Regulatory Toxicology and Pharmacology 77 (2016) 13–24

Contents lists available at ScienceDirect

Regulatory Toxicology and Pharmacology

journal homepage: www.elsevier.com/locate/yrtph

# Principles and procedures for implementation of ICH M7 recommended (Q)SAR analyses☆

Alexander Amberg [a], Lisa Beilke [b], Joel Bercu [c], Dave Bower [d], Alessandro Brigo [e], Kevin P. Cross [d], Laura Custer [f], Krista Dobo [g], Eric Dowdy [c], Kevin A. Ford [h], Susanne Glowienke [i], Jacky Van Gompel [j], James Harvey [k], Catrin Hasselgren [d], Masamitsu Honma [l], Robert Jolly [m], Raymond Kemper [n], Michelle Kenyon [g], Naomi Kruhlak [o], Penny Leavitt [f], Scott Miller [d], Wolfgang Muster [e], John Nicolette [p], Andreja Plaper [q], Mark Powley [o], Donald P. Quigley [d], M. Vijayaraj Reddy [r], Hans-Peter Spirkl [a], Lidiya Stavitskaya [o], Andrew Teasdale [s], Sandy Weiner [t], Dennie S. Welch [p], Angela White [k], Joerg Wichard [u], Glenn J. Myatt [d, *]

[a] *Sanofi-Aventis Deutschland GmbH, Frankfurt, Germany*
[b] *Toxicology Solutions, San Diego, CA, USA*
[c] *Gilead, Foster City, CA, USA*
[d] *Leadscope, Columbus, OH, USA*
[e] *Roche Pharmaceutical Research & Early Development, Pharmaceutical Sciences, Roche Innovation Center, Basel, Switzerland*
[f] *Bristol-Myers Squibb, New Brunswick, NJ, USA*
[g] *Pfizer, Groton, CT, USA*
[h] *Genentech, South San Francisco, USA*
[i] *Novartis Institutes for Biomedical Research, Basel, Switzerland*
[j] *Janssen, Beerse, Belgium*
[k] *GlaxoSmithKline, Ware, Herts, UK*
[l] *National Institute of Health Sciences, Tokyo, Japan*
[m] *Eli Lilly and Company, Indianapolis, IN, USA*
[n] *Vertex, Boston, MA, USA*
[o] *FDA Center for Drug Evaluation and Research, Silver Spring, MD, USA*
[p] *AbbVie Inc., North Chicago, IL, USA*
[q] *KRKA, Novo Mesto, Slovenia*
[r] *Merck Research Laboratories, West Point, PA, USA*
[s] *AstraZeneca, Macclesfield, Cheshire, UK*
[t] *Janssen, Spring House, PA, USA*
[u] *Bayer HealthCare, Berlin, Germany*

ARTICLE INFO

*Article history:*
Received 3 February 2016
Accepted 5 February 2016
Available online 11 February 2016

*Keywords:*
ICH M7

ABSTRACT

The ICH M7 guideline describes a consistent approach to identify, categorize, and control DNA reactive, mutagenic, impurities in pharmaceutical products to limit the potential carcinogenic risk related to such impurities. This paper outlines a series of principles and procedures to consider when generating (Q)SAR assessments aligned with the ICH M7 guideline to be included in a regulatory submission. In the absence of adequate experimental data, the results from two complementary (Q)SAR methodologies may be combined to support an initial hazard classification. This may be followed by an assessment of additional information that serves as the basis for an expert review to support or refute the predictions. This paper

☆ FDA CDER Disclaimer: The findings and conclusions in this manuscript have not been formally disseminated by the FDA and should not be construed to represent any agency determination or policy. The mention of commercial products, their sources, or their use in connection with material reported herein is not to be construed as either an actual or implied endorsement of such products by the Department of Health and Human Services.

* Corresponding author. Present address: Leadscope, Inc., 1393 Dublin Road, Columbus, OH 43215, USA.
*E-mail address:* gmyatt@leadscope.com (G.J. Myatt).

http://dx.doi.org/10.1016/j.yrtph.2016.02.004

Impurities
(Q)SAR
Mutagenic impurities
Ames test
Toxicity databases
Expert review

elucidates scenarios where additional expert knowledge may be beneficial, what such an expert review may contain, and how the results and accompanying considerations may be documented. Furthermore, the use of these principles and procedures to yield a consistent and robust (Q)SAR-based argument to support impurity qualification for regulatory purposes is described in this manuscript.



