JRC Scientific and Technical Reports

Analysis of the Cramer classification scheme for oral systemic toxicity - implications for its implementation in Toxtree

Silvia Lapenna and Andrew Worth

EUR 24898 EN - 2011

JRC
EUROPEAN COMMISSION

ihcp
Institute for Health
and Consumer Protection

The mission of the JRC-IHCP is to protect the interests and health of the consumer in the framework of EU legislation on chemicals, food, and consumer products by providing scientific and technical support including risk-benefit assessment and analysis of traceability.

Europe Direct is a service to help you find answers
to your questions about the European Union

Freephone number (*):

00 800 6 7 8 9 10 11

(*) Certain mobile telephone operators do not allow access to 00 800 numbers or these calls may be billed.

A great deal of additional information on the European Union is available on the Internet.
It can be accessed through the Europa server http://europa.eu/

ABSTRACT

In the application of the Threshold of Toxicological Concern (TTC) concept to non-cancer endpoints, the decision tree proposed by Cramer, Ford and Hall in 1978, commonly referred to as the Cramer scheme, is probably the most widely used approach for classifying and ranking chemicals according to their expected level of oral systemic toxicity. The decision tree categorises chemicals, mainly on the basis of chemical structure and reactivity, into three classes indicating a high (Class III), medium (Class II) or low (Class I) level of concern. Each Cramer class is associated with a specified human exposure level, below which chemicals are considered to present a negligible risk to human health. In the absence of experimental hazard data, these exposure threshold (TTC) values have formed the basis of priority setting in the risk assessment process. To facilitate the application of the TTC approach, the original Cramer scheme, and an extended version, have been implemented in Toxtree, a freely available software tool for predicting toxicological effects and mechanisms of action. Building on previous work by Patlewicz and coworkers, this report provides some suggestions for improving the Cramer scheme based on a review of the scientific literature, a survey of Toxtree users, and an analysis of lists of body and food components incorporated in Toxtree.

LIST OF ABBREVIATIONS

EFSA   European Food Safety Authority
EMA   European Medicines Agency
EU   European Union
FAO   Food and Agriculture Organization
FDA   United States Food and Drug Administration
ILSI   International Life Sciences Institute
JECFA   Joint FAO/WHO Expert Committee on Food Additives
JRC   Joint Research Centre
NOEL   No observed effect level
QSAR   Quantitative Structure-Activity Relationship
SA   Structural alert
TTC   Threshold of Toxicological Concern
WHO   World Health Organization

CONTENTS

1. Introduction to the TTC concept and the Cramer decision tree

1.1 Introduction

The Threshold of Toxicological Concern (TTC) concept refers to the establishment of a generic oral exposure level for (groups of) chemicals below which there is expected to be no appreciable risk to human health (Barlow, 2005). The TTC approach is intended for use as a screening tool for chemicals for which substance-specific toxicity data are not available or routinely required in regulatory submissions, for example metabolites and degradation products. Originally, the approach was used in the assessment of indirect food additives (contact substances) and food flavourings; and subsequently, it has been investigated and proposed for use in a wide range of regulatory areas.

To facilitate the consistent and transparent application of the TTC approach, including the assessment of both cancer and non-cancer endpoints, the JRC has developed the Toxtree software (http://ihcp.jrc.ec.europa.eu/our_labs/computational_toxicology/qsar_tools/toxtree), in collaboration with various partners, including IdeaConsult Ltd (Bulgaria), Curios-IT (The Netherlands) and the Istituto di Sanita' (Italy). The most widely used TTC approach for non-cancer endpoints is the Cramer decision tree, developed by Cramer, Ford and Hall in 1978 (Cramer et al, 1978). An initial evaluation of the Toxtree implementation of the Cramer scheme was carried out by Patlewicz and coworkers (2008), and this led to several recommendations for improving the Cramer scheme and its computational interpretation in Toxtree.

This paper summarises additional findings and observations on the Cramer scheme, based on a literature review and on a survey of Toxtree users. The survey was carried out in collaboration with the European Food Safety Authority (EFSA) and the results were used to support the development of an EFSA Opinion on the applicability of TTC in the food and feed safety area (EFSA, 2011).

1.2 Cramer decision tree

In the application of the TTC concept to non-cancer endpoints, the Cramer decision tree is probably the most commonly used approach for classifying and ranking chemicals on the basis of their expected level of oral toxicity. It was proposed by Cramer, Ford and Hall in 1978 (Cramer et al, 1978) as a priority setting tool in the safety assessment of food additives which would make expert judgements more transparent, explicit and rational, and thus more reproducible and trustworthy. The scheme was derived from the authors' earlier experience in classifying food flavours (Oser & Hall, 1977) and their subsequent work in evaluating a range of carcinogens, pesticides and industrial chemicals (Cramer et al, 1978).

The original Cramer decision tree consists of 33 questions, each answered 'yes' or 'no' and leading to another question or to the final classification into one of the three classes (I, II and III) as follows:

Class I   Substances with simple chemical structures and for which efficient modes of metabolism exist, suggesting a low order of oral toxicity.

Class II   Substances which possess structures that are less innocuous than class I substances, but do not contain structural features suggestive of toxicity like those substances in class III.

Class III   Substances with chemical structures that permit no strong initial presumption of safety or may even suggest significant toxicity or have reactive functional groups.

The logic of the sequential questions (Figure 1) was based on the then available knowledge on toxicity and on how chemical structures were metabolised in mammalian metabolic pathways. The questions

relate mostly to chemical structure, but natural occurrence in the body (Q1) and in food (Q22) are also taken into consideration (Table 1). The tree is intended for use with all ingested, structurally defined organic and metallo-organic substances.

The Cramer scheme was tested against 81 chemicals including pesticides, drugs, food additives and industrial chemicals with known no observed effect level (NOEL) values reported in terms of dietary concentrations in short-terms or chronic studies (Cramer et al. 1978). Although there was overlap in the range of magnitudes of the NOELs between the three structural classes, it was clear that the NOELs of Class I substances were generally higher than those of Class III, with those of Class II being in between. Noteworthy, there was no underestimation of toxicity when compared with the available chronic oral toxicity data.

1.3 Derivation of human exposure threshold values

The Cramer decision tree was subsequently used by Munro and coworkers with the purpose of deriving human exposure levels (TTC values) for toxicity endpoints other than carcinogenicity (Munro et al., 1996). The Munro dataset comprised over 613 organic chemicals with associated 2941 NOEL values derived from a variety of non-cancer endpoints from sub-chronic, chronic, reproductive and developmental toxicity studies carried out in rodents and rabbits. The authors assigned the dataset substances to one of the three classes based on the Cramer scheme, and derived human exposure threshold values by taking the lower fifth percentile value of the distribution of NOELs for each of the three Cramer classes, multiplying by 60 to convert the values expressed as mg/kg bw per day into mg/person per day, and then dividing by a factor of 100 to ensure a margin of safety. On this basis, Munro and coworkers proposed TTC values of 1800, 540 and 90 µg/person/day for class I, II and III, respectively.

In addition to the above-mentioned TTC levels for non-cancer endpoints, specific (and lower) TTC levels have also been derived for compounds with structural alerts for genotoxicity (0.15 µg/day) and for organophosphates (18 µg/day) (Kroes et al., 2004), the general idea being that these lower threshold values should be applied in a tiered assessment approach before the Munro non-cancer threshold values.

The TTC levels proposed by Munro are now widely used in the food safety area, for example in the international evaluation of flavouring substances (first applied by the Joint FAO/WHO Expert Committee on Food Additives (JECFA) in 1997 [WHO, 1997])

1.4 Investigations on the applicability of the TTC approach

Application of the TTC approach requires knowledge of the chemical structure of the substance concerned as well as information on human exposure. It is intended for use as a screening tool for chemicals for which substance-specific toxicity data are not available. At present, the approach is used in the EU (EFSA) and internationally (JECFA) in the evaluation of flavouring substances in food, and is also used in the EU in the evaluation of pesticide metabolites in groundwater (SCP, 2000).

In addition, the approach has been investigated for its possible application in a wide range regulatory areas, including environmental risk assessment (De Wolf et al., 2005), consumer products (Blackburn et al. 2005), cosmetics (Kroes et al., 2007), impurities in pharmaceutical preparations (Müller et al. 2006) and plant-protection product metabolites (Melching-Kollmuß et al. 2010; CRD 2010).

There is also considerable ongoing research. For example, the ILSI Research Foundation (http://www.ilsi.org/ResearchFoundation/) is working with the US EPA and the biocides industry on of the application of the TTC approach to antimicrobial pesticides. Within the EU-funded COSMOS

project (http://www.cosmostox.eu/), the applicability of current and novel classification schemes to cosmetics is being explored.

Figure 1. Decision tree of the Toxtree - Cramer scheme (original) plug-in. Yes branch in green. No branch in red. Terminal nodes (labelled 1, 2 & 3) refer to Cramer classifications I, II and III.

Figure 2. Decision tree of the Toxtree - Cramer scheme with extensions plug-in. Yes branch in green. No branch in red. Terminal nodes (labelled 1, 2 & 3) refer to Cramer classifications I, II and III.

Table 1. Questions of the Cramer decision tree: the original scheme (Q1-33) and the extended scheme (Q40-44) as implemented in the Toxtree software

Question No.	Question title	If YES, assign label	If YES, go to rule	If NO, assign label	If NO, go to rule
1	Normal constituent of the body	Low (Class I)			2
2	Contains functional groups associated with enhanced toxicity	High (Class III)			3
3	Contains elements other than C,H,O,N,divalent S		4		43
4	Elements not listed in Q3 occurs only as a Na,K,Ca,Mg,N salt, sulphamate, sulphonate, sulphate, hydrochloride ...		40	High (Class III)	
40	Possibly harmful organophosphate or organophosphothionate...	High (Class III)			41
41	Removes phosphates	Low (Class I)			7
7	Heterocyclic		8		16
8	Lactone or cyclic diester		9		10
9	Lactone, fused to another ring, or 5- or 6-membered α,β-unsaturated lactone?	High (Class III)			[Open chain]
[Open chain]	Open chain		20		[Hetero-cyclic]
20	Aliphatic with some functional groups (see explanation)		21		22
21	3 or more different functional groups	High (Class III)			44
44	Free α,β-unsaturated heteroatom...	High (Class III)			18
18	One of the list (see explanation)	Intermediate (Class II)		Low (Class I)	
22	Common component of food	Intermediate (Class II)			33
33	Has sufficient number of sulphonate or sulphamate groups	Low (Class I)		High (Class III)	
[Hetero-cyclic]	Heterocyclic		10		23
10	3-membered heterocycle	High (Class III)			11
11	Has a heterocyclic ring with complex substituents.		33		12
12	Heteroaromatic		13		22
13	Does the ring bear any substituents?		14	High (Class III)	
14	More than one aromatic ring		15		22
15	Readily hydrolysed		[Heterocyclic]		33
[Hetero-cyclic]	Heterocyclic		22		16
16	Common terpene	Low (Class I)			17

4

Question No.	Question title	If YES, assign label	If YES, go to rule	If NO, assign label	If NO, go to rule
17	Readily hydrolysed to a common terpene		[Terpene]		19
[Terpene]	Common terpene		18		19
19	Open chain		20		23
23	Aromatic		27		24
27	Rings with substituents		28	High (Class III)	
28	More than one aromatic ring		29		30
29	Readily hydrolised		[Aromatic]		33
[Aromatic]	Aromatic		30		19
30	Aromatic Ring with complex substituents		31		[Aromatic]
31	Is the substance an acyclic acetal or ester of substances defined in Q30?		[Aromatic]		32
[Aromatic]	Aromatic		18		19
32	Contains only the functional groups listed in Q30 or Q31 and those listed below.	Intermediate (Class II)			22
24	Monocarbocyclic with simple substituents		18		25
25	Cyclopropane, …	Intermediate (Class II)			26
26	Monocycloalkanone or a bicyclocompound	Intermediate (Class II)			22
43	Possibly harmful divalent sulphur (not detected via Q3)...	High (Class III)			5
5	Simply branched aliphatic hydrocarbon or a common carbohydrate	Low (Class I)			6
6	Benzene derivative with certain substituents	High (Class III)			42
42	Possibly harmful analogue of benzene...	High (Class III)			7

Questions 40-44 (number underlined) are part of the extended version of the Cramer decision tree scheme. The 'Cramer rules with extensions' plug-in was developed by Curios-IT, The Netherlands, on behalf of the JRC.

2. Computer-based implementation of the Cramer decision tree

While the Cramer classification scheme undoubtedly served to improve consistency between the toxicological evaluations made by different experts, its paper-based application presupposes a working knowledge of organic chemistry and biochemistry, as the rulebase relies primarily on features of chemical structure, chemical reactivity, toxicity and metabolism, and inevitably involves a degree of subjectivity. Therefore, following a recommendation made in a JRC workshop (Saliner et al, 2005), the JRC commissioned the development of a software tool, Toxtree, to facilitate the consistent application of the Cramer scheme. Toxtree is an open-source software and is freely downloadable from the JRC website (http://ihcp.jrc.ec.europa.eu/our_labs/computational_toxicology/qsar_tools/toxtree) and from the sourceforge.net site: (http://toxtree.sourceforge.net/). The Toxtree implementation of the Cramer scheme can, in principle, be applied to organic molecules, organic salts and organometallics, and structurally well-defined oligomers and polymers. However, there are some scientifically motivated exclusions to the use of Cramer scheme due to insufficient data in the underlying TTC datasets (EFSA, 2011).

It should be noted that the computer-based implementation of the Cramer scheme in Toxtree and other software tools (e.g. OECD QSAR Toolbox) has inevitably involved some decisions by the programmer, such as the chemically-based interpretation of the original rules, and the establishment of pre-defined "look-up lists" of normal body constituents and common food components. The performance of the Cramer scheme in Toxtree has been evaluated by Patlewicz et al. (2008).

Toxtree (v. 2.1.0 and later) includes three rulebases relevant to TTC assessment: 1) the Cramer decision tree (Figure 1) as described in the original paper (Cramer et al., 1978); 2) the Cramer rulebase with extensions (Figure 2); and 3) the TTC decision tree by Kroes et al. (2004). Table 1 lists the questions and the architecture of the Cramer scheme (including extensions).

The Extended Cramer rulebase was introduced in 2009 to overcome possible misclassifications of several compounds which, following the Cramer decision tree scheme, were classified as Class I or Class II by Munro et al. (1996) even though the authors reported low NOEL values upon oral administration (indicating relatively high toxicity). The extended Cramer rulebase works by assigning compounds to Class I, II, or III, according to the original Cramer rules (questions 1-33) and five extra ones (questions 40-44). Also, the extended Cramer rulebase has an expanded list of natural body constituents for question 1 (over 400 unique compounds, with no hormones) with respect to the original Cramer scheme plug-in (67 compounds), in line with Cramer's assumption that, with the exclusion of hormones, human endogenous compounds pose a minor risk to health (Cramer et al., 1978).

The TTC decision tree by Kroes et al. (2004) results in three possible outcomes: a) substance would not be expected to be a safety concern; b) negligible risk (low probability of a life-time cancer risk greater than 1 in 10^6); and c) risk assessment requires compound-specific data. It incorporates the Benigni/Bossa rules for the identification of genotoxic carcinogens (Benigni et al., 2008), and requires the user to input the estimated daily intake.

Toxtree provides a convenient computer-based and consistent means of applying the Cramer scheme (and related rulebases), which was developed in the pre-computational era. Any discrepancy between the Toxtree classification and the expected Cramer assignment can be traced to a specific rule by examining the Toxtree pathway (this is termed the "verbose explanation" appearing in the 'reasoning' window of the software main screen), which gives the answer to each question followed by the final classification (e.g. 1N, 2N, 3Y, 4N – Class III).

3. Survey on the Cramer decision tree in Toxtree

As the Cramer scheme in Toxtree is being increasingly used, particularly by industry and regulatory bodies, it was decided to further evaluate its scope and limitations in relation to the TTC approach.

We carried out a survey in the form of a structured questionnaire to Toxtree users, with a view to: a) identifying decision rules for which clarification was needed; b) obtaining recommendations to revise, remove or add a given rule; and c) identifying bugs or inconsistencies in the Toxtree implementation of the Cramer rulebase. The questionnaire is reported in Appendix 1. The comments received and recommended follow-up actions are summarised in Appendix 2.

Although a relatively small number of completed questionnaires were returned (5 replies out of 12 recipients), most responders provided detailed comments and raised additional questions on specific issues related to interpretation of the Cramer scheme and the classification results they obtained with given test compounds.

According to the results of the questionnaire, the Cramer scheme is currently being used in the evaluation of trace drinking water contaminants, flavouring substances, migrants from packaging materials and excipients and impurities (including leachables and extractables) in drug products or personal and household care products. The main observations concerning the scientific refinement of the Cramer scheme can be summarised as follows:

1) Many of the original Cramer rules are written in a confusing and inter-dependent way, which leads to difficulties in the rationalisation of the predictions they make. These rules should be rewritten in a clearer way, possibly with modification and re-ordering.

2) Two rules are not based on chemical features, but simply make reference to look-up lists of chemicals (Q1, normal body constituents; Q22, common food components). The Cramer scheme should be revised by removing these two questions. The use of look-up lists to assess whether specified substances can be regarded as "safe" (or even "unsafe") should be carried out separately (outside of the Cramer scheme *per se*), as part of the overall the TTC assessment scheme. These look-up lists can be adapted (i.e. extended, reduced or even not used by a given regulatory body) based on expert knowledge and experience. Ideally, the lists should be subject to regulatory peer-review.

3) Some rules make ambiguous references to chemical features (e.g. steric hindrance) which need to be clarified and possibly revised/deleted.

4) The Cramer rulebase with extensions (which already contains an expanded list for Q1) does not appear to be widely used.

5) Several studies (e.g. Appendix 3) have identified outliers in the Cramer classification scheme, for example Class I outliers that have low NOELs). These outliers should be taken into account in any revisions to the Cramer scheme, with a view to making the classification scheme more discrminating in terms of NOEL values.

6) It might be desirable to establish an international forum for agreeing changes to the Cramer scheme.

4. Analysis of the look-up lists for common body and food components

The Toxtree user manual refers to look-up lists of common body and food components as follows:

> *"Cramer rules #1 and #22 depend explicitly on user-defined lists of compounds, which are normal constituents of the body or common components of food. We provide example lists of such compounds in the files bodymol.sdf and foodmol.sdf respectively... Please, note that the bodymol.sdf and foodmol.sdf files are provided mainly as an example. They contain currently only a very limited number of "Normal constituents of the body" and "Common components of food" respectively, following an expert advice. Users should consider expanding these files with appropriate molecules."*

Since some of the comments received in the survey recommended that the look-up lists of common body and food components should be excluded from the Cramer scheme, we decided to explore how the Cramer scheme would classify these substances if the look-up lists (bodymol and foodmol) were not taken into account. This was carried out by processing the bodymol and foodmol structures in Toxtree, using both the unmodified form of the Cramer rulebase, and a modified form in which the look-up files were replaced with empty (structure-free) files. The method and results are described in more detail in Appendix 4.

A total of 548 compounds, including 440 bodymol and 108 foodmol compounds, were analysed. The results show that a high proportion (75%; 409 out of 548) of the compounds in the look-up lists are potentially of medium (Class II) or high (Class III) concern as classified by the Toxtree 1.6 - Cramer with extensions plug-in (modified, without look up lists), whereas 139 compounds (25%) are of low concern. Of the 409 compounds classified in Cramer Classes II or III, 114 compounds are predicted as genotoxic by the Benigni-Bossa module for carcinogenicity and mutagenicity (Benigni et al, 2008). Thus, in the context of an overall TTC assessment scheme, such as the Kroes decision tree (Kroes et al. 2004), these 114 compounds would be removed in a pre-Cramer step, and would therefore not receive a Cramer classification. In addition, the Benigni-Bossa rulebase also predicts 10 of 139 Cramer Class I compounds as genotoxic. The pre-screening by the Benigni-Bossa rulebase would therefore result in124 compounds being removed from further consideration, while 424 compounds would enter the Cramer scheme.

5. Concluding remarks and recommendations

The Cramer scheme was proposed in the late 1970s, before the development of what is now understood by the TTC approach and before the advent of computer-based tools for interpreting chemical structure and applying structure-activity relationships. Subsequently, in the 1990s, Munro and colleagues proposed the association between Cramer classes I, II and III and human exposure thresholds for non-cancer endpoints of 1800, 540 and 90 µg/person/day, respectively (Section 1.3). On the basis of various independent analyses, using different datasets and endpoints of concern, these threshold values have been found to be robust and protective for human health (EFSA, 2011). Nevertheless, an alternative set of threshold values could be derived on the basis of a more extensive toxicity database and a classification scheme reflecting current knowledge of toxicity and metabolism, and incorporating structural features identified by modern structure-activity modelling tools.

Despite its practical applicability in protecting human health, the Cramer scheme has a number of scientific limitations, namely: a) it is dated, being based on the knowledge of the late 1970s; b) there is considerable overlap in the NOEL distributions of the three classes, which means that the Cramer classes do not discriminate well between substances of different toxic potencies; and c) Cramer Class II is ill-defined and sparsely populated for most chemical types, and thus of questionable added value.

In the short-term, it is recommended that the Cramer scheme should be rewritten, making it more transparent and easy-to-understand. Such an exercise should take into account the rules that work well, but also make changes (rule additions, rule deletions, revisions in rule scope and ordering, where necessary). In re-writing the Cramer scheme, it may be sufficient to include only two classes, corresponding to high and low concern, where the high concern class corresponds to current Cramer classes II and III. Lists of "safe" or "unsafe" substances should not be part of the scheme. Thus, the bodymol and foodmol lists of common endogenous body components and food ingredients should be inactivated.

In the longer term, the Cramer scheme should be significantly rewritten / abandoned in favour of a purely structure-based classification scheme that takes into account current research based on the use of multivariate statistical and data-mining methods to uncover new structural features that may be useful in setting human exposure thresholds. Some preliminary investigations in this direction have been published (e.g. Bassan et al, 2011). Again, lists of "safe" or "unsafe" substances should not be part of the non-cancer classification scheme, but should be considered as part of the overall TTC assessment scheme. Similarly, the inclusion of additional thresholds relating to (genotoxic) carcinogens and other endpoints of high concern (e.g. neurotoxicity) should be part of the overall TTC assessment scheme.

Finally, given the considerable international interest in the use of the Cramer decision tree, it would be desirable to establish an international platform to discuss and agree changes to the Cramer scheme, which would then form the basis for new plug-ins to the Toxtree software.

6. Acknowledgements and Disclaimer

7. References

Benigni R, Bossa C, Jeliazkova N, Netzeva T & Worth A (2008). The Benigni / Bossa rulebase for mutagenicity and carcinogenicity – a module of Toxtree. JRC Technical Report, EUR 23241 EN. Available at: http://ihcp.jrc.ec.europa.eu/our_labs/computational_toxicology/publications

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). Accessible from: http://www.efsa.europa.eu/fr/supporting/pub/159e.htm.

Blackburn K, Stickney JA, Carlson-Lynch HL, McGinnis PM, Chappell L & Felter SP (2005). Application of the threshold of toxicological concern approach to ingredients in personal and household care products. *Regulatory Toxicology and Pharmacology* **43**, 249-259.

Cramer GM, Ford RA & Hall RL (1978). Estimation of toxic hazard--a decision tree approach. *Food and Cosmetics Toxicology* **16**, 255-276.

EFSA (2011). Draft Scientific Opinion on Exploring options for providing preliminary advice about possible human health risks based on the concept of Threshold of Toxicological Concern (TTC). Available at: http://www.efsa.europa.eu/en/consultations/call/110712a.htm

ICH (2011). Impurities: guidelines for residual solvents Q3C(R5), step 4 version. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. Available at: http://www.ich.org/products/guidelines/quality/article/quality-guidelines.html

Kroes R, Renwick AG, Cheeseman M, Kleiner J, Mangelsdorf I, Piersma A, Schilter B, Schlatter J, van Schothorst F, Vos JG & Wurtzen G (2004). Structure-based thresholds of toxicological concern (TTC): guidance for application to substances present at low levels in the diet. *Food and Chemical Toxicology* **42**, 65-83.

Kroes R, Renwick AG, Feron V, Galli CL, Gibney M, Greim H, Guy RH, Lhuguenot JC & van de Sandt JJ (2007). Application of the threshold of toxicological concern (TTC) to the safety evaluation of cosmetic ingredients. *Food and Chemical Toxicology* **45**, 2533-2562.

Melching-Kollmuss S, Dekant W & Kalberlah F (2010). Application of the "threshold of toxicological concern" to derive tolerable concentrations of "non-relevant metabolites" formed from plant protection products in ground and drinking water. *Regulatory Toxicology and Pharmacology* **56**, 126-134.

Muller L, Mauthe RJ, Riley CM, Andino MM, Antonis DD, Beels C, DeGeorge J, De Knaep AG, Ellison D, Fagerland JA, Frank R, Fritschel B, Galloway S, Harpur E, Humfrey CD, Jacks AS, Jagota N, Mackinnon J, Mohan G, Ness DK, O'Donovan MR, Smith MD, Vudathala G & Yotti L (2006). A rationale for determining, testing, and controlling specific impurities in pharmaceuticals that possess potential for genotoxicity. *Regulatory Toxicology and Pharmacology* **44**, 198-211.

Munro IC, Ford RA, Kennepohl E & Sprenger JG (1996). Correlation of structural class with no-observed-effect levels: a proposal for establishing a threshold of concern. *Food and Chemical Toxicology* **34**, 829-867.

Naven RT & Derzi M (2011). Poster presentation at the Genotoxic Impurities Conference, 28-29 June 2011, Berlin, Germany.

Oser BL & Hall RL (1977). Criteria employed by the expert panel of FEMA for the GRAS evaluation of flavouring substances. *Food and Cosmetics Toxicology* **15**, 457-466.

Patlewicz G, Jeliazkova N, Safford RJ, Worth AP & Aleksiev B (2008). An evaluation of the implementation of the Cramer classification scheme in the Toxtree software. *SAR and QSAR in Environmental Research* **19**, 495-524.

Phillips JC, Purchase R, Watts P & Gangolli SD (1987). An evaluation of the decision tree approach for assessing priorities for safety testing of food additives. *Food Additives & Contaminants* **4**, 109-123.

8. Appendices

Appendix 1. Questionnaire on the use of Toxtree Cramer

The Cramer classification scheme is implemented as a rulebase in Toxtree, a standalone software application which is downloadable from the Joint Research Centre (JRC) website.

The aim of this questionnaire is to gain an overview of how the Toxtree – Cramer rulebase is being used in decision making, to develop better guidance on its application, and to guide the further development of the software.

The information provided will also be used in the drafting of an European Food Safety Authority (EFSA) Opinion on the applicability of the TTC concept in food safety assessment. Please feel free to forward the questionnaire to interested persons.

Please e-mail the completed form to: Dr Silvia Lapenna, EC Joint Research Centre
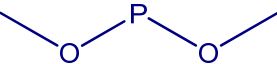
E-mail: silvia.lapenna@ec.europa.eu

Question	Answer
1) Please provide your name and organisation, or indicate that you wish your responses to be kept anonymous.	
2) For what purpose(s) do you (or does your organisation) use the Cramer scheme (e.g. what kinds of compounds and in which regulatory context)?	
3) Have you encountered examples of chemicals for which the Toxtree interpretation of a specific rule or set of rules was in disagreement with your knowledge and/or when the Toxtree output (Cramer class) was contrary to your expectation? If yes, please give the chemical name and structure, and describe the discrepancy.	
4) How do you assess "difficult" substances such as salts, metallic compounds, and polymers using Toxtree?	
5) Have you ever used the extended Cramer scheme which is available as a Toxtree plug-in?	
6) Have you encountered any software problems, such as bugs or processing errors, while working with the Toxtree? If yes, please specify.	
7) Do you have any general suggestions on how Toxtree could be improved (not limited to the Cramer rulebase)?	
8) Do you have any suggestions on how the Cramer scheme (as opposed to Toxtree) could be improved?	
9) Please add any additional comments or suggestions here.	
10) Please indicate if any information provided above cannot be shared outside the JRC.	

Appendix 2. Comments / observations on Cramer rulebase and its Toxtree implementation

Cramer question	Rule definition and explanation	Observation from literature or questionnaire	Type of comment & proposed follow-up
Q1	Is the substance a normal constituent of the body, or an optical isomer of such?	The look-up list could be extended (Patlewicz et al. 2008). Alternatively, the look-up list could be removed from the Cramer decision tree scheme, and considered separately as additional information (JRC observation).	Rulebase + software The list of normal body constituents (bodymol.sdf file) contains 68 compounds (original Cramer) and 440 compounds (Cramer with extensions).
		Some of the listed natural body constituents may pose toxicity concerns and should not be classified as 'low concern' substances, e.g. potent receptor binders or their precursors (e.g. dopamine, L-dopa, tyramine), prostaglandin-H2, and reactive species (e.g. H_2O_2). It was found that 352 out of 440 chemicals on the bodymol list would trigger a class II or III classification, if they were not treated as class I because they are constituents of the body.	Rulebase + software 1) Remove any chemical look-up lists from the Cramer classification scheme (Q1 and Q22), as inconsistent with the remainder of the scheme, which is based on structure or reactivity (hydrolysis/metabolism). In this way, toxicity classification will result only from its chemical structure and its biokinetics. 2) Keep these lists as supplementary information in the context of the overall TTC evaluation.
		The JECFA Cramer Classification of flavouring substances often differs from the output of Toxtree, mainly because of differences regarding their qualification as natural body constituents	Rulebase + software Obtain lists of flavouring substances and check against the bodymol file.
Q2	Does the substance contain any of the following functional groups associated with enhanced toxicity: an aliphatic secondary amine or a salt thereof, cyano, N-nitroso, diazo, triazeno or quaternary nitrogen, except in any of the following forms: >C=N+R2, >C=N+H2 or the hydrochloride or sulphate salt of a	-	-

Cramer question	Rule definition and explanation	Observation from literature or questionnaire	Type of comment & proposed follow-up
	primary or tertiary amine?		
Q3	Does the structure contain elements other than C, H, O, N or divalent S?	Examples from EFSA: In the following molecules, the S in aromatic ring is not recognised as divalent: O=C(C1=C(SN=N2)C2=CC=C1)SC O=C(O)N[C@H](C(N[C@@H](C1=NC2=CC=C(F)C=C2S1)C)=O)C(C)C C1(C2=CSC=N2)=NC3=CC=CC=C3N1	Answer to Q3 was found to be wrong: it is a divalent S even if it is part of an aromatic ring. On the other hand, answer to Q3 is correct when you run an unsaturated ring containing S (divalent). This problem has been fixed in Toxtree v. 2.1.0 and later, by replacing valency calculation based on CDK atom typing, instead of counting bond order (1.5 for aromatic bonds).
Q43	Does the compound contain a non-natural divalent sulphur (not detected via Q3)?	-	-
Q4	Do all elements not listed in Q3 occur only as (a) a Na,K,Ca,Mg or N salt of a carboxylic acid, or (b) a sulphate or hydrochloride of an amine, or (c) a Na,K, or Ca sulphonate, sulphamate or sulphate?	Q4 does not list phosphates Examples by Patlewicz (predicted classs II instead of I): disodium 5' guanylate disodium 5-inosate sodium steroyl lactate [Na+].O=C(OC(C)C(=O)OC(C)C(O)=O)CCCCCCCCCCCCCCCCC	Rulebase + Clarification Because Q4 does not list phosphate salts (in the original Cramer scheme and in Toxtree-Cramer scheme), any compound with a P element is assigned as Class III. However, this causes many safe natural compounds / known food additives to be misclassified as higher concern. To solve this, Q4 has been changed (and Q40 added) in the extended Cramer scheme plug-in (see below).
Q4 (in Cramer rules with extensions plug-in)		Q4 should include phosphates, but not other organophosphoro derivatives.	Rulebase + Clarification (in Cramer rules with extensions plug-in only) In the base Cramer scheme, Q4 classifies any compound with a P as Class III. However, it is known that many compounds with phosphates occur in natural compounds. To enable fewer false positive class III predictions, the Cramer scheme with extensions plug-in was developed. However, it was shown not to be implemented correctly. The correct implementation is as follows: 1) Q4 of the extended Cramer tree should be revised to allow Phosphorus elements too to proceed to Q40, i.e. answer "Yes" to compounds that contain no other possibly harmful group than a P. 2) Q40 should filter out compounds with dangerous P group, i.e. **uncharged** organophosphate, phosphonates, phosphamates, phospohoroamido, and hydroxymethylphosphinyl derivatives. In other words, Q40 should consider any P that is not part of a negatively charged

Cramer question	Rule definition and explanation	Observation from literature or questionnaire	Type of comment & proposed follow-up
			phosphate as "Yes" =>High/Class III The phrasing of Q4 of the extended Cramer tree should be updated accordingly, i.e. to allow Phosphorus elements too (of which risky ones are filtered successively by Q40).
		Examples of misprediction reported by EFSA: P(CC1=CC(C(C)(C)C)=C(O)C(C(C)(C)C)=C1)(OCC)=O.[O-]P(CC2=CC(C(C)(C)C)=C(O)C(C(C)(C)C)=C2)(OCC)=O.[Ca+2] Answer to Q4 is No, resulting in class III while should be Yes, as it contains Ca, and proceed to the next question	Software
Q40	Is any element not listed in Q3 an uncharged organophosphate Recognises possibly harmful (uncharged) organophosphate or organophosphothionate and put it to class III (but let charged PO4 through Q41)	Unclear rule	Clarification e.g. add phosphoric triamides O=P(N(C)C)(N(C)C)N(C)C, as these are Q40Y. Rulebase: investigate mispredictions for possible rule refinement
Q41	All phosphate groups that can occur in natural compounds are hydrolysed and removed. Each resulting fragment considered individually in the tree	Unclear rule? What does "hydrolysed and removed" mean? (EFSA observation)	Clarification Unlike most Cramer questions, Q41 is not a yes/no question. Instead, Q41 only splits any natural phosphate-like moiety from the molecule and: 1) removes the phosphates (PO4) from the compound, and 2) puts the remaining fragments through the rest of the tree via the "answer" "No" (via Q7), one by one (Residue 1, 2, etc.).
		In the example "No" (see the structure below), which kind of Phosphorus are you considering? (EFSA observation) 	Clarification This example is indeed confusing as this compound would actually not reach Q41, as Q40 would classify such a compound as class III. A better example is needed to illustrate a compounds that lacks a synthetic/natural phosphate.
Q5	Is it a simply branched acyclic aliphatic	Misclassifcation of sucrose monopalmiate and monsterate as class	Rulebase

15

Cramer question	Rule definition and explanation	Observation from literature or questionnaire	Type of comment & proposed follow-up
	hydrocarbon or a common carbohydrate?	II instead of I (Patlewicz et al. 2008).	Reinterpretation of "simply branched"? Ambiguity of common carbohydrate?
Q6	Is the substance a benzene derivative bearing substituents consisting only of: (a) hydrocarbon chains or 1'hydroxy or hydroxy ester-substituted hydrocarbon chains and (b) one or more alkoxy groups, one of which must be para to the hydrocarbon chain in (a)? This places in class III safrole, myristicin and related substances.	-	-
Q42	Does the compound consist of one aromatic ring, with at most one heavy atom connected to each aromatic atom? Assigns compounds that consist of a single aromatic ring with zero to six single atom-substituents as Class III.	Concerns over inclusion of monomethylated benzenes	Rulebase On the (non)toxicity of, say, toluene: while Munro handled a NOEL value of 500 mg/kg/day for toluene, lower values have also been proposed (e.g. <150 mg/kg/day by Wilkins-Haug L.Teratology. 1997;55(2):145-51). The rule can be rewritten such that only small phenols are captured.
Q7	Is the substance heterocyclic?	-	-
Q8	Is it a lactone or cyclic diester? Separates the lactones and cyclic diesters from other heterocyclic compounds	There should be no restriction to 5- and 6-membered rings for lactones in Q8 in Toxtree	Software This was recognised as a bug and fixed in Toxtree v. 2.1.0 and later.
Q9		Discrepancy (JRC observation) The 4-membered saturated lactone (C1OC(=O)C1C) used as the Yes Class III actually gives No and Class I This inconsistency is also in the original Cramer paper	Software and possible rulebase revision Need to better define this rule – what about 4-membered lactones?
Q10	Is it a 3-membered heterocycle? This places substances like epoxides and ethylenamine in class III	-	-

Cramer question	Rule definition and explanation	Observation from literature or questionnaire	Type of comment & proposed follow-up
Q11	Does it have a heterocyclic ring with complex substituents? Disregarding only the heteroatoms on any one ring, does that heterocyclic ring contain or bear substituents other than *simply branched* (class I) hydrocarbons (including bridged chains and monocyclic aryl or alkyl structures),alkyl alcohols, aldehydes,acetals,ketones,ketals,acids,esters (including cyclic esters other than lactones),mercaptans, sulphides, methyl ethers, hydroxy or single rings (hetero or aryl) with no substituents other than those just listed? Questions 11-15 separate out various categories of heteroaromatic substances. Under 11, set aside and do not consider the atom(s), usually O,N and S, making the ring heterocyclic.If there is more than one hetero ring, regard each ring separately, with the remainder of the structure as substituents of that hetero ring.Other than the heterocyclic atoms, does the ring carry anything besides the simple groups listed? If so, the answer is YES,and the next question 33. If not, then classify further by Q12 etc. Bridged chain derivatives may be represented by structures like the bicyclic ether 1,4 cineole while monocyclic aryl derivatives may be represented by compounds like benzaldehyde propylene glycol acetal or 3-phenyl-2-furancarboxaldehyde.	Incorrect classification of caffeine as class III instead of class II (Patlewicz et al. 2008).	Rulebase Reconsider scope of question 11
Q12	This question separates the aromatic heterocyclics for the purpose of considering whether they are polynuclear (Q14) or unsubstituted (Q13).	-	-
Q13	Does the ring bear any substituents?	-	-

Cramer question	Rule definition and explanation	Observation from literature or questionnaire	Type of comment & proposed follow-up
Q14	Does the structure contain more than one *aromatic* ring?	-	-
Q15	Is it *readily hydrolysed* to mononuclear residues?If YES, treat the mononuclear heterocylic residues by Q.22 and any carbocyclic residue by Q16.	-	-
Q16	Is it a *common terpene* - hydrocarbon, alcohol, aldehyde or -carboxylic acid (not a ketone)? Q16 and Q17 deal with terpenes. A hydrocarbon terpene that is a *common terpene* and has not already been put in class I by Q5, would go into class I by Q16.	-	-
Q17	Is the substance *readily hydrolysed* to a *common terpene* -alcohol, aldehyde or carboxylic acid? If the answer is YES, treat the hydrolysed residues separately and proceed to Q18 for the terpene moiety and to Q19 for any non-terpenoid moiety). Since there may be substances that are hydrolysed to two or more residues, one of which is terpene, treat the residues separately from Q18 onward.	Recognition of terpenes could be refined (Patlewicz et al. 2008). e.g. isobornylacetate (O=C(OC1CC2CCC1(C)C2(C)(C))C) predicted as class I instead of II, since it is incorrectly assumed to be readily hydrolysed	Software
Q18	Is the substance one of the following: (a) a vicinal diketone; or a ketone or ketal of a ketone attached to a terminal vynil group (b) a secondary alcohol or ester of a secondary alcohol attached to a terminal vinyl group (c) allyl alcohol or its acetal, ketal or ester derivative (d) allyl mercaptan, an allyl sulphide, an allyl thioester or allyl amine (e) acrolein, a methacrolein or ther acetals	Why are acrylate esters in Class I but acrylic acid is in class II (acrylic acid is specifically assigned to CC II in the Cramer et al publication).	Clarification and possible rulebase revision If you use the Cramer rules with extensions, the estimates for acrylate esters and acrylic acid are Class III. This is because Q44 recognises a small moiety that correlates with low NOELs and false class I/II predictions (an alpha,beta-unsaturated heteroatom moiety) and assigns such compounds as Class III.

Cramer question	Rule definition and explanation	Observation from literature or questionnaire	Type of comment & proposed follow-up
	(f) acrylic or methacrylic acid (g) an acetylenic compound (h) an acyclic *aliphatic* ketone,ketal or ketoalcohol with no other functional groups and with four or more carbons on either side of the keto group (i) a substance in which the *functional groups* are all *sterically hindered* Q18 examines the terpenes and later the open-chain and mononuclear substances by reference) to determine whether they contain certain structural features generally thought to be associated with some enhanced toxicity.		
Q44	Does the compound contain a free alpha,beta-unsaturated functional group? The functional group consists of a carbon with an attached heteroatom (O in case of an alcohol or ester). Here, 'free' means that position beta from this group contains an sp2 [CH2] or an sp1 [CH].	Unclear rule Questionable predictions with acrylic acid, methacrylic acid and their esters	Clarification and possible rulebase revision and software implementation Need to further explain how Q44 works and check its implementation. Q44 looks for a substructure pattern that can be simplified to [CH2]=C~C~O where ~ means any bond. This pattern [CH2]=C~C~O matches [CH2]CC=O and/or [CH2]CCO. Check with test set including phthalate esters, e.g. diallyl phthalate (NOEL = 150mg/kg/day).
Q19	Is the substance *open chain*? Q19-21 deal with open-chain substances.	-	-
Q20	Is the structure a linear or *simply branched aliphatic* compound, containing any one or combination of only the following *functional groups* : (a) four or less, each, of alcohol, aldehyde, carboxylic acid or esters and/or (b) one each of one or more of the following: acetal, either ketone or ketal but not both, mercaptan, sulphide (mono- or poly-), thioester, polyoxyethylene [(-	Difficult to interpret Incorrect classification of propargyl alcohol as class III instead of class II	Rulebase Need to reconsider the scope of this rule, or even break it down into multiple rules

Cramer question	Rule definition and explanation	Observation from literature or questionnaire	Type of comment & proposed follow-up
	OCH2CH2-)x with x = 4], or primary or tertiary amine This question should be answered YES if the structure contains one or any possible combination of alcoholic, aldehydic or carboxylic acid or ester groups, provided there are no more than four of any one kind. It should be answered YES if the structure contains in addition to, or instead of, those just listed, any assortment of no more than one each of the following:acetal, either ketone or ketal but not both, mercaptan, mono- or polysulphide, thioester, polyoxyethylene, primary or tertiary amine. Answer the question NO if the structure contains more than four of any of the first set of groups, more than one of the second set, or any substituent not listed.		
Q21	Does the structure contain >= 3 different types of functional groups Puts into Class III aliphatic compounds with >= 3 different types of functional groups (excluding polyesters and similar substances)	Incorrect prediction from Q21 (*reported by the Toxtree developers*): the acid and its sodium salt are incorrectly predicted to belong to different Cramer classes. According to Q4 a sodium salt should be treated as a free acid.	Software This was recognised as a bug and fixed in Toxtree v. 2.1.0 and later.
Q22	Is the substance a common component of food (C) or *structurally related* to a common component of food? Returns true if the query is isomorphic to one of the structures in the foodmol file	See comments in Q1	Rulebase and software The foodmol file currently lists 108 natural food constituents (Toxtree v.1.6 and later).
Q23	Is the substance aromatic? Q 23-26 deal with alicyclic substances	-	-
Q24	Is the substance monocarbocycic with simple substituents? Is the substance monocarbocyclic (excluding cyclopropane or cyclobutane	Class I misclassified as class II (Patlewicz et al. 2008), e.g. calcium cyclamate [Ca++].[O-]S(=O)(=O)NC1CCCCC1.	Software Toxtree counted two rings instead of one. Alert software developer – IdeaConsult Ltd

Cramer question	Rule definition and explanation	Observation from literature or questionnaire	Type of comment & proposed follow-up
	and their derivatives) with ring or *aliphatic* side chains,unsubstituted or containing only alcohol, aldehyde, side-chain ketone, acid, ester, or Na, K or Ca sulphonate or sulphamate, or acyclic acetal or ketal?	[O-]S(=O)(=O)NC2CCCCC2	
Q25	Is the substance (a) a cyclopropane or cyclobutane with only the substituents mentioned in Q24 or (b) a mono- or bicyclic sulphide or mercaptan?	-	-
Q26		-	-
Q27	Do(es) the ring(s) have any substituents? Q27-31 deal with aromatic compounds.	-	-
Q28	Does the structure contain more than one *aromatic* ring?	-	-
Q29	Is it *readily hydrolysed* to mononuclear residues? If YES, treat the individual aromatic mononuclear residues by Q30 and any other residue by Q19.	-	-
Q30	Does it contain an aromatic ring with complex substituents? Disregarding ring hydroxy or methoxy does the ring bear substituents *other* than 1-5-carbon *aliphatic* groups, either hydrocarbon or containing alcohol, ketone, aldehyde, carboxyl or simple esters that may be hydrolised to ring substituents of =5 carbons? (If a simple ester that may be hydrolysed, treat the aromatic portion by Q.18 and the residue by Q19.) This should be answered NO if the ring bears only aliphatic groups of =5 carbons, which are either hydrocarbons containing the groups listed.If the ring bears any other	Unclear rule Cut-off of 5 carbons should be reconsidered, and maybe set at 6. Example (Patlewicz et al. 2008): 4-hexylresorcinol CCCCCCC1=C(O)C=C(O)C=C1 misclassified as class II instead of I Example provided by EFSA: O=C(OCC)C1=CC=CC=C1OCC (ethyl 4-ethoxybenzoate) gives class III.: ester hydrolysis results in the aromatic part and an ethanol. Ethanol is numbered as "Residue 2": this Residue 2 get a classification of class II. Nevertheless, the aromatic part is predicted to be	Clarification and possible rulebase revision and software implementation Redefine rule, test it with selected substances, and implement changes in future software release. Current implementation: With Q30, a simple ester, i.e. Ar-COO-R, where R is a 1-5-carbon aliphatic, triggers Yes, and is hydrolyzed, and the two portions go through Q31 and Q32. This is not correct, since the Cramer rule wants the Ar part of such ester to go to Q18, and the non-Ar part to Q19. Proposed correction: do not allow hydrolysis already in Q30 (e.g. in the example provided by EFSA: Residue 2 / EtOH answers Yes to Q30, should go unhydrolised to Q31 and there be hydrolised (so that the Ar part goes to Q18, and the non-Ar part to Q19)

Cramer question	Rule definition and explanation	Observation from literature or questionnaire	Type of comment & proposed follow-up
	substituents than those listed, the question should be answered YES and proceed to Q31	class III, and so the entire compound gets assigned this class	
Q31	Is the substance an acyclic acetal or ester of substances defined in Q30? If YES, assume hydrolysis and treat the non-aromatic residues by Q19 and the aromatic residue by Q18.) This question is designed to see whether the substance would fit within the definition of Q30 if it were not an acetal, a ketal or an ester.In other words, would the substance carry only the groups listed in Q30.	-	-
Q32	Does the substance contain only the *functional groups* listed in Q30, or their derivatives listed in Q31, but with any or all of the following: (a) a single fused non-aromatic carbocyclic ring (b) aliphatic substituent chains longer than 5 carbon atoms, or (c) a polyoxyethylene [(-OCH2CH2-)x, with x = 4] chain either on the aromatic ring or on an aliphatic side chain? Part (a) is intended to allow simple derivatives of tetralin into class II while putting polycyclic compounds such as steroids ultimately into class III except those that may be normal food components. Part (b) allows compounds with permitted functional groups but longer side chains into class II instead of sending them eventually into class III. Part (c) puts short-chain polyoxyethylene derivatives of aryl compounds into class II rather than class III.	In the question explanation: (c)a polyoxyethylene [(-OCH2CH2-)x, with x = 4] chain either on the aromatic ring or on an aliphatic side chain?	Clarification While this restriction is implemented since the first Toxtree versions, the rule explanation should be clarified by saying $1<x<4$ (and not x = 4).

Cramer question	Rule definition and explanation	Observation from literature or questionnaire	Type of comment & proposed follow-up
Q33	Does the substance bear on every major structural component at least one Na, K or Ca sulphonate or sulphamate for every 20 or fewer carbon atoms, without any free primary amines except those adjacent to the sulphonate or sulphamate. Na,K,Ca sulphonate and sulphamate salts have a strong tendency to decrease toxicity by promoting solubility and rapid excretion. This is particularly noticeable, for example, with some of the food colourings. It is important that the substance bears sufficient sulphonate groups, including one on each major structural fragments into which the original compound might be metabolized. This question serves to steer sulphonated compounds except those with amines non-adjacent to the sulphonate into a presumptively less toxic classification than the compounds would occupy if unsulphonated.	Unclear rule Some class III substances (e.g. Na 5-aminonaphthalene-2-sulphonate) are misclassified as class I (Patlewicz et al. 2008). Toxtree also fails to assign as class I compounds where the sulphonate group is adjacent to a primary amine	Software
General comment		Would be good to have Toxtree set up to run the entire TTC tiered approach with one click – the software would determine if the chemical should be excluded, identify structural alerts (SAs) for genotoxicity and then determine the Cramer Class. It might be helpful to be able to determine the Cramer Class even if a chemical does have SA's but there is a potential for mis-use if the user doesn't realize that they have to consider SA's before advancing to the Cramer Classes	These suggestions have been implemented into the Toxtree plug-in for the Kroes TTC decision tree (Toxtree 2.1.0 and later)
General comment		The monomethyl and monoethyl ether derivatives of ethylene glycol have	Rulebase and software

Cramer question	Rule definition and explanation	Observation from literature or questionnaire	Type of comment & proposed follow-up
		been shown to cause haematological effects in experimental animals and also to be teratogenic following exposure by oral and other routes (see review in ECETOC 1985). Thus, Phillips et al. (1978) suggested to put all glycol ethers including ethylene glycol monoethyl and monomethyl ethers into class II or III	Phillips' suggestions could be implemented, but where (revise Q21?).
		Review Cramer treatment of Class I outliers, e.g. acetone, phenols and methoxymethanol, which have low NOELs	
		It might be desirable to establish an international forum for agreeing changes to the Cramer scheme.	Rulebase and software

Footnote: Questions 40-43 refer to the Extended Cramer scheme

Appendix 3. Comparison of Cramer and ICH classifications for residual solvents in pharmaceutical products

In a recent study by Naven & Derzi (2011)[1], the safety thresholds derived by Munro using the Cramer scheme were compared with the ICH classification for residual solvents in pharmaceutical products (ICH, 2011). Using a test set of 59 ICH solvents and 9 non-classified compounds (internally evaluated), the ability of the Cramer decision tree to discriminate between low risk and high risk compounds was assessed (Toxtree 2.1.0, Cramer rules with extension plug-in). The authors concluded that the Cramer scheme, while requiring updating and refining to reflect today's chemical space, overall was effective for distinguishing these compounds classes. However, when the structural feature contributions to low risk- and high risk-class discrimination were determined, through a Dragon Descriptor Function Group Count analysis, only a limited number of structure-toxicity relationships could be found. This was attributed to the complexity and multi-factorial progression of chronic toxicity and the statistical limitation of NOEL values as toxicological endpoint (e.g. the NOEL value is affected by the choice of test doses). Furthermore, the Cramer decision tree pathway for each compound was analysed to identify the Cramer rules which were reflective of toxicity, from which it was concluded that *broad* rules were more useful to make the high- and low-toxicity distinction. The following rules were considered to be most useful:

High toxicity:
- Questions identifying the presence of uncommon elements
- Questions identifying the presence of reactive functional groups
- Answer NO to Q33: does the substance bear a sufficient number of SO3 groups? (Na,K,Ca sulphonate and sulphamate salts have a strong tendency to decrease toxicity by promoting solubility and rapid excretion).

Low toxicity:
- Q1 natural body constituent
- Q22 food constituent
- Answer NO to Q18 (which examines the terpenes and later the open-chain and mononuclear substances to determine whether they contain certain structural features generally thought to be associated with some enhanced toxicity (e.g. heterocycles, polyaromatic, etc)

Five compounds were reported to be Cramer scheme Class 1 outliers (i.e. having low NOELs): including acetone, an alkoxy ethanol, a poly-ether alkane, 2,6-dimethyl phenol and an aryl,alkyl secondary alcohol, pointing to possible improvements of the scheme.

---

[1] The authors are grateful to Russell Naven (Pfizer) for his comments and for providing a copy of the poster presentation (Naven & Derzi, 2011).

Appendix 4. Analysis of the bodymol (body components) and foodmol (food constituents) compounds in Toxtree

Summary

Cramer rules #1 and #22 depend on user-defined lists of compounds, which are normal constituents of the body or common components of food, respectively. The user-defined lists are encoded in the bodymol.sdf and foodmol.sdf files, which can be modified (extended, reduced, removed) based on expert judgement.

In this study, we explored how the Cramer scheme would classify these substances if the look-up lists (bodymol and foodmol) were not taken into account. This was carried out by processing the bodymol and foodmol structures in Toxtree 1.6, using both the unmodified form of the Cramer rulebase, and a modified form in which the look-up files were replaced with empty (structure-free) files. In addition, we investigated how the Benigni-Bossa scheme would classify the same substances in terms of their predicted genotoxic potential.

A total of 548 compounds, including 440 bodymol and 108 foodmol compounds, were analysed. The results show that a high proportion (75%; 409 out of 548) of the compounds in the look-up lists are potentially of medium (Class II) or high (Class III) concern as classified by the Cramer scheme, whereas 139 compounds (25%) are of low concern. Of the 409 compounds classified in Cramer Classes II or III, 114 compounds are predicted as genotoxic by the Benigni-Bossa module for carcinogenicity and mutagenicity (Benigni et al, 2008). Thus, in the context of an overall TTC assessment scheme, such as the Kroes decision tree (Kroes et al. 2004), these 114 compounds would be removed in a pre-Cramer step, and would therefore not receive a Cramer classification. In addition, the Benigni-Bossa rulebase also predicts 10 of 139 Cramer Class I compounds as genotoxic. The pre-screening by the Benigni-Bossa rulebase would therefore result in 124 compounds being removed from further consideration, while 424 compounds would enter the Cramer scheme. In this further analysis, 12/13 compounds resulted in class II/III respectively, which are all foodmol compounds.

Method

To create the modified version of the Cramer rulebase, we used Toxtree 1.6, and not later versions, because in v. 2.1.0 and later the bodymol and foodmol sdf files are not stored in the application folder as they are in 1.6, but they are packaged inside one of the Toxtree-Cramer jar files which are extracted into a temporary folder when the Cramer rulebase is applied. This means that the bodymol and foodmol files cannot be easily manipulated by the user in version 2.2.0. The bodymol and foodmol files can also be obtained from Sourceforge:

> Bodymol file: http://toxtree.svn.sourceforge.net/viewvc/toxtree/bodymol.sdf?view=log
> Foodmol file: http://toxtree.svn.sourceforge.net/viewvc/toxtree/foodmol.sdf?view=log

A combined file containing 548 bodymol and foodmol substances was analysed in:
  a) Toxtree 1.6 (unmodified)
  b) Toxtree 1.6 (modified, without look up lists), in which the bodymol and foodmol files were emptied.

The 548 bodymol and foodmol substances were also assessed for their potential genotoxicity by using the Benigni-Bossa module for carcinogenicity and mutagenicity in Toxtree, including the embedded QSAR models. In this case, the latest downloadable version of the software, Toxtree 2.2.0 (October 2010), was used, since it includes several bug fixes compared with Toxtree 1.6.

In applying the Benigni-Bossa rulebase, the results were interpreted as follows: if any of the BB QSAR models were applicable, and if at least one QSAR model gave a positive outcome (i.e. "Potential S. typhimurium TA100 mutagen based on QSAR", or "Potential carcinogen based on QSAR"), a positive prediction (1) was assigned, whereas if all QSAR results were negative ("Unlikely to be a S. typhimurium TA100 mutagen based on QSAR" or "Unlikely to be a carcinogen based on QSAR") a negative outcome (0) was assigned. If none of the QSAR models were applicable, the outcome of the SA analysis was considered.

Results – Cramer analysis

The outcome of applying the original Cramer rulebase and the Cramer rulebase with Extensions in Toxtree 1.6 is summarised in Table 2.

Table 2. Cramer classifications for 548 common body or food components

	Toxtree 1.6 / Original Cramer rules (unmodified)			Toxtree 1.6 / Cramer rules with Extension (unmodified)			Toxtree 1.6 / Cramer with Extension plug-in (modified, without look up lists)		
	class I	class II	class III	class I	class II	class III	class I	class II	class III
Bodymol (440 compounds)	439	0	1*	439	0	1*	88	13	339
Foodmol (108 compounds)	76	20	12	76	16	16	51	5	52
Total (548 compounds)	515	20	13	515	16	17	139	18	391

*This is an error, due to an incorrect structure for protoheme in the bodymol file

Unmodified Cramer scheme

- When the unmodified Cramer scheme is applied to bodymol compounds, all of them are expected to be classified in Class I (since they trigger a 1Y response). In practice, however, protoheme triggered the following classification: 1N,2N,3Y,4N (class 3). This is an error, due to an incorrect structure for protoheme in the bodymol file (bodymol.sdf; #221).
- When the unmodified Cramer scheme is applied to the foodmol compounds, any hazard class (I, II or III) can result. This is because, in contrast to the assessment of the bodymol file, which is the first question and always verified, the Cramer scheme does not set verification of the foodmol file (Q22) as mandatory. Even when the compound is present in the foodmol.sdf, the flow of the questions may not pass through Q22. Several different decision-tree evaluation patterns can therefore result, and the final classification vary according to the chemical structure:
  - The analysis stops when answering to Q4 since the compound contains elements outside of the common organic subset, as defined in Q4. In this case, Class III is assigned, and Q22 is not verified. For example, "1N,2N,3Y,4N" (saccharin).
  - The analysis takes a pathway which does not include Q22 and therefore foodmol.sdf is not verified. For example, "1N,2N,3N,5N,6N,7Y,8N,10N,11Y,33N" (folic acid).
  - The analysis involves hydrolysis or other metabolic reactions. For example: "1N,2N,3N,5N,6N,7N,16N,17N,19N,23Y,27Y,28N,30Y(31N,32Y)(31N,32N,22N,33" (aspartame) and "1N,2N,3Y,4Y,40N,41N(7Y,8N,10N,11N,12Y,13Y,14N,22N,33N)" (vitamin B6). In this case, starting from the rule where the reaction occurs (i.e. 30Y for aspartame), the subsequent analysis is performed not on the original query compound,

but on each of the reaction products, which is verified against foodmol.sdf, and might not be in the list (22N).

- When the unmodified Cramer scheme was applied to the 108 foodmol compounds, 20 compounds resulted of intermediate concern (Class II), of which 16 triggered a positive answer to Q22, while the remaining four had a more complex classification pathway. None of the original query foodmol compounds had a 22N result, indicating that no structural errors were present in foodmol.sdf (22N occurs for one of the hydrolysis products of aspartame). Of the 20 Class II foodmol compounds with the original Cramer scheme, four resulted in Class III when the Cramer rules with extensions plug-in (unmodified) was used instead of the original scheme, owing to a positive answer to either question 42 (dimethylpyrazine, maltol, methylpyrazine) or question 43 (cystine) of the extended rulebase (Table 3).
- Interestingly, 44 foodmol compounds would be classified as low concern (Class I) following matching with the bodymol list (1Y).

Modified (without look-up lists) Cramer rulebase with extensions

- When the modified version (without look-up lists) of the Extended Cramer scheme was applied, Class II/III classifications were obtained for 352 bodymol substances (cf. 1 compound in the unmodified version of the scheme) and for 57 foodmol compounds (cf. 32 compounds in the unmodified version of the scheme).

Table 3. Foodmol compounds classified in Class II by the original Cramer scheme that reclassifed in Class III by the Extended Cramer rulebase in Toxtree 1.6

Compound	Original Cramer scheme result	Extended rulebase result
cystine	1N,2N,3N,5N,6N,7N,16N,17N,19Y,20N,22Y	1N,2N,3N,43Y
dimethylpyrazine	1N,2N,3N,5N,6N,7Y,8N,10N,11N,12Y,13Y,14N,22Y	1N,2N,3N,43N,5N,6N,42Y
maltol	1N,2N,3N,5N,6N,7Y,8N,10N,11N,12N,22Y	1N,2N,3N,43N,5N,6N,42Y
methylpyrazine	1N,2N,3N,5N,6N,7Y,8N,10N,11N,12Y,13Y,14N,22Y	1N,2N,3N,43N,5N,6N,42Y

Results – Benigni-Bossa analysis

For the combined dataset of 548 bodymol/foodmol compounds:

- the Benigni-Bossa QSAR models were applicable to 17 compounds (QSAR6/8 for 15 aromatic amines, and QSAR13 for 2 α,β-unsaturated aldehydes), of which 11 gave a positive mutagenic/carcinogenic prediction and 6 a negative prediction;
- 127 compounds triggered a SA for genotoxic carcinogenicity;
- no SA for non-genotoxic carcinogenicity was fired.

The breakdown of the results obtained with the Benigni-Bossa module in Toxtree according to Cramer classification is given in Table 4.

Table 4. Cramer and Benigni-Bossa classifications of the bodymol and foodmol compounds (Toxtree 1.6 / Cramer with Extension plug-in, without look-up lists *vs.* Toxtree 2.2.0 / Benigni-Bossa rulebase)

Bodymol and foodmol compounds	Positive prediction in at least one Benigni-Bossa QSAR model	Negative prediction in at least one Benigni-Bossa QSAR model	Benigni-Bossa Structural Alert(s) for genotoxic carcinogenicity	Overall prediction: potentially genotoxic
class I (139 compounds)	0	1	11	10
class II (18 compounds)	0	0	3	3
class III (391 compounds)	11	5	113	111

In Table 4, it can be seen that:

- 114 Class II/III compounds are potential carcinogens based on the Benigni-Bossa rulebase as implemented in Toxtree 2.2.0;
- 10 Class I compounds triggered a Benigni-Bossa Structural Alert for genotoxic carcinogenicity. These compounds are 2 foodmol compounds, menadione (vitamin K3) and vanillin, and 8 bodymol compounds: 3,4-dihydroxyphenyl)acetaldehyde, succinate, 4-aminobutanal, 4-methylpentanal, acetaldehyde, glyoxylate, glyceraldehyde and malonate.

It has to be noted that SAs represent a first approximation, or 'coarse-grain' approach to SAR analysis, while QSAR models are fine-tuned estimations. The current QSAR models within the Benigni-Bossa rulebase in Toxtree only apply to a limited number of compound classes (aromatic amines and α,β-unsaturated aldehydes, with restrictions). In fact, only one (cis-9-retinal) out of 139 Class-I compounds were suitable for the Benigni-Bossa QSAR analysis. Cis-9-retinal was predicted as "unlikely to be a Salmonella mutagen carcinogen based on QSAR" (QSAR13), while the α,β unsaturated carbonyls SA for genotoxic carcinogenicity was triggered.

Results – Combined application of Benigni-Bossa and Cramer rules

In an overall TTC assessment scheme, such as the one proposed by Kroes et al (2004), rules for the prediction of genotoxic potential would be applied before the Cramer rules, with the consequence that potentially genotoxic compounds re associated with a lower TTC value of 0.15 μg/day. Table 5 shows the outcomes of applying the Benigni-Bossa and Cramer classification schemes in sequence.

Table 5. Results obtained by the stepwise application of the Benigni-Bossa and Cramer (with extensions) rulebases to 548 bodymol and foodmol compounds

No of compounds entering step	Step 1 Predicted genotoxic	Step 2 Cramer class I	Step 2 Cramer class II	Step 2 Cramer class III
548	124	N/A	N/A	N/A
424 (original Cramer with look-up lists)	N/A	399	12	13
424 (modified Cramer without look-up lists)	N/A	129	15	280

Abstract

In the application of the Threshold of Toxicological Concern (TTC) concept to non-cancer endpoints, the decision tree proposed by Cramer, Ford and Hall in 1978, commonly referred to as the Cramer scheme, is probably the most widely used approach for classifying and ranking chemicals according to their expected level of oral systemic toxicity. The decision tree categorises chemicals, mainly on the basis of chemical structure and reactivity, into three classes indicating a high (Class III), medium (Class II) or low (Class I) level of concern. Each Cramer class is associated with a specified human exposure level, below which chemicals are considered to present a negligible risk to human health. In the absence of experimental hazard data, these exposure threshold (TTC) values have formed the basis of priority setting in the risk assessment process. To facilitate the application of the TTC approach, the original Cramer scheme, and an extended version, have been implemented in Toxtree, a freely available software tool for predicting toxicological effects and mechanisms of action. Building on previous work by Patlewicz and coworkers, this report provides some suggestions for improving the Cramer scheme based on a review of the scientific literature, a survey of Toxtree users, and an analysis of lists of body and food components incorporated in Toxtree.

How to obtain EU publications

Our priced publications are available from EU Bookshop (http://bookshop.europa.eu), where you can place an order with the sales agent of your choice.

The Publications Office has a worldwide network of sales agents. You can obtain their contact details by sending a fax to (352) 29 29-42758.

The mission of the JRC is to provide customer-driven scientific and technical support for the conception, development, implementation and monitoring of EU policies. As a service of the European Commission, the JRC functions as a reference centre of science and technology for the Union. Close to the policy-making process, it serves the common interest of the Member States, while being independent of special interests, whether private or national.

JRC
EUROPEAN COMMISSION

9 789279 208041

Publications Office