigeneration NOAEL 2 mg/kg bw/day; reduced fertility and parturition effects (rat and mouse).	Annex 1 not included Increased hydronephrosis (rat) without maternal toxicity and extra ribs (rabbit) with maternal toxicity NOAEL 13 mg/kg bw/day (rat)

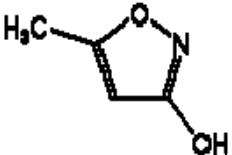
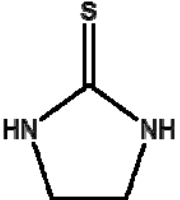
No	Compound	Pesticide class	CAS No	Classification Current EU classification	ARfD mg/kg bw	Basis for ARfD	Comments
13	<p>Flonicamid</p> 	Pyridine	158062-67-0	Proposed R63 Not classified.	0.025	Rabbit developmental	In the rat teratology study, the maternal NOAEL was 100 mg/kg bw/day, based on effects observed in the kidneys and liver. The developmental NOAEL was also 100 mg/kg bw/day, related to an increased incidence of skeletal variations, namely extra cervical ribs. The length of the rib was considered as adverse, even occurring in the presence of slight maternal toxicity. In the rabbit teratologystudy, the maternal NOAEL was 7.5 mg/kg bw/day, based on reduced body weight gain. There were some indications of foetotoxicity at a dose level without maternal toxicity (foetuses with one or more visceral malformations), and the resulting developmental NOAEL was 2.5 mg/kg bw/day
14	<p>Flumioxazin</p> 	N-phenyl phthalamide	103361-09-7	R61, category 1 (2)?? Repr. Cat. 2; R61	0.05	Rat, developmental toxicity study (NOAEL 10 mg/kg bw/d) SF 200	Teratogenic and foetotoxic in absence of maternal toxicity in rat, but not in rabbit Rat : ↓ live fetuses, fetal BW, cardiovascular abnormalities (primarily VSD), wavy ribs, curvature of the scapula, ↓ ossified sacrococcygeal vertebral bodies at 30 mg/kg bw/d

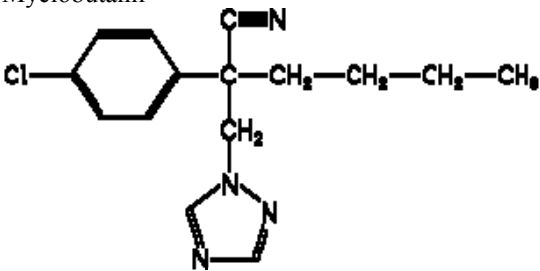
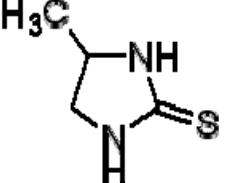
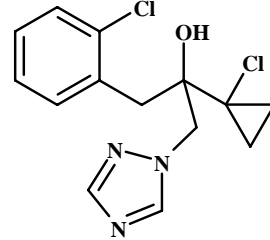
No	Compound	Pesticide class	CAS No	Classification Current EU classification	ARfD mg/kg bw	Basis for ARfD	Comments
15	<p>Fluzilazole</p> 	Triazole	85509-19-9	<p>R61 Category 1 (2) ?? Repr. Cat. 2; R61</p>	0.005	<p>NOAEL 0.5 mg/kg bw/d for Developmental toxicity in rat (vaginal discharge; increased placental weight; increase in rudimentary 7th cervical ribs</p>	<p>JMPR 2007 – ARfD of 0.02 mg/kg bw based on the NOAEL of 2 mg/kg bw/d for skeletal anomalies in an oral developmental toxicity study in rats, SF 100. Info on maternal toxicity would be needed ECB, 1998-Repr Cat. 2; R61: it was noted that serious effects were evident in both rats and rabbits which could not be accounted for either by arguments of non-specific maternal toxicity or species-specificity NOAEL for teratogenicity in rat 50 mg/kg bw/day – at maternally toxic doses specific malformations noted were cleft palate and absent renal papillae.</p>

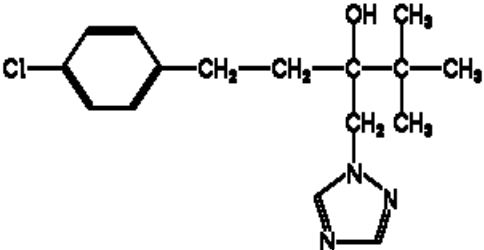
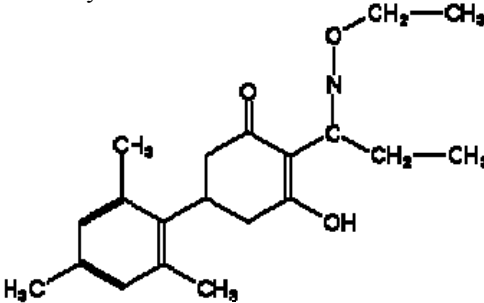
No	Compound	Pesticide class	CAS No	Classification Current EU classification	ARfD mg/kg bw	Basis for ARfD	Comments
16	<p>Glufosinate-ammonium</p>	Phosphic acid	77182-82-2	R61 Category 2 (ISPR) Repr. Cat. 3; R63	0.021 (women child bearing age) 0.045 (general populatio n)	Developmental study rabbit SF 300 Dog studies	<p>ARfD for women of child bearing potential; 6.3 mg/kg bw/day/300 i.e. 0.021 mg/kg bw/day</p> <p>ARfD for the general population; 4.5 mg/kg bw/day/100 i.e. 0.045 mg/kg bw/day</p> <p>Would be interesting on which data classification is based. The severity of the reproductive toxicity was discussed at the Expert Meeting in May 2004. The meeting concluded that there are severe developmental toxicity induced by glufosinate-ammonium seen as pre- and post implantation losses, vaginal bleedings, abortions and dead foetuses not induced by maternal toxicity. However, the meeting also concluded that the underlying mechanism for this could not be identified but that the reduced glutamine synthetase activity might be involved. The meeting agreed that the data was sufficient to conclude on and that no further studies were needed. Furthermore, the meeting agreed with the rapporteur Member State of the proposed classification of glufosinate-ammonium as a Category 2 substance T; R61 "Toxic: may cause harm to the unborn child".</p>

12

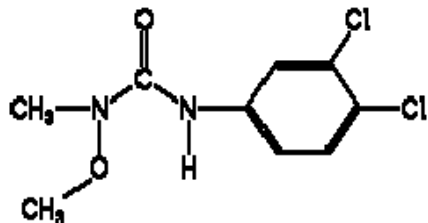
The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a service level agreement between the European Food Safety Authority and the European Community. The present document is published complying with the transparency principle to which the European Food Safety Authority is subject. It may not be considered as an output adopted by EFSA. EFSA reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors

No	Compound	Pesticide class	CAS No	Classification Current EU classification	ARfD mg/kg bw	Basis for ARfD	Comments
17	Hymexazole 	Oxazole	10004-44-1	R63 No classification	0.5	Developmental study rabbit Reduced foetal weight (rat, rabbit), increased post-implantation loss (rabbit), skeletal variations (rat), heart/great vessels malformations (incomplete inferior vena cava, rabbit) Developmental NOAELs: Rat: 100 mg/kg bw/d; Rabbit: 50 mg/kg bw/d	Developmental effects in maternal nontoxic doses: maternal NOAELs: Rat: 500 mg/kg bw/d; Rabbit: 150 mg/kg bw/d Effects at relatively high doses.
18	Mancozeb and maneb common metabolite ETU (ethylene thiourea) 		96-45-7	R61 Category 1 (2)? Repr. Cat. 2; R61	0.05	Rat developmental NOAEL 5 mg/kg bw/d	ETU induced meningoencephalocele, meningorrhagia, meningorrhea, hydrocephalus, obliterated neural canal, abnormal pelvic limb posture with equinovarus, and short or kinky tail after 10 mg/kg or more in all rat experiments (Khera, 1973) Ethylenethiourea (ETU) is a potent teratogen in the rat but not in the mouse or any other species tested (Daston et al, 1989). ARfD 1993 JMPR 0.003 mg/kg bw, NOAEL 0.3 mg/kg bw/d rat developmental skeletal variations

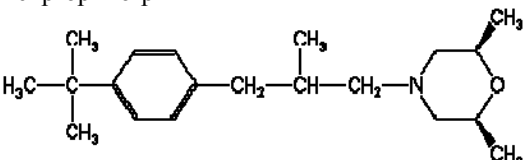
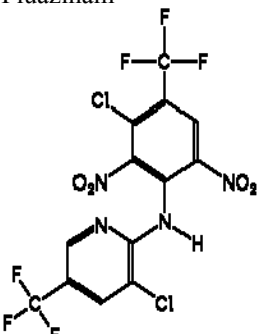
No	Compound	Pesticide class	CAS No	Classification Current EU classification	ARfD mg/kg bw	Basis for ARfD	Comments
19	Myclobutanil 	Triazole	88671-89-0	R63 Category 2 (3)?? Repr. Cat. 3; R63	0.31	Developmental rat NOAEL 31 mg/kg bw/d	Embryotoxic effects – altered viability index with a concomitant increase in resorptions per litter and litters with more than 2 resorptions. Maternal tox? Altered viability index without maternal toxicity The relevant parental NOAEL is 94 mg/kg bw/day, while the relevant developmental NOAEL is 31 mg/kg bw/day.
20	Propineb metabolite PTU (propylene thiourea) 		2122-19-2	Category 2 Repr. Cat. 3; R63	0.003	Developmental toxicity, rat	
21	Metabolite Desthio-prothiconazole 		120983-64-4	R61 category 2 No classification (not discussed at ECHA)	0.01	Developmental NOAEL of 1 mg/kg bw/d (rat, rabbit)	JMPR 2008 ARfD 0.01 mg/kg bw for women of childbearing age; 1 mg/kg bw for the general population.

No	Compound	Pesticide class	CAS No	Classification Current EU classification	ARfD mg/kg bw	Basis for ARfD	Comments
22	Tebuconazole 	Triazole	107534-96-3	R63 Category 2 (3)?? Repr. Cat. 3; R63	0.03	Developmental LOAEL of 10 mg/kg bw/d in mouse teratogenicity study SF 300.	Mice - open eye, runts, cleft palate without maternal toxicity (maternal NOAEL 100 mg/kg bw/d) JMPR 2010 ARfD 0.3 mg/kg bw based on maternal and developmental NOAEL 30 mg/kg bw/d in rats and rabbits
23	Tralkoxydim 	Cyclohexadione	87820-88-0	R63 Category 3 No classification	0.01	Rat developmental study NOAEL 1 mg/kg bw/d	Increased incidence of external/visceral defects (oedema, pale spleen and cleft palate) Maternal NOAEL 3 mg/kg bw/d.

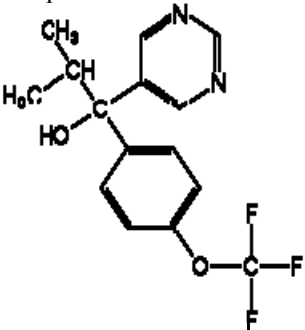
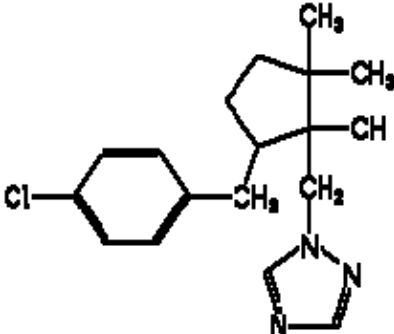
C[C@H]1C(=O)N(c2cc(Cl)cc(Cl)c2)C1=O

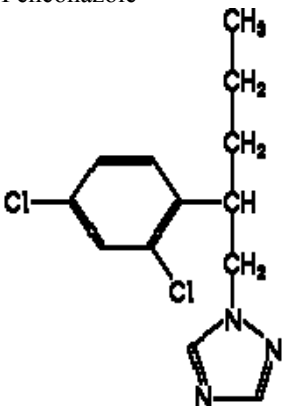
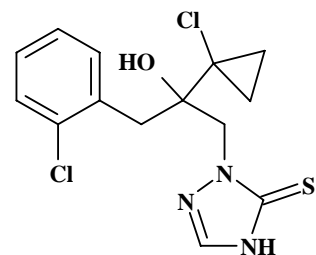
No	Compound	Pesticide class	CAS No	Classification Current EU classification	ARfD mg/kg bw	Basis for ARfD	Comments
26	<p>Linuron</p> 	Urea	330-55-2	R61 Category 1 Repr. Cat. 2; R61	0.03	Rabbit developmental NOAEL 10 mg/kg bw/d SF 300; Fetotoxicity at doses inducing maternal toxicity.	<p>EFSA conclusion not available COM conclusions</p> <p>Data is not convincing.</p> <p>Reproductive malformations in male rat (reduced anogenital distance, retained nipples, hypospadias, malformed epididymis, testis atrophy, ...) Endocrine disruptor</p>

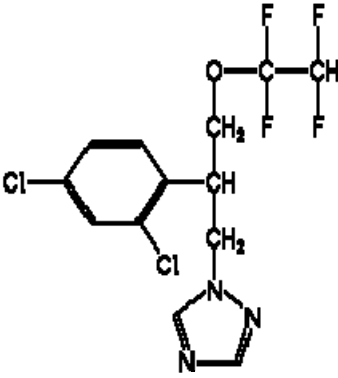
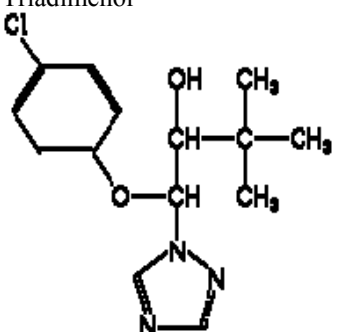
11 “WEAKER” DEVLEOPMENTAL POSITIVES SELECTED BY EFSA

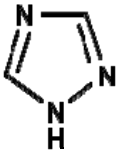
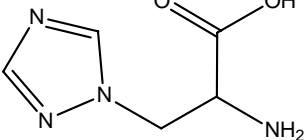
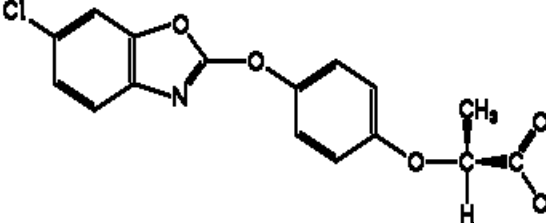
No	Compound	Pesticide class	CAS No	Classification Current EU classification	ARfD mg/kg bw	Basis for ARfD	Comments
27	Fenpropimorph 	Morpholine	67564-91-4	R63 Category 2 (3)?? Repr. Cat. 3; R63	0.03	Rabbit developmental (AF: 500)	Increased incidence of malformations in rats (cleft palates) and rabbits (skeleton malformations : shortening of long bones, position anomalies of the limbs, cleft palate, fused sternbrae) at maternal toxic doses
28	Fluazinam 	Phenylpyri dinamine	79622-59-6	Proposed R63 No classification	0.07	Developmental rabbit NOAEL 7 mg/kg bw/day (LOAEL 12 mg/kg bw) based on litter losses, skeletal abnormalities and placental abnormalities	Foetal: postimplantation loss↑ (rat and rabbit), ossification incomplete (rat and rabbit), fetal and placental weight ↓ (rat), significant abnormalities at maternal toxic doses (cleft palate in rats, placental and skeletal abnormalities in rabbits) Specific malformations only at maternal toxic doses, only in rat but other effects start already at lower dose

The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a service level agreement between the European Food Safety Authority and the European Community. The present document is published complying with the transparency principle to which the European Food Safety Authority is subject. It may not be considered as an output adopted by EFSA. EFSA reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors

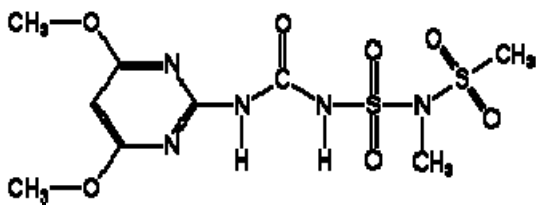
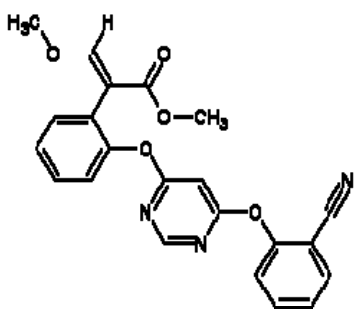
No	Compound	Pesticide class	CAS No	Classification Current EU classification	ARfD mg/kg bw	Basis for ARfD	Comments
29	Flurprimidol 	Pyrimidinyl carbinol	56425-91-3	Proposed R63 No classification	0.09	Developmental, rabbit supported by developmental, rat study	Rat: Malformations (microphthalmia), increased incidence of variants and abnormal fetuses Rabbit: Increased incidence of variants and abnormal fetuses Maternal NOAELs: Rat: 10 mg/kg bw/day; Rabbit: 9.0 mg/kg bw/day Developmental NOAELs: Rat: 10 mg/kg bw/day; Rabbit: 9.0 mg/kg bw/day
30	Metconazole 	Triazole	125116-23-6	R63 Category 2 (3)?? Repr. Cat. 3; R63	0.01	Developmental rabbit NOAEL 4 mg/kg bw/d, SF 400	Increased hydrocephaly incidence (maternal NOAEL 4 mg/kg bw/d) in rabbits. Rat: embryo/foetotoxicity (↑ post-implantation loss, ↓ litter size, foetal weight, ↑ placental weight) and skeletal ossification variations at the highest dose tested, ↑ incidence of bilateral hydronephrosis. Malformations may depend on purity of the substance (no identif. of a specific impurity)

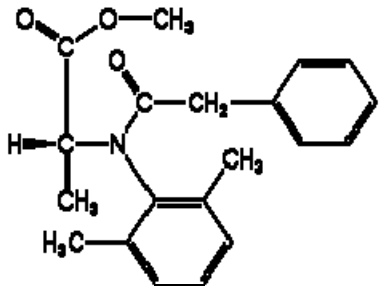
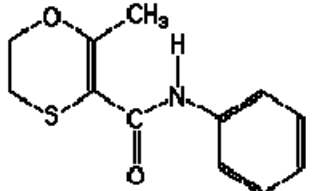
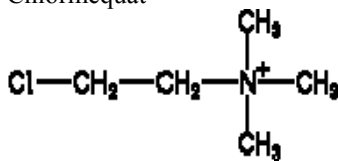
No	Compound	Pesticide class	CAS No	Classification Current EU classification	ARfD mg/kg bw	Basis for ARfD	Comments
31	Penconazole 	Triazole	66246-88-6	R63 category 3 No classification	0.5	Developmental rabbit (maternal NOAEL 50 mg/kg bw/d)	R63 Reduced foetal weightt., delayed ossification and skeletal variations, bilateral microphthalmia in one rabbit strain, all findings at maternal toxic doses
32	Prothioconazole 	Triazole	178928-70-6	R63 category 3 No classification (not finalized in ISPRA, not yet discussed at ECHA)	0.2	Developmental rat NOAEL of 20 mg/kg bw/d (combined from 2 studies)	R63 proposed based on increased incidences of microphthalmia (at 1000 mg/kg, maternal toxic dose) in one out of two rat strains tested JMPR 2008 ARfD 0.8 mg/kg bw for women of childbearing age; unnecessary for the general population

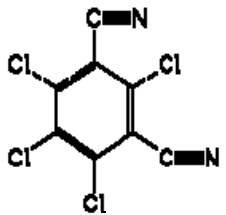
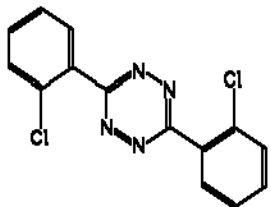
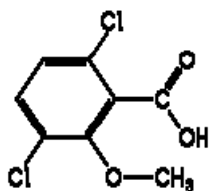
No	Compound	Pesticide class	CAS No	Classification Current EU classification	ARfD mg/kg bw	Basis for ARfD	Comments
33	Tetraconazole 	Triazole	112281-77-3	R63 No classification	0.05	Maternal NOAEL 5 mg/kg bw/d in rat developmental study	R63 Extra ribs, hydrourether and hydronephrosis (developmental NOAEL 22.5 mg/kg bw/d in rats).
34	Triadimenol 	Triazole	55219-65-3	R63? No classification	0.05	Overall NOAEL rat chronic study + acute and subchronic neurotoxicity study + multigeneration study, SF 100	Rats: ↑ incidence of extra ribs and ↑ placental weight at maternally toxic dose (↓ bodyweight gains). Cleft palate at high dose level Rabbits: ↑ post-implantation losses, ↓ litter size, ↓ foetal weight, ↑ abnormal or incomplete ossification at maternally toxic dose (↓ bodyweight gains and food consumption).

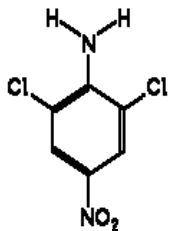
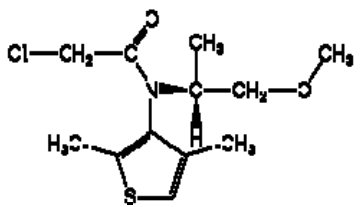
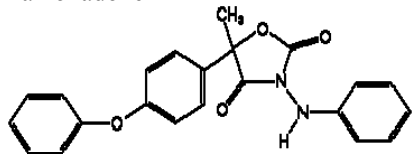
No	Compound	Pesticide class	CAS No	Classification Current EU classification	ARfD mg/kg bw	Basis for ARfD	Comments
35	Triazole common metabolite 1,2,4-triazole 	Triazole	288-88-0	R63 category 3 Repr. Cat. 3; R63	0.06	Rat developmental SF 500 NOAEL 30 mg/kg bw/d SF 500 (classified R62, 63, limited data base, data gaps for DNT, endocrine disruption)	JMPR 2008 ARfD 0.3 mg/kg bw Rabbit, study of developmental toxicity (maternal and developmental NOAELs 30 mg/kg bw/d)
36	Triazole common metabolite triazole alanine 	Triazole	10109-05-4	No classification	0.1	Rat developmental SF 1000 based on NOAEL 100 mg/kg bw/d, SF 1000 (limited data base, no rabbit developmental study)	JMPR 2008 ARfD not necessary incl. triazole acetic acid
37	Fenoxaprop-P 	Aryl oxy phenoxy propionate	113158-40-0	Proposed R63 No classification	0.1	Rat developmental	Fetal toxicity: delayed ossification particularly in cranial bones (rat); NOAEL 10 mg/kg (maternal toxicity NOAEL 32 mg/kg)

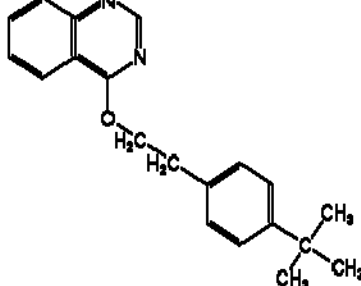
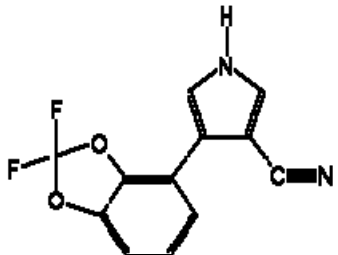
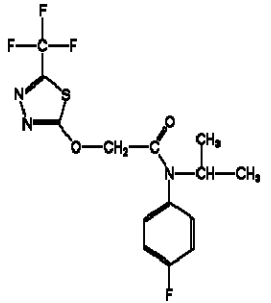
39 DEVELOPMENTAL NEGATIVES SELECTED BY EFSA

No	Compound/ chemical structure	Chemical class	CAS No	Comments
1	Amidosulfuron 	Pyrimidinylsulfonylurea	120923-37-7	
2	Azoxystrobin 	strobilurin	131860-33-8	

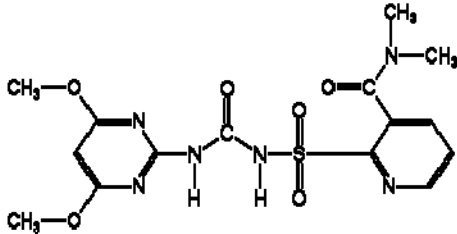
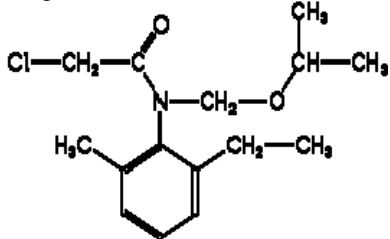
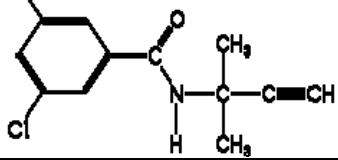
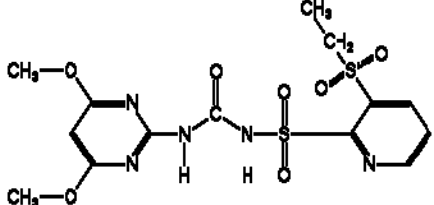
No	Compound/ chemical structure	Chemical class	CAS No	Comments
3	Benalaxyl-M 	anilide	98243-83-5	
4	Carboxin 	anilide, oxathin	5234-68-4	
5	Chlormequat 	growth retardants	7003-89-6	

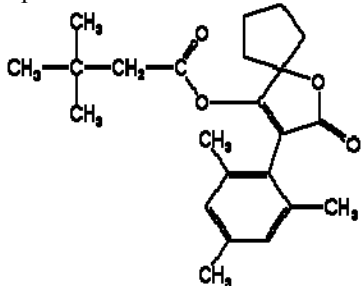
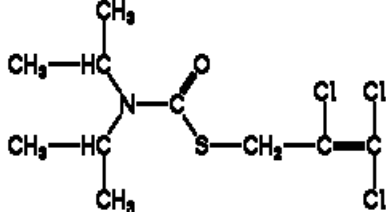
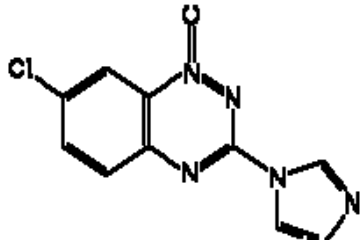
No	Compound/ chemical structure	Chemical class	CAS No	Comments
6	<p>Chlorothalonil</p> 	aromatic	1897-45-6	
7	<p>Clofentezine</p> 	tetrazine	74115-24-5	
8	<p>Dicamba</p> 	benzoic acid	1918-00-9	

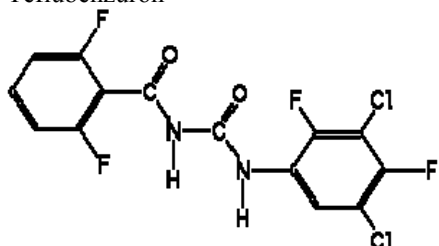
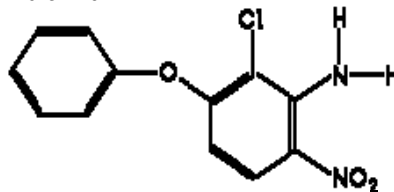
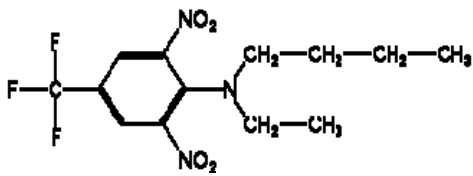
No	Compound/ chemical structure	Chemical class	CAS No	Comments
9	Dicloran 	aromatic	99-30-9	
10	Dimethenamide-P 	amide	163515-14-8	
11	Famoxadone 	Dicarboximide, oxazole	131807-57-3	

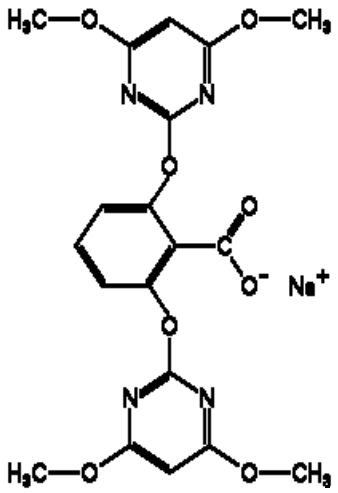
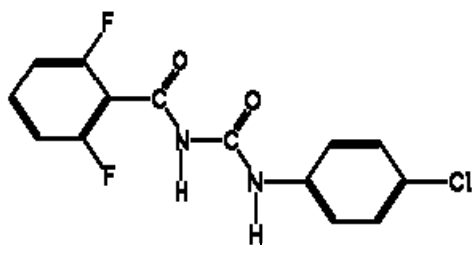
No	Compound/ chemical structure	Chemical class	CAS No	Comments
12	Fenazaquin 	unclassified acaricide	120928-09-8	
13	Fludioxonil 	pyrrole	131341-86-1	
14	Flufenacet 	Anilide	142459-58-3	

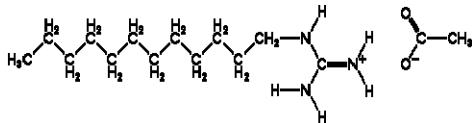
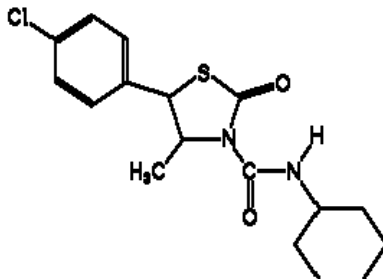
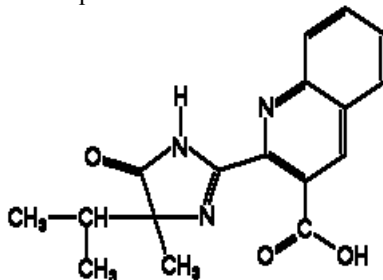
No	Compound/ chemical structure	Chemical class	CAS No	Comments
15	Fosthiazate 	organophosphate	98886-44-3	
16	Metaflumizone 	unclassified insecticide	139968-49-3	
17	Metrafenone 	Aryl phenyl ketone	220899-03-6	

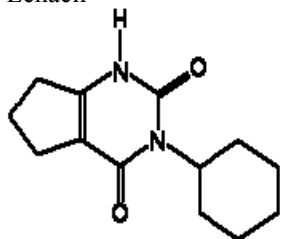
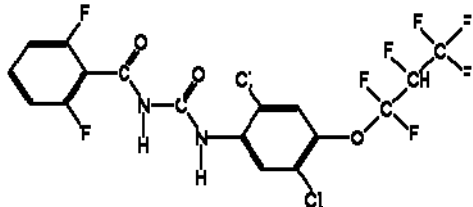
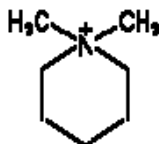

No	Compound/ chemical structure	Chemical class	CAS No	Comments
18	Nicosulfuron 	pyrimidinylsulfonylurea	111991-09-4	
19	Propisochlor 	chloroacetanilide	86763-47-5	
20	Propyzamide 	amide	23950-58-5	
21	Rimsulfuron 	pyrimidinylsulfonylurea	122931-48-0	

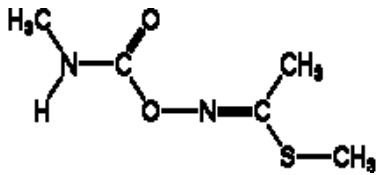
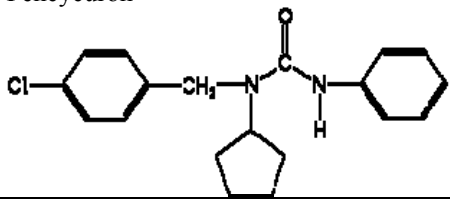
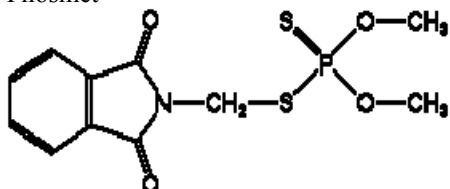
No	Compound/ chemical structure	Chemical class	CAS No	Comments
22	<p>Spiromesifen</p> 	Tetronic acid	283594-90-1	
23	<p>Tri-allate</p> 	thiocarbamate	2303-17-5	
24	<p>Triazoxide</p> 	imidazole	72459-58-6	

No	Compound/ chemical structure	Chemical class	CAS No	Comments
25	Teflubenzuron 	benzoylphenylurea	83121-18-0	
26	Aclonifen 	nitrophenyl ether herbicide	74070-46-5	EFSA conclusion
27	Benfluralin 	dinitroaniline herbicide	1861-40-1	EFSA conclusion

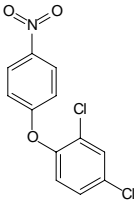
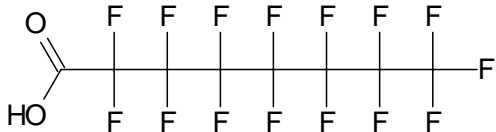
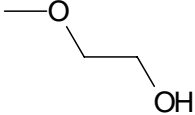
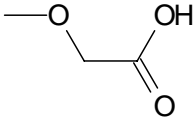
No	Compound/ chemical structure	Chemical class	CAS No	Comments
28	<p>Bispyribac sodium</p> 	pyrimidinyloxybenzoic acid herbicide	125401-92	EFSA conclusion
29	<p>Diflubenzuron</p> 	benzoylphrnylurea insecticide	35367-38-5	EFSA conclusion

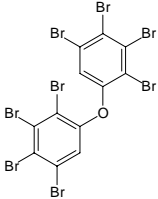
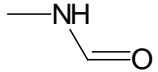
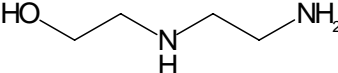
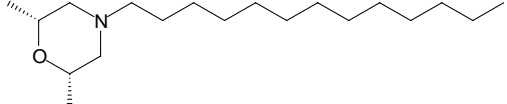
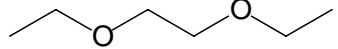
No	Compound/ chemical structure	Chemical class	CAS No	Comments
30	Dodine 	aliphatic nitrogen fungicide	2439-10-3	EFSA conclusion
31	Hexythiazox 	thiazolidine acaricide	78587-05-0	EFSA conclusion
32	Imazaquin 	imidazolinone herbicide	81335-37-7	EFSA conclusion

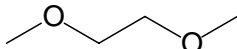
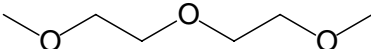
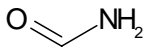
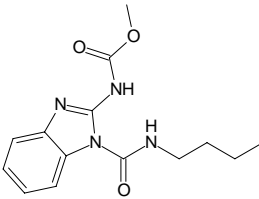
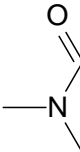
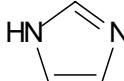
No	Compound/ chemical structure	Chemical class	CAS No	Comments
33	<p>Lenacil</p> 	uracil herbicide	2164-08-1	EFSA conclusion
34	<p>Lufenuron</p> 	benzoylphrnylurea insecticide	103055-07-8	EFSA conclusion
35	<p>Mepiquat</p> 	growth inhibitor	15302-91-7	EFSA conclusion
36	<p>Metamitron</p> 	triazinone herbicide	41394-05-2	EFSA conclusion

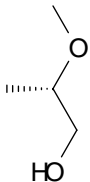
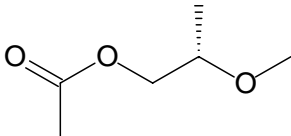
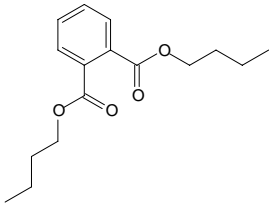
No	Compound/ chemical structure	Chemical class	CAS No	Comments
37	Methomyl 	oxime carbamate insecticide	16752-77-5	EFSA conclusion
38	Pencycuron 	urea fungicide	66063-05-6	EFSA conclusion
39	Phosmet 	organothiophosphate	732-11-6	EFSA conclusion

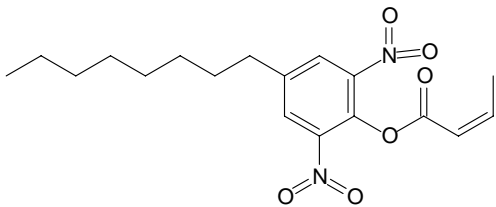
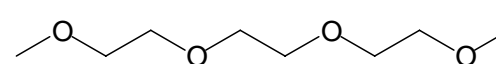
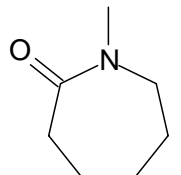
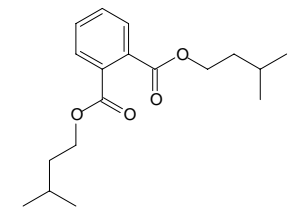
B. RIVM dataset of chemicals classified for developmental toxicity

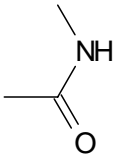
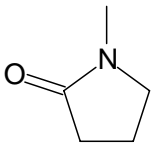
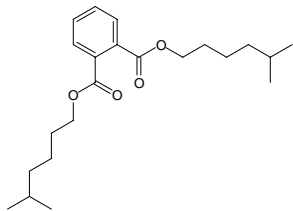
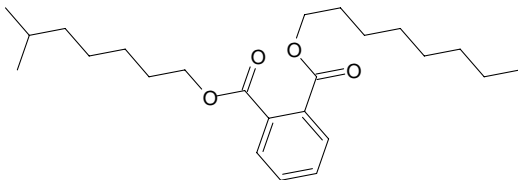
No	2D STRUCTURE	CAS	NAME	EU classification
1		1836-75-5	Nitrofen (ISO), 2,4-Dichlorophenyl 4-nitrophenyl ether	R61
2		3825-26-1	Ammonium pentadecafluorooctanoate	R61
3		109-86-4	2-Methoxyethanol	R61
4		625-45-6	Methoxyacetic acid	R61

No	2D STRUCTURE	CAS	NAME	EU classification
5		32536-52-0	Octabromobiphenyl ether	R61
6		123-39-7	N-Methylformamide	R61
7		111-41-1	2-(2-Aminoethylamino)ethanol	R61
8		24602-86-6	Tridemorph	R61
9		629-14-1	1,2-Diethoxyethane	R61

No	2D STRUCTURE	CAS	NAME	EU classification
10		110-71-4	1-2-Dimethoxyethane	R61
11		111-96-6	Bis(2-Methoxyethyl)ether	R61
12		75-12-7	Formamide	R61
13		17804-35-2	Benomyl	R61
14		68-12-2	N,N-Dimethylformamide, dimethyl formamide	R61
15		288-32-4	Imidazole, N, N'-1,2-Ethenediyl-methanimidamide	R61

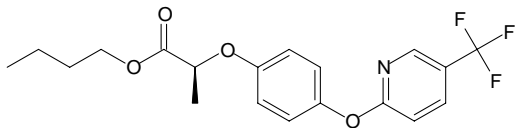
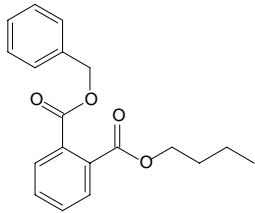
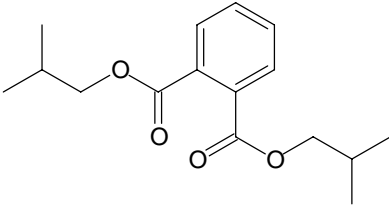
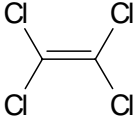
No	2D STRUCTURE	CAS	NAME	EU classification
16		1589-47-5	2-Methoxypropanol	R61
17		70657-70-4	2-Methoxypropyl acetate	R61
18		84-74-2	Dibutyl phthalate	R61

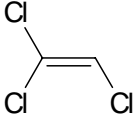
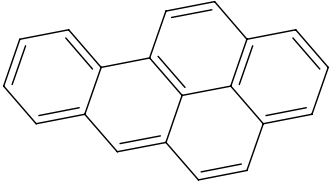
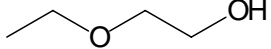
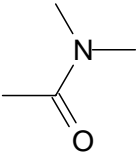
No	2D STRUCTURE	CAS	NAME	EU classification
19		39300-45-3	Dinocap	R61
20		112-49-2	1,2-bis(2-Methoxyethoxy)ethane	R61
21		2556-73-2	N-Methylcaprolactam	R61
22		605-50-5	di-n-Pentylphthalate	R61

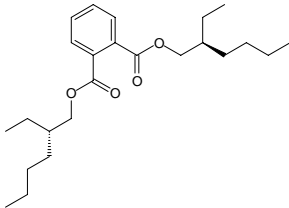
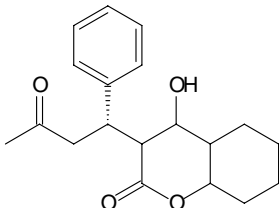
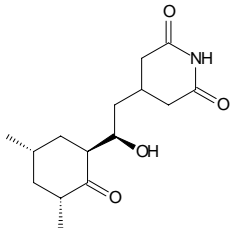
No	2D STRUCTURE	CAS	NAME	EU classification
23		79-16-3	N-Methylacetamide	R61
24		872-50-4	1-Methyl-2-pyrrolidone	R61
25		71888-89-6	1,2-Benzenedicarboxylic acid; di-C6-8-branched alkylesters, C7-rich	R61
26		68515-42-4	1,2-Benzenedicarboxylic acid, di-C7-11-branched and linear alkylesters	R61

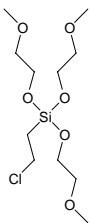
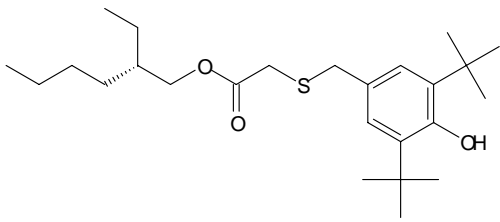
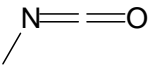
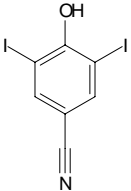
41

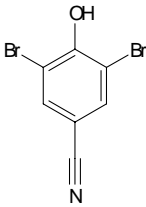
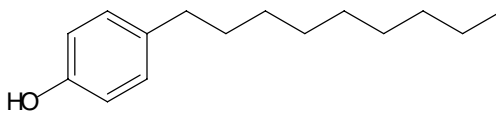
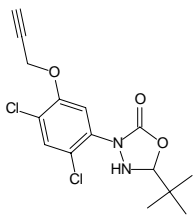
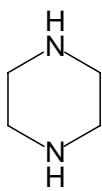
The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a service level agreement between the European Food Safety Authority and the European Community. The present document is published complying with the transparency principle to which the European Food Safety Authority is subject. It may not be considered as an output adopted by EFSA. EFSA reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors

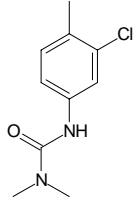
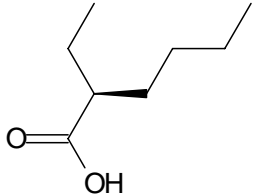
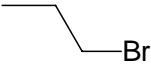
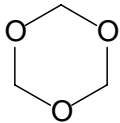
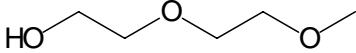
No	2D STRUCTURE	CAS	NAME	EU classification
27		69806-50-4	Fuazifop-butyl	R61
28		85-68-7	Butyl benzyl phthalate	R61
29		84-69-5	Diisobutyl phthalate	R61
30		127-18-4	Tetrachloroethylene	R61

No	2D STRUCTURE	CAS	NAME	EU classification
31		79-01-6	Trichloroethylene, Trichloroethene	63
32		50-32-8	Benzo[a]pyrene	R61
33		110-80-5	2-Ethoxyethanol	R61
34		127-19-5	N,N-Dimethylacetamide	R61

No	2D STRUCTURE	CAS	NAME	EU classification
35		117-81-7	Bis(2-ethylhexyl) phthalate (DEHP)	R61
36		81-81-2	Warfarin	R61
37		66-81-9	Cycloheximide	R61

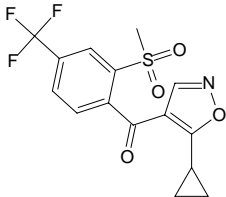
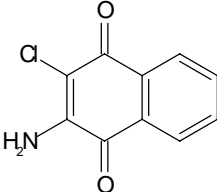
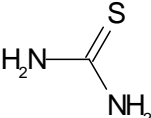
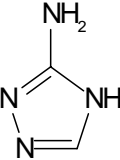
No	2D STRUCTURE	CAS	NAME	EU classification
38		37894-46-5	Etacelasil, 6-(2-Chloroethyl)-6-(2-methoxyethoxy)-2,5,7,10-tetraoxa-6-silaundecane	R61
39		80387-97-9	2-Ethylhexyl[[[3,5-bis(1,1-dimethylethyl)-4-hydroxyphenyl]methyl]thio]acetate	R61
40		624-83-9	Methyl isocyanate	R63
41		1689-83-4	Ioxynil (ISO) and its salts, 4-Hydroxy-3,5-diiodobenzonitrile	R63

No	2D STRUCTURE	CAS	NAME	EU classification
42		1689-84-5	Bromoxynil (ISO) and its salts, 3,5-Dibromo-4-hydroxybenzonitrile, Bromoxynil phenol	R63
43		25154-52-3	Nonylphenol	R63
44		39807-15-3	Oxadiargyl	R63
45		110-85-0	Piperazine	R63

No	2D STRUCTURE	CAS	NAME	EU classification
46		15545-48-9	Chlorotoluron	R63
47		149-57-5	2-Ethylhexanoic acid	R63
48		106-94-5	1-Bromopropane	R63
49		110-88-3	1,3,5-Trioxan, Trioxymethylene	R63
50		111-77-3	2-(2-Methoxyethoxy)ethanol	R63

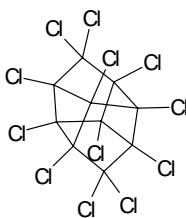
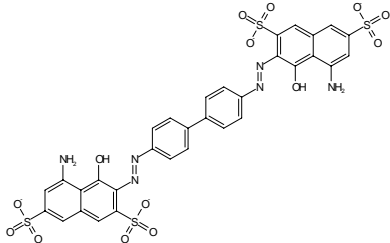
47

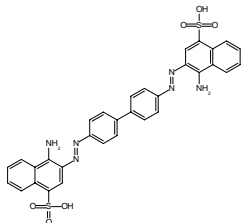
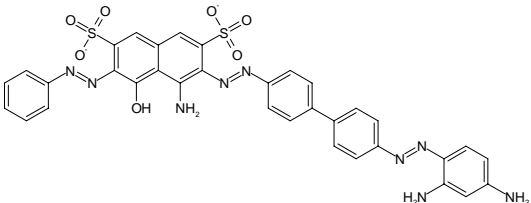
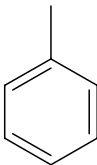
The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a service level agreement between the European Food Safety Authority and the European Community. The present document is published complying with the transparency principle to which the European Food Safety Authority is subject. It may not be considered as an output adopted by EFSA. EFSA reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors

No	2D STRUCTURE	CAS	NAME	EU classification
51		141112-29-0	Isoxaflutole	R63
52		2797-51-5	Quinoclamine	R63
53		62-56-6	Thiourea, Thiocarbamide	R63
54		61-82-5	Amitrole (ISO), 1,2,4-Triazol-3-ylamine	R63

48

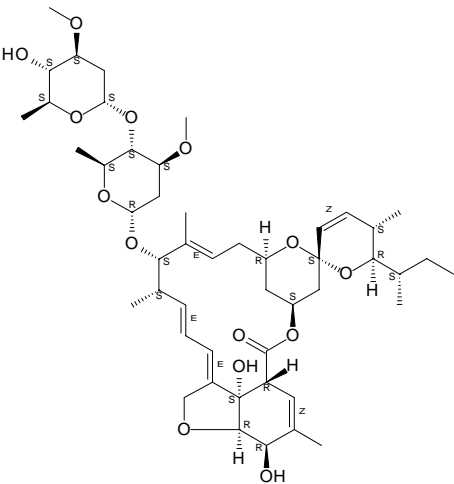
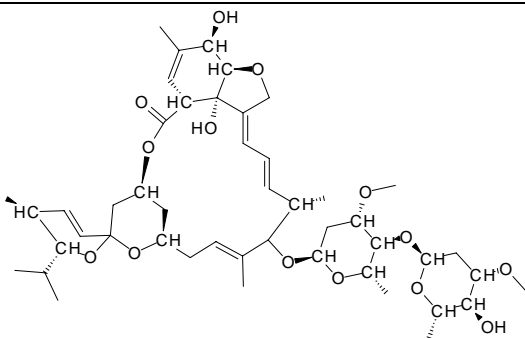
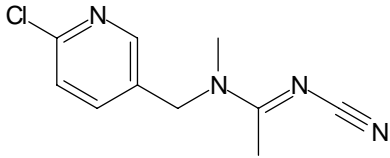
The present document has been produced and adopted by the bodies identified above as author(s). This task has been carried out exclusively by the author(s) in the context of a service level agreement between the European Food Safety Authority and the European Community. The present document is published complying with the transparency principle to which the European Food Safety Authority is subject. It may not be considered as an output adopted by EFSA. EFSA reserves its rights, view and position as regards the issues addressed and the conclusions reached in the present document, without prejudice to the rights of the authors

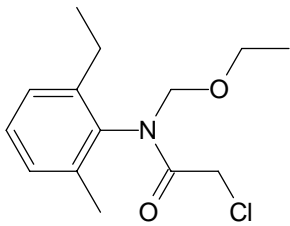
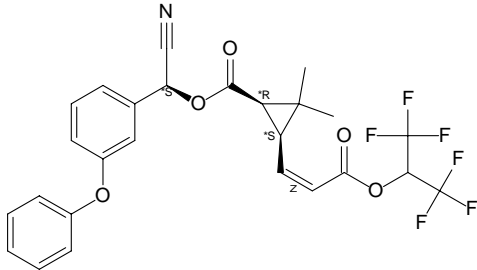
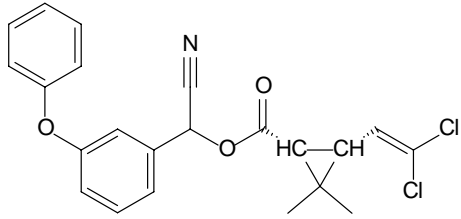
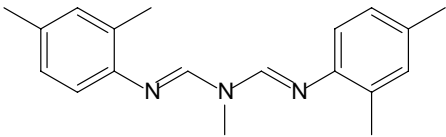
No	2D STRUCTURE	CAS	NAME	EU classification
55		2385-85-5	Mirex, Dodecachloropentacyclo[5.2.1.0.2,6.0.3,9.0.5,8]decane	R63
56		2602-46-2	C.I. Direct Blue 6, Tetrasodium 3,3'-[[1,1'-biphenyl]-4,4'-diylbis(azo)]bis[5-amino-4-hydroxynaphthalene-2,7-disulphonate]	R63

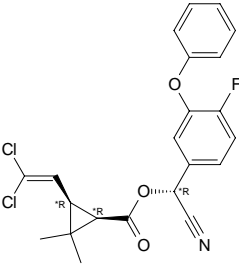
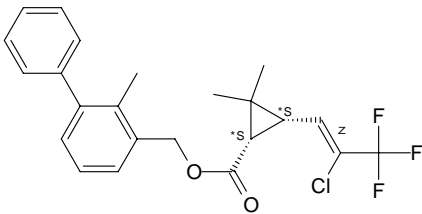
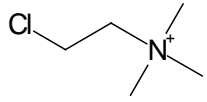
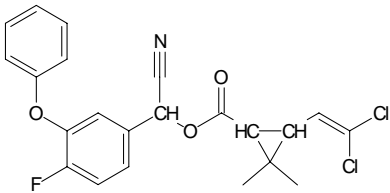
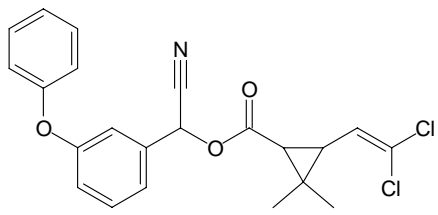
No	2D STRUCTURE	CAS	NAME	EU classification
57		573-58-0	C.I. Direct Red 28, Disodium 3,3'-[[1,1'-biphenyl]-4,4'-diylbis(azo)]bis(4-aminonaphthalene-1-sulphonate)	R63
58		1937-37-7	C.I. Direct Black 38 Disodium 4-amino-3-[[4'-[(2,4-diaminophenyl)azo]][1,1'-biphenyl]-4-yl]azo]-5-hydroxy-6-(phenylazo)naphtalene-2,7-disulphonate	R63
59		108-88-3	Toluene	R63

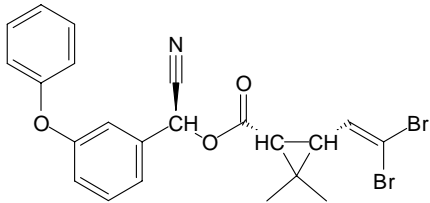
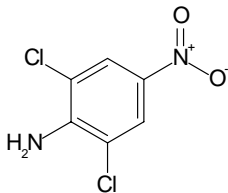
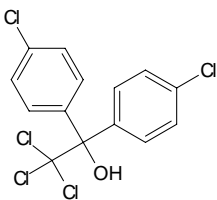
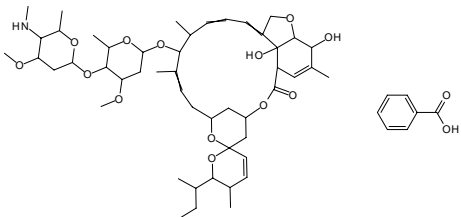
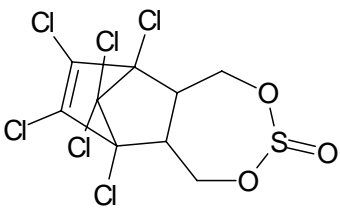
C. Neurotoxicity dataset (EFSA set of pesticides)

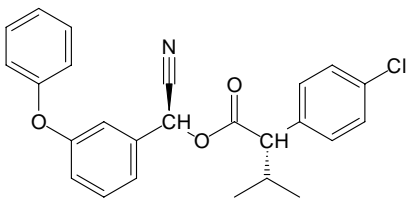
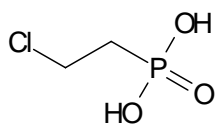
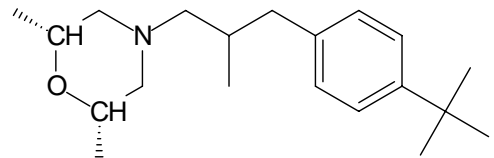
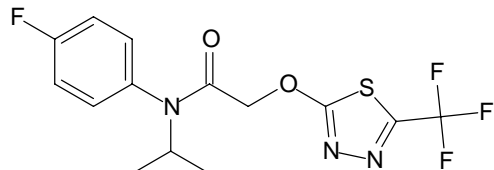
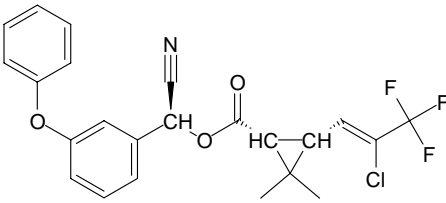
42 NEUROTOXICITY POSITIVES SELECTED BY EFSA

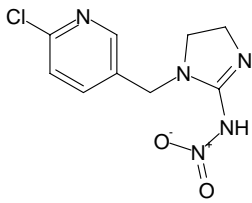
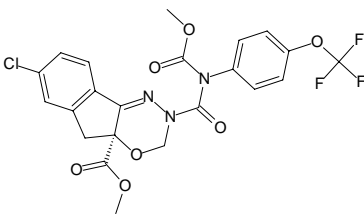
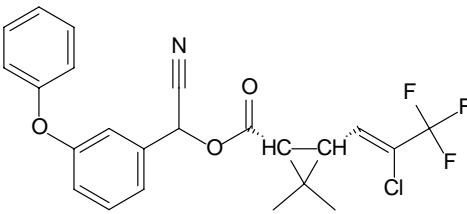
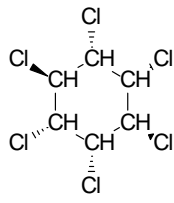
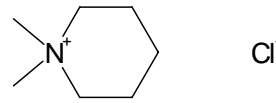
No	Compound / 2D Structure	Pesticide class	CAS No
1a	 <p>avermectin B1a</p>	antibiotic, avermectin, milbemectin	71751-41-2
1b	 <p>avermectin B1b</p>	antibiotic, avermectin, milbemectin	71751-41-2
2	 <p>acetamiprid</p>	neonicotinoide	135410-20-7

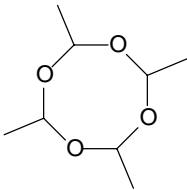
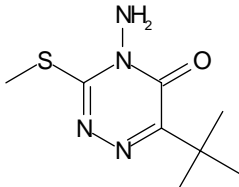
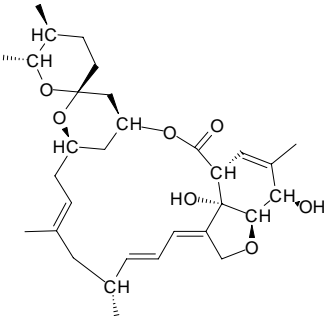
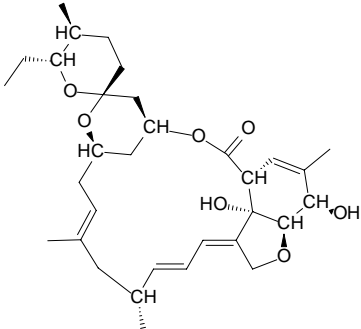
No	Compound / 2D Structure	Pesticide class	CAS No
3	 acetochlor	Acetamide, chloroacetanilide	34256-82-1
4	 acrinathrin	pyrethroid	101007-06-1
5	 a-cypermethrin	pyrethroid	67375-30-8
6	 amitraz	amidine	33089-61-1

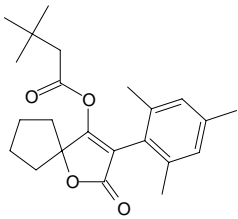
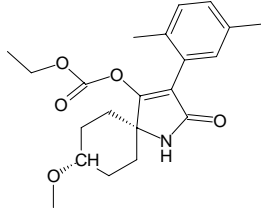
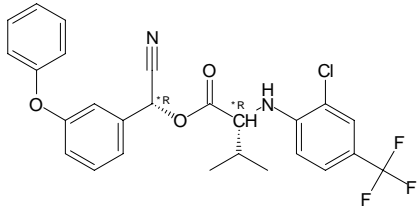
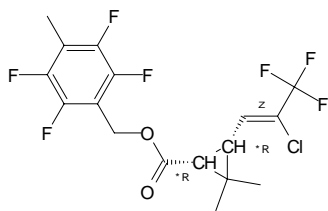
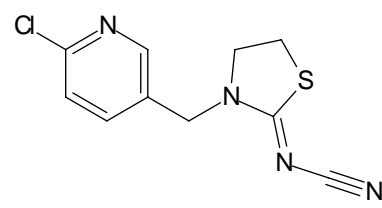
No	Compound / 2D Structure	Pesticide class	CAS No
7	 b-cyfluthrin (1R,3R,αR)	pyrethroid	68359-37-5
8	 bifenthrin (1S,3S)	pyrethroid	82657-04-3
9	 chlormequat	quaternary ammonium	7003-89-6
10	 cyfluthrin	pyrethroid	68359-37-5
11	 cypermethrin	pyrethroid	52315-07-8

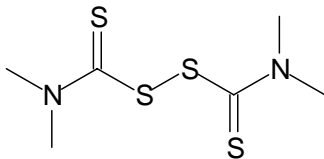
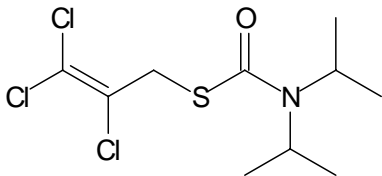
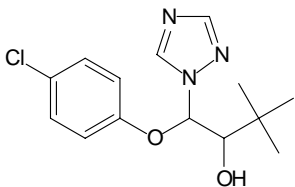
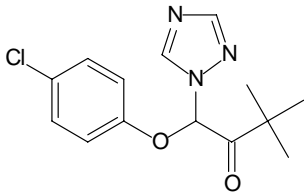
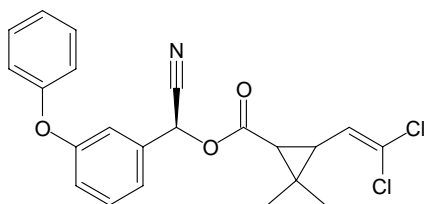
No	Compound / 2D Structure	Pesticide class	CAS No
12	 deltamethrin	pyrethroid	52918-63-5
13	 dichloran	chlorinated nitroaniline	99-30-9
14	 dicofol	organochlorine	115-32-2
15	 emamectin benzoate	antibiotic, avermectin, milbemectin	137512-74-4
16	 endosulfan	organochlorine	115-29-7

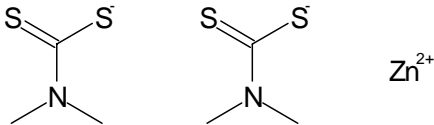
No	Compound / 2D Structure	Pesticide class	CAS No
17	 esfenvalerate	pyrethroid	66230-04-4
18	 ethephon	ethylene generator	16672-87-0
19	 fenpropimorph	morpholine	67306-03-0,
20	 flufenacet	oxyacetamide	142459-58-3
21	 gamma-cyhalothrin	pyrethroid	76703-62-3

No	Compound / 2D Structure	Pesticide class	CAS No
22	 imidacloprid	neonicotinoide	105827-78-9
23	 indoxacarb	oxadiazines	144171-61-9
24	 lambda-cyhalothrin	pyrethroid	91465-08-6
25	 lindane	organochlorine	58-89-9
26	 mepiquat chloride	piperidine, quaternary ammonium	24307-26-4

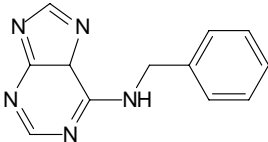
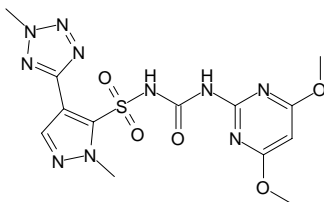
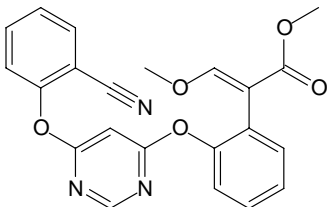
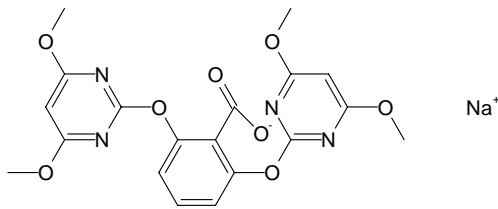
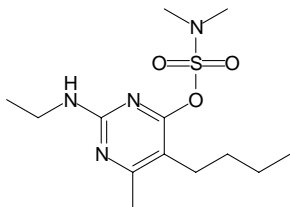
No	Compound / 2D Structure	Pesticide class	CAS No
27	 metaldehyde	acetaldehyde	9002-91-9
28	 metribuzin	triazine	21087-64-9
29a	 milbemectin A3	antibiotic, avermectin, milbemectin	51596-10-2
29b	 milbemectin A4	antibiotic, avermectin, milbemectin	51596-11-3

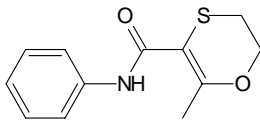
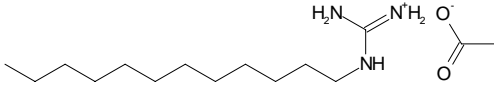
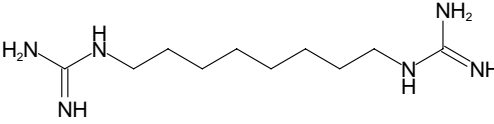
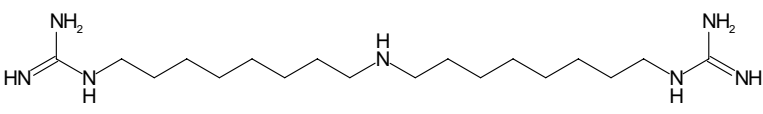
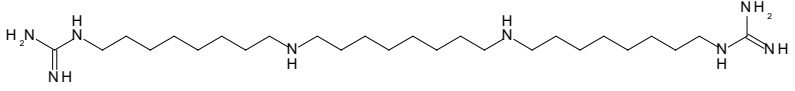
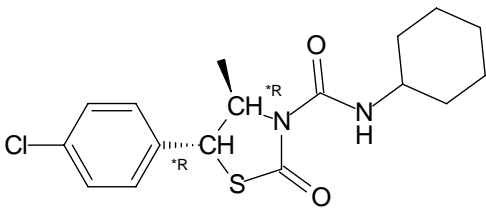
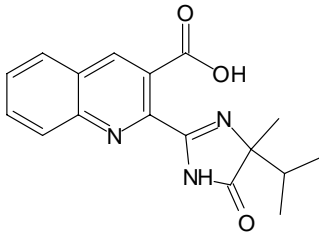
No	Compound / 2D Structure	Pesticide class	CAS No
30	 spiromesifen	tetronic tetramic derivative and acid	283594-90-1
31	 spirotetramat	tetronic tetramic derivative and acid	203313-25-1
32	 tau-fluvalinate (R-cyano)	pyrethroid	102851-06-9
33	 tefluthrin (Z-(1R,3R))	pyrethroid	79538-32-2
34	 thiacloprid	neonicotinoide	111988-49-9

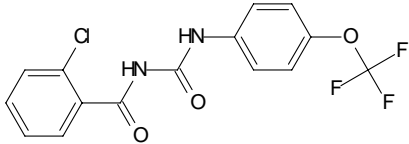
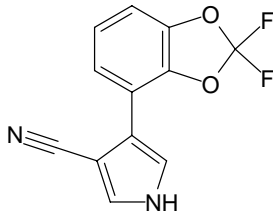
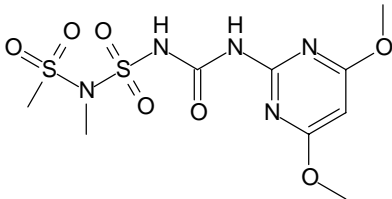
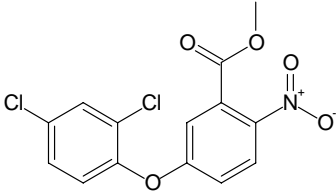
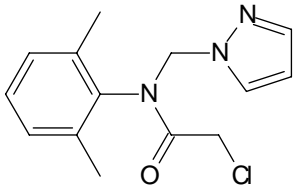
No	Compound / 2D Structure	Pesticide class	CAS No
35	 thiram	dithiocarbamate	137-26-8
36	 tri-allate	thiocarbamate, organochlorine	2303-17-5
37	 triadimenol	triazole, conazole	55219-65-3
38	 triadimefon	triazole, conazole	43121-43-3
39	 zeta-cypermethrin	pyrethroid	52315-07-8

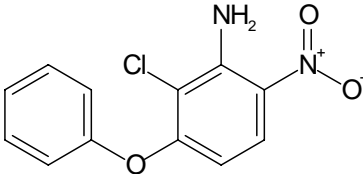
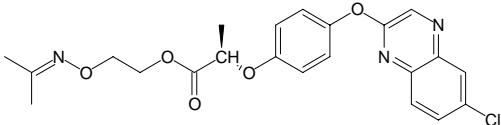
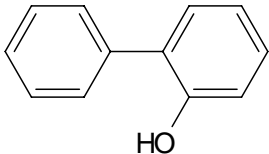
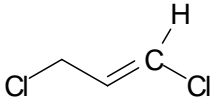
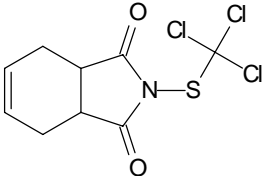
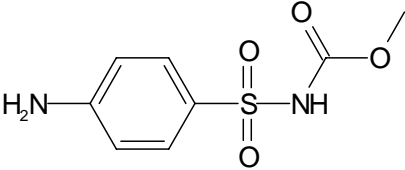
No	Compound / 2D Structure	Pesticide class	CAS No
40	 ziram	dithiocarbamate	137-30-4

23 NEUROTOXICITY NEGATIVES SELECTED BY EFSA

No	Compound / 2D Structure	Pesticide class	CAS No
1	 6-benzyladenine	cytokinins	1214-39-7
2	 azimsulfuron	pyrazole herbicide, pyrimidinylsulfonyl urea herbicide	120162-55-2
3	 azoxystrobin	strobilurin fungicides	131860-33-8
4	 bispyribac sodium	pyrimidinylbenzoic acid herbicide	125401-92-5
5	 bupirimate	pyrimidine fungicide	41483-43-6

No	Compound / 2D Structure	Pesticide class	CAS No
6	 carboxin	anilide fungicides, oxathiin fungicides	5234-68-4
7	 dodine	aliphatic nitrogen fungicides	2439-10-3
8a	 guazatine n=0	aliphatic nitrogen fungicide	108173-90-6
8b	 guazatine n=1	aliphatic nitrogen fungicide	108173-90-6
8c	 guazatine n=2	aliphatic nitrogen fungicide	108173-90-6
9	 hexythiazox	mite regulators, thiazolidine acaricides	78587-05-0
10		imidazolinone herbicides	81335-37-7

No	Compound / 2D Structure	Pesticide class	CAS No
	imazaquin		
11	 triflumuron	benzoylphenylurea chitin synthesis inhibitors	64628-44-0
12	 fludioxonil	pyrrole fungicides	131341-86-1
13	 amidosulfuron	pyrimidinylsulfonylurea herbicides	120923-37-7
14	 bifenox	nitrophenyl ether herbicides	42576-02-3
15	 metazachlor	chloroacetanilide herbicides, pyrazole herbicides	67129-08-2

No	Compound / 2D Structure	Pesticide class	CAS No
16	 acлонifen	nitrophenyl ether herbicides	74070-46-5
17	 propaquizafop	aryloxyphenoxypropionic herbicides	111479-05-1
18	 2-phenylphenol	unclassified fungicides	90-43-7
19	 (1E,Z)-1,3-dichloropropene	unclassified nematicides	542-75-6
20	 captane	phthalimide	133-06-2
21		carbamate herbicide, sulfonamide herbicide	3337-71-1

	QSAR Analysis of Developmental Toxicity and Neurotoxicity
--	-----------------------------------------------------------

No	Compound / 2D Structure	Pesticide class	CAS No
	asulam		

D. Results of Leadscape structural fragment analysis

Substructure	PPP	EFSA neurotoxicity test set	EFSA developmental test set (extended)	ToxRef test set	CAESAR training	Leadscape training
Amino acids	25	1	3	9	40	186
Bases, nucleosides	0	0	1	2	5	45
Benzenes	486	39	81	230	199	1094
Carbocycles	14	2	1	7	32	68
Carbohydrates	5	0	0	0	10	73
Elements	91	1	5	46	2	64
Functional groups	766	65	133	355	290	1818
alcohol	101	8	26	29	107	600
aldehyde	13	0	0	0	2	13
alkene	165	24	14	60	63	423
alkyne	10	0	4	8	9	27
allene	1	0	0	0	0	0
amidine	17	4	3	12	5	43
amines	174	27	46	133	179	1042
azide	0	0	0	0	0	0
boron groups	0	0	0	1	0	0
carbamate	32	5	6	37	6	58
carbonyl	380	40	69	222	187	1070
carboxamide	82	7	25	55	85	385
carboxylate	112	23	14	68	27	248
carboxylic acid	99	4	10	31	59	282
ether	233	34	48	124	55	531
guanidine	11	4	1	5	3	31
halide	354	37	59	197	66	438
hydrazine	34	2	5	25	7	59
hydroxylamine	25	1	2	14	5	23
imine	3	0	0	1	9	47

Substructure	PPP	EFSA neurotoxicity test set	EFSA developmental test set (extended)	ToxRef test set	CAESAR training	Leadscope training
iminomethyl	189	10	5	26	20	124
isocyanate	0	0	1	0	0	1
ketone	46	1	6	19	37	240
mercaptan	2	0	0	0	1	7
misc nitrogen groups	16	4	4	11	4	49
misc oxygen groups	3	1	0	0	1	9
misc sulfur groups	11	3	0	6	1	16
nitrile	34	13	9	22	0	23
nitro	42	4	5	23	9	67
nitroso	0	0	0	0	0	12
organometal	0	0	0	4	0	0
phosphorous groups	17	1	0	10	0	42
quinones	5	0	1	1	2	16
sulfide	34	1	2	16	36	91
sulfonamide	39	3	3	27	25	112
sulfonic acid	0	0	3	1	1	17
sulfonate	4	0	3	2	2	20
sulfone	17	0	2	6	0	16
sulfonyl group	63	4	7	34	27	151
sulfonyl halide	0	0	0	0	0	1
sulfoxide	2	0	0	1	1	15
thiocarboxamide	0	0	0	0	0	4
thiocarboxylates	2	0	0	3	1	5
thioxomethyl	12	2	4	12	6	37
urea	46	4	14	28	14	107
Heterocycles	340	30	66	189	182	1026
azepine	1	0	1	1	5	17
azetidine	0	0	0	0	26	25
aziridine	0	0	0	0	1	9
benzimidazole	5	0	2	3	0	29
benzimidazole, 2-oxo	0	0	0	0	0	3

Substructure	PPP	EFSA neurotoxicity test set	EFSA developmental test set (extended)	ToxRef test set	CAESAR training	Leadscope training
1,4-benzodiazepine	0	0	0	0	9	30
1,4-benzodioxin	0	0	0	0	0	4
1,3-benzodioxole	2	1	1	3	2	11
benzofuran	0	0	0	0	0	5
benzopyran	11	0	0	3	0	33
benzopyran, 2-oxo	9	0	1	1	6	7
benzopyran, 4-oxo	1	0	0	1	1	11
benzopyrazole	0	0	0	0	0	1
5,1-benzothiazepine, 2-oxo	0	0	0	0	0	1
1,4-benzothiazine	0	0	0	0	5	28
1,3-benzothiazole	9	0	0	1	0	4
1,2-benzothiazole	0	0	0	0	0	2
1,2-benzothiazole, trioxo	0	0	0	0	1	1
benzothiophene	0	0	0	0	0	2
1,4-benzoxazine	1	0	1	1	0	3
1,3-benzoxazole	2	0	1	1	0	0
1,2-benzoxazole	0	0	0	0	0	3
beta lactam	0	0	0	0	26	24
bicyclic amines	1	0	0	0	19	32
1,3-diazine(H)	8	0	1	3	12	62
1,2-diazine(H)	7	0	0	6	0	1
1,4-diazepine	0	0	0	0	9	39
1,3-diazepine	0	0	0	0	0	2
1,2-diazepine	0	0	0	0	0	1
1,4-dioxane	1	0	0	0	1	7
1,3-dioxane	2	0	0	0	1	1
1,3-dioxolane	9	1	1	6	3	21
1,4-dithiane	2	0	0	0	0	0
1,3-dithiolane	3	0	0	0	0	0
epoxide	4	0	1	0	1	25
furan	9	0	0	1	5	21

Substructure	PPP	EFSA neurotoxicity test set	EFSA developmental test set (extended)	ToxRef test set	CAESAR training	Leadscope training
6,6-fused N-rings	2	0	0	1	3	31
5,6-fused N-rings	5	0	0	1	1	44
5,5-fused N-rings	0	0	0	0	0	8
imidazole	12	0	4	7	9	97
imidazolidine	10	3	3	10	6	35
indazole	0	0	0	0	0	1
indole	3	0	0	0	4	54
isoindole, 1,3-dioxo	4	0	1	1	0	6
isoindole, 1-oxo	4	0	1	1	1	7
isothiazolidine	1	0	0	1	0	0
isoquinoline	0	0	0	0	0	2
isothiazole	0	0	0	0	0	2
isoxazole	4	0	2	2	5	14
isoxazolidine	3	0	0	2	0	0
morpholine	6	1	5	1	1	21
1,3,4-oxadiazole	3	0	1	1	0	1
1,2,5-oxadiazole	0	0	0	0	0	1
1,2,3-oxadiazole	0	0	0	0	0	1
1,4-oxazepine	0	0	0	0	0	4
1,3-oxazepine	0	0	0	0	0	1
1,2,4-oxadiazole	1	0	0	0	0	0
1,4-oxazine	7	1	6	2	1	24
1,3-oxazine	0	0	0	0	0	4
oxazole	2	0	1	1	0	5
oxazolidine	7	0	2	5	3	18
oxepin	3	0	0	0	0	9
oxetane	1	0	0	1	0	4
oxolane	22	6	3	6	18	92
piperazine	1	0	1	0	6	91
piperidine	6	1	3	3	32	150
pteridine	1	0	0	0	5	4

Substructure	PPP	EFSA neurotoxicity test set	EFSA developmental test set (extended)	ToxRef test set	CAESAR training	Leadscope training
purine	2	0	0	0	4	17
purine, 2,6-dioxo	0	0	0	0	0	9
pyran(H)	18	5	1	4	8	120
pyrazine	7	1	0	1	6	11
pyrazine(H)	2	0	1	1	6	95
pyrazole	15	2	0	8	0	10
pyrazolidine	1	0	0	0	2	8
pyridazine	2	0	0	1	1	6
pyridine	50	4	6	39	19	116
pyridine, 1,4-dihydro	2	0	0	1	0	40
pyridine(H)	10	1	3	7	32	222
pyrimidine	43	6	8	22	17	79
pyrimidine, 2,4-dioxo	0	0	1	2	4	27
pyrrole	6	1	1	2	5	58
pyrrolidine	14	2	3	8	6	113
						22
pyrrolidine, 2-oxo	10	2	3	8	2	
quinazoline	1	0	1	0	1	8
quinazoline, 4-oxo	2	0	0	0	3	8
quinoline	5	1	1	5	3	17
quinoline, 2-oxo	1	0	0	0	0	5
quinoline, 4-oxo	1	0	0	0	0	17
quinoxaline	5	1	0	1	1	2
3,4-ring systems	0	0	0	0	22	65
rings size 4-7 O+S	2	1	1	2	0	1
z-rings size 8-14	1	1	0	1	0	16
spiro amines	1	1	0	0	0	0
spiro ethers	4	6	2	2	1	12
spiro lactones	0	1	1	1	1	3
spiro lactams	0	1	0	0	0	0
1,2,4,5-tetrazine	2	0	1	1	0	0

Substructure	PPP	EFSA neurotoxicity test set	EFSA developmental test set (extended)	ToxRef test set	CAESAR training	Leadscope training
tetrazole	1	1	0	0	1	17
1,2,4-thiadiazine, dioxo	0	0	0	0	7	5
1,3,4-thiadiazole	5	1	1	4	1	4
1,2,4-thiadiazole	1	0	0	1	0	0
1,2,3-thiadiazole	2	0	0	2	0	0
1,2,5-thiadiazole	0	0	0	0	0	4
1,4-thiazepine	0	0	0	0	14	21
1,3-thiazine	0	0	0	0	12	3
1,2-thiazine	0	0	0	0	0	6
1,4-thiazine	0	0	0	0	5	29
thiane(H)	3	0	0	1	0	5
thiazole	13	0	0	4	4	22
thiazolidine	7	2	2	4	14	30
thiepin	0	0	0	0	1	2
thiolane	0	0	0	0	0	6
thiophene	4	0	1	2	3	29
1,3,5-triazine	21	0	0	19	1	5
1,2,4-triazine	1	0	1	0	0	1
1,3,5-triazine(H)	7	0	0	3	0	2
1,2,3-triazine	2	0	0	0	0	1
1,2,4-triazine(H)	3	1	1	2	0	3
1,2,3-triazole	0	0	0	0	1	1
1,3,4-triazole	43	2	18	26	2	21
1,2,4-triazole	38	2	16	21	2	18
1,2,4-triazolidine	5	0	2	5	0	4
Naphthalenes	10	0	3	2	2	36
Natural products	1	0	0	0	21	91
Pharmacophores	770	63	131	358	289	1800
Protective groups	42	3	12	9	26	166
Spacer groups	224	25	43	110	130	712