## 1. Introduction

The ICH M7 guideline ("Assessment and control of DNA reactive (mutagenic) impurities in pharmaceuticals to limit potential carcinogenic risk") provides a framework for assessing and controlling DNA reactive impurities in a pharmaceutical product (ICH M7, 2015a). The guideline describes the process whereby actual and potential impurities or degradation products likely to be present in the drug substance and drug product are identified and outlines how a hazard assessment should be performed. When no adequate experimental mutagenicity and/or carcinogenicity results are available, a structure-based computational toxicology or (Q)SAR[1] analysis may be able to predict the mutagenic potential of an impurity. The hazard assessment process leads to the assignment of each impurity to one of five classes described in Table 1. Briefly, class 1 impurities are to be controlled "… *at or below compound-specific acceptable limit*" (ICH M7, 2015b), class 2 or 3 impurities are to be controlled at or below acceptable limits (appropriate Threshold of Toxicological Concern or TTC) and classes 4 and 5 are to be treated as non-mutagenic impurities (ICH M7, 2015a; Kasper and Müller, 2015).

Prior to the publication of ICH M7, many regional guidance documents and scientific papers were published, each contributing to the thought process followed in a mutagenic impurity risk assessment (EMA, 2006, 2010; FDA, 2008; Müller et al., 2006). This included regulatory guidance documents from the European Medicines Agency (EMA, 2006, 2010) and a draft guidance from the US Food and Drug Administration (FDA, 2008) that outlined a methodology for assessing DNA-reactive compounds based on available data as well as mutagenicity predictions from (Q)SAR models. Sutter et al. (2013) outlined the different (Q)SAR methodologies available and highlighted the importance of applying expert knowledge to predictions, a concept also discussed by Dobo et al. (2012), Kruhlak et al. (2012), Naven et al. (2012), Barber et al. (2015) and Stavitskaya et al. (2015). Dobo et al. (2012) demonstrated improved accuracy with expert input on negative predictions. Powley (2015), Greene et al. (2015), Stavitskaya et al. (2015) and Barber et al. (2015) recently provided additional details concerning the use of expert knowledge in the context of an ICH M7 (Q)SAR analysis and Powley (2015) provided general recommendations concerning the format and content of a (Q)SAR analysis report to support regulatory submission.

The ICH M7 guideline is currently being implemented throughout the pharmaceutical industry and international regulatory agencies. A number of specific difficulties are being encountered that are not fully addressed in existing publications. These include: (1) the process of assessing the adequacy of sufficient *in vivo* and/or *in vitro* data; (2) the generation of an overall assessment from the two (Q)SAR methodologies which individually generate positive, negative, or inconclusive predictions as well as out-of-domain classifications; (3) when to apply expert knowledge that could potentially refute a (Q)SAR prediction; (4) what rationale may be considered for use in such an expert review; and (5) an outline for a standardized report to ensure the results are consistently documented, transparent and complete.

Fig. 1 summarizes the process of implementing a (Q)SAR analysis of potential mutagenic impurities. The first step is to collect any relevant data from public sources (such as from the literature) for each impurity. This information can be supplemented with relevant in-house test results. In general, adequate negative bacterial mutagenicity and/or carcinogenicity laboratory data are sufficient to assign the impurity to class 5, whereas adequate positive data would result in assigning the impurity to classes 1 or 2. The adequacy of the data used in these classifications should be critically reviewed. In the absence of adequate data, a (Q)SAR analysis may be used for this class assignment. The (Q)SAR results are used to assign the impurity to ICH M7 classes 3–5. This may include the generation of an expert review to accept or refute any predictions. Positive overall assessments are assigned to class 3, with negative overall assessments generally assigned to class 5; however, where a specific argument based on shared alerts with a compound known to be non-mutagenic is made, these compounds may be assigned to class 4.

This paper outlines a number of practical principles and procedures that can be used in generating a (Q)SAR assessment aligned with ICH M7 as part of a regulatory submission, including accompanying expert analysis. The paper provides a brief overview of the process of identifying and reviewing available data from public and in-house databases as well as the literature. In the absence of adequate data, the principles for combining the (Q)SAR results from

**Table 1**
Definition of the ICH M7 hazard classifications.

| Class | Definition |
|---|---|
| 1 | Known mutagenic carcinogens |
| 2 | Known mutagens with unknown carcinogenic potential (bacterial mutagenicity positive,[a] no rodent carcinogenicity data) |
| 3 | Alerting structure, unrelated to the structure of the drug substance; no mutagenicity data |
| 4 | Alerting structure, same alert in drug substance or compounds related to the drug substance (e.g., process intermediates) which have been tested and are non-mutagenic |
| 5 | No structural alerts, or alerting structure with sufficient data to demonstrate lack of mutagenicity or carcinogenicity |

[a] Or other relevant positive mutagenicity data indicative of DNA-reactivity-related induction of gene mutations (e.g., positive findings in *in vivo* gene mutation studies).

[1] The term "(Q)SAR" refers to (Quantitative) Structure-Activity Relationship and is used as an acronym for computational models that predict a biological response (such as mutagenicity) based on the chemical structure of the test molecule. The term collectively refers to both quantitative and non-quantitative structure-activity relationships by placing the "Q" in parentheses.

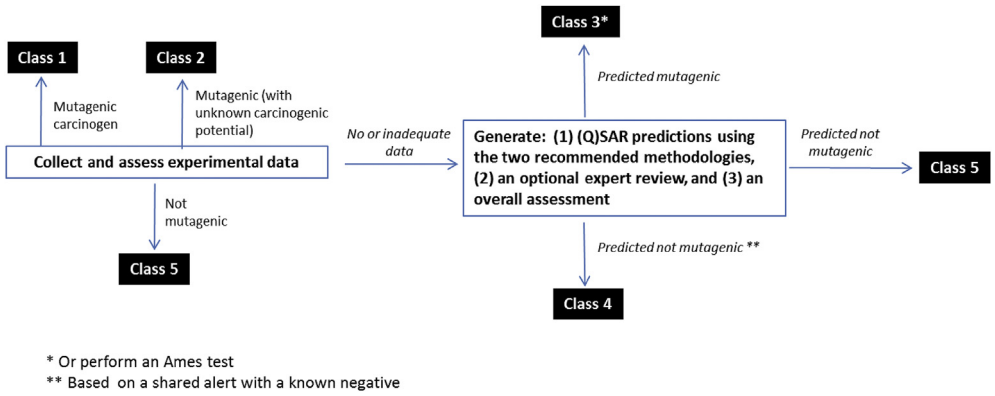



**Fig. 1.** Flow chart depicting an ICH M7 (Q)SAR assessment.

complementary methodologies will be described. This paper will discuss when and how to generate a supplemental expert review that may concur with or refute any prediction. A series of case studies are presented to illustrate the different principles and procedures described. Many of these case studies are from pharmaceutical projects but have not been reviewed or accepted by a regulatory authority unless stated otherwise. This paper will also provide suggestions detailing the contents of an expert analysis and delineate its inclusion in a regulatory submission.

## 2. Assessing available data

According to the ICH M7 guideline (ICH M7, 2015a), the first step in the hazard assessment is "… *database and literature searches for carcinogenicity and bacterial mutagenicity data …*" Since data may have been generated within a pharmaceutical manufacturer's organization, a search of proprietary in-house data may be performed alongside open access or commercial database searches. Table 2 lists a number of open access and commercial databases containing mutagenicity and/or carcinogenicity data. Since it is unrealistic to search all possible databases individually, utilizing a database containing up-to-date information from many of these sources provides a useful alternative. A number of such services are described in Table 2.

In addition ICH recently published a draft addendum to ICH M7. Included within this addendum are a series of permissible limits for a range of commonly used reagents (ICH M7, 2015b).

The focus of ICH M7 is on DNA-reactive impurities, which are generally identified using the Bacterial Reverse Mutation Assay, commonly referred to as the Ames assay (OECD, 1997). An Ames assay may have been performed on the specific impurity, either by the pharmaceutical manufacturer or identified from a search of open access or commercial databases. Any results from a database search should return information necessary to understand the adequacy of the study. An adequately performed negative bacterial mutagenicity study is generally sufficient to assign the impurity to class 5, which is treated as a non-mutagenic impurity. Positive results may be used to assign the impurity to class 2 (known mutagens with unknown carcinogenic potential). The adequacy of any Ames data used in both the class 2 or class 5 assignments should be critically reviewed as discussed in Greene et al. (2015), in line with the principles of Klimisch (Klimisch et al., 1997) as well as be generally consistent with the discussion in Note 2 of the ICH M7 guideline (ICH M7, 2015a). These publications indicate that the Ames test data should be available for inspection and should include at least five strains of bacteria, including four strains of *S. typhimurium* (TA1535; TA1537 or TA97a or TA97; TA98; and TA100) as well as *Escherichia coli* WP2 strains or *Salmonella typhimurium* TA102 (which are similar in mutation detection), exposed to the test substance both in the presence and absence of an appropriate metabolic activation system, with concentrations for soluble non-cytotoxic substances up to 5 mg/plate or 5 μl/plate. Studies pre-dating the publication of the OECD guideline are generally acceptable when they were performed in a manner consistent with the OECD guideline (OECD, 1997).

Pending sufficient justification (e.g., difficult to synthesize impurities), data from other study designs, using fewer test strains or lower drug concentrations, may be used when the quality of the data and study design is considered appropriate. Decisions to accept suboptimal assays may be influenced by an analysis of the risk versus the benefit. Deviations from the standard test protocols are acceptable in certain situations, for example, where a limited number of strains have been tested yet it has been shown that those strains are sensitive to any identified structural alert, as outlined in Note 2 of the ICH M7 guideline (ICH M7, 2015a). The assessment of data may also take into account structural classes that result in false positives under certain experimental conditions, such as an interaction between a test material containing an acid halide or sulfonyl halide and DMSO in the Ames test (Amberg et al., 2015). It should be noted that Ames data tested on a limited number of strains may be considered as part of the weight of evidence in any accompanying expert analysis. Validation statistics of limited strain models can be used to support the expert analysis (Diehl et al., 2000; Zeiger et al., 1985). Other reported genetic toxicity testing battery results are not generally relevant in this context, but may be considered on a case-by-case basis when no or inadequate Ames data are available, such as, positive mouse lymphoma studies with increases in large colonies, when the assay and data meet up to date criteria for positive results (OECD 490, 2015).

The ICH M7 Addendum (Step 2) discusses what factors constitute an adequate rodent carcinogenicity study (ICH M7, 2015b). An adequate negative rodent carcinogenicity study is sufficient to categorize the impurity as class 5. A positive result with evidence of a mutagenic mechanism from an adequately performed study may be used to categorize the compound as class 1 (known mutagenic carcinogen). There may also be situations where a compound is positive in the rodent carcinogenicity study and negative in the bacterial mutagenicity study. For example, carcinogens that are negative in the bacterial mutation study may act through a non-mutagenic mechanism such as by causing hormonal imbalance or proliferative changes leading to cancer. When mechanisms are clearly demonstrated, these cases are considered outside the scope of ICH M7. When a genotoxic threshold is demonstrated per ICH M7 in an *in vivo* follow-up test e.g. rat micronucleus, a Permissible Daily

**Table 2**
Databases containing information on carcinogenicity and mutagenicity data.

| Database | Description | Reference |
|---|---|---|
| ATSDR | Open access database from the Agency for Toxic Substances and Disease Registry (ATSDR) includes toxicological profiles for the hazardous substances including genotoxicity | ATSDR, 2015 |
| CCRIS | Open access database covering chemical carcinogens, including structures and experimental data, covering the period 1985–2011 | Young, 2002; CCRIS, 2011 |
| CPDB | Open access Carcinogenicity Potency DataBase covering the period 1980–2011 | Gold, 1997, 2001, 2005; CPDB 2011 |
| DSSTox | Open access Distributed Structure-Searchable Toxicity (DSSTox) Database Network including content from other sources (e.g. CPDB, ISSCAN) | DSSTox-Archive, 2012 |
| ECHA | Open access European Chemicals Agency (ECHA) database containing actual data and read across results for chemicals manufactured and imported in Europe | ECHA, 2015 |
| ExPub | Commercial application that includes access to the GENE-TOX and CCRIS databases | ExPub, 2015 |
| GENE-TOX | GENE-TOX provides genetic toxicology (mutagenicity) test data from expert peer review of open scientific literature for more than 3000 chemicals from the United States Environmental Protection Agency (EPA) | GENE-TOX, 1998 |
| IARC | Open access International Agency for Research on Cancer (IARC) monographs including carcinogenicity classification | IARC, 2015 |
| IPS INCHEM | Open access International Program on Chemical Safety search for variety of summary documents | INCHEM, 2015 |
| IRIS | Open access data from the EPA in support of human health risk assessment, focusing on hazard identification and dose –response assessment | IRIS, 2015 |
| ISSCAN | Open access database on chemical carcinogens, including structures and experimental data from Istituto Superiore di Sanità | Benigni et al., 2008 |
| JECDB | Open access Japanese Existing Chemical Data Base (JECDB) containing high production volume chemicals | JECDB, 2015 |
| Leadscope | Commercial genetic toxicity and rodent carcinogenicity databases from numerous sources (including US FDA CDER product approval reviews, FDA CFSAN, NTP, CCRIS, and so on) as well as ongoing data harvesting from the literature. Currently includes genetic toxicity data for 11,028 compounds and 179,732 test results and rodent carcinogenicity data for 3598 compounds and 11,538 test results. | Leadscope, 2015 |
| MultiCASE | QSAR model training sets containing mutagenicity and rodent carcinogenicity data from public and proprietary sources including the FDA, GENETOX, NTP, CCRIS and IARC. | MultiCASE, 2015 |
| NTP | Open access database of National Toxicology Program results | Tennant, 1991; NTP, 2015 |
| PAN | Open access Pesticide Action Network (PAN) Pesticide Database | PAN, 2014 |
| Pharma Pendium | Commercial toxicity data from FDA and EMA approval documents | Pharmapendium, 2015 |
| RTECS | Commercial database available through third parties (e.g. Leadscope) currently containing 10,517 Tumorigenic studies for 3724 compounds and 46,385 Mutation studies for 13,343 compounds | Sweet, 1999; RTECS, 2015 |
| ToxNet/ ChemIDPlus | Open access on-line toxicity search system from the US National Library of Medicine with access to archived versions of CCRIS, GENE-TOX, CPDB | Wexler, 2001; ToxNet, 2015 |
| TRACE from BIBRA | Commercial service for TRACE includes information from peer-reviewed toxicology and nutrition journals as well as secondary sources and websites. In addition to the primary literature on the health effects of chemicals, TRACE covers official publications and evaluations issued by authoritative groups. | Anderson, 2000; BIBRA, 2015; Robinson, 2000 |
| VITIC from Lhasa Limited | Commercial data from published and unpublished sources (15,000 records for carcinogenicity and nearly 95,000 records with mutagenicity Ames data) from a number of sources including IARC Monographs, European Chemicals Bureau (IUCLID) and NTP. | VITIC, 2015 |

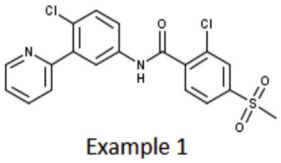
Exposure (PDE) approach may be considered (ICH M7, 2015a).

### 2.1. Case study 1: identifying a compound with historical data

In case study 1, a public database search identified a historical bacterial mutagenicity study with a negative result for the impurity, as shown in Fig. 2. This search identified a 5-strain Ames study by which the compound may be assigned to class 5 due to sufficient evidence for absence of mutagenicity in an adequately performed *in vitro* reverse mutation assay.

## 3. Generating (Q)SAR predictions

In the absence of sufficient experimental mutagenicity and/or carcinogenicity data for a specific impurity, the ICH M7 guideline recommends the use of (Q)SAR models for evaluating the mutagenic potential. This (Q)SAR assessment should utilize models that focus on "… *bacterial mutagenicity predictions …*" and the guideline suggests the use of the two complementary methodologies: "*expert rule-based*" *and* "*statistical-based.*" The guideline goes on to state that the "… *(Q)SAR models … should follow the general validation*



Example 1

| | |
|---|---|
| **Study call:** | Negative |
| **Title:** | GDC-0449.1: Salmonella-Escherichia coli/Mammalian-Microsome Reverse Mutation Assay with a Confirmatory Assay |
| **Reference:** | http://www.accessdata.fda.gov/drugsatfda_docs/nda/2012/203388Orig1s000PharmR.pdf#page=66 |
| **Study type:** | bacterial mutagenesis |
| **Source:** | cder |
| **Species:** | Salmonella typhimurium (12); Escherichia coli (3) |
| **Strains:** | TA98 (3); TA100 (3); TA1535 (3); TA1537 (3); WP2uvrA (3) |
| **Metabolic activation:** | Present (10); Absent (5) |
| **Metabolic activation system:** | S9 (10) |
| **Test calls:** | Negative (15) |
| **Dose summary:** | 0.01 mg (15), 0.0333 mg (15), 0.1 mg (15), 0.333 mg (15), 1.0 mg (15), 2.0 mg (15), 5.0 mg (15) |
| **Study Report:** | Study Report.pdf |

**Fig. 2.** Example 1 showing the results of a database search.

*principles set forth by the Organisation for Economic Co-operation and Development (OECD).*" (OECD, 2007a)

Commonly used statistical-based models include the Leadscope Genetox Statistical QSAR, CASE Ultra from MultiCASE, Inc., and Sarah Nexus from Lhasa Limited and commonly used expert rule-based methodologies include the Leadscope Genetox Expert Alerts and Derek Nexus from Lhasa Limited. The most recent version of each model is preferred for the (Q)SAR analysis; however, it is generally accepted that there are limited changes between different versions and that in practice there are few if any reported changes in overall predictions, in particular of negative predictions being reversed. Recommendations for setting computational model parameters have been provided by Stavitskaya et al., 2013. For example (at the time of publication), with the Leadscope expert-rule based methodology (Leadscope Model Applier: Genetox Expert Alerts Suite), the domain assessment should be turned on, and with the Leadscope statistical-based methodology (Leadscope Model Applier: Genetox Statistical (Q)SAR Suite) probabilities $\geq 0.6$ set to positive, probabilities $< 0.4$ set to negative and the domain assessment turned on.

(Q)SAR models adhering to OECD principles would ideally generate the following prediction results that can be used directly to assess the individual impurities: positive (predicted to be mutagenic) and negative (predicted to be non-mutagenic). However, there are a number of reasons why a (Q)SAR model does not always generate such a classification. The first reason is that the system may determine that the impurity is *out-of-domain*, that is, it is incapable of making a prediction since the system does not adequately cover the structural features of the impurity (OECD validation principle #3). The second reason is that the prediction results may be categorized as equivocal or indeterminate due to weak or conflicting evidence, such that a definitive prediction cannot be made with adequate confidence. The third is where a prediction system is technically unable to process certain types of chemicals, such as for coordination compounds.[2]

## 4. Considerations for an overall assessment and expert review

### 4.1. Overview

The ICH M7 guideline states that the "… *absence of structural alerts …*" from the two suggested (Q)SAR methodologies is sufficient to assign the impurity to class 5. Since any individual methodology may generate results such as a positive prediction, a negative prediction, an inconclusive prediction, or an out-of-domain assignment, it is important to consider how these individual results may be used to derive an overall mutagenic or non-mutagenic assessment consistent with the language in the guideline. The ICH M7 guideline goes on to state that the results from the (Q)SAR methodologies may, if warranted, be examined further. This expert review may provide "… *additional supportive evidence on relevance of any positive, negative, conflicting or inconclusive prediction and provide a rationale to support the final conclusion.*" (ICH M7, 2015a) This review has been shown to improve performance (Stavitskaya et al., 2015; Sutter et al., 2013) and provide a basis for refuting the (Q)SAR results (Powley, 2015; Stavitskaya et al., 2015; Barber et al., 2015). The following sections outline a series of general principles that describe (1) how an overall assessment may be performed, (2) when an expert review may be provided, and (3) what such an expert analysis may contain.

[2] A coordination complex or metal complex consists of a central atom or ion (generally metallic) and a surrounding array of bound molecules or ions.

### 4.2. Negative assessments and expert reviews

A regulatory evaluation of potentially mutagenic impurities should allow for the analysis of many compounds while maintaining a high degree of sensitivity. This can reasonably be achieved using negative predictions from two recommended (Q)SAR methodologies for each compound, without the need for a detailed expert analysis, as long as the methodologies use an automated domain assessment. If additional verification is desired, a rapid visual inspection of the results by the expert can be used to verify that no valid alerts for mutagenicity with a plausible mechanism were overlooked by the two (Q)SAR methodologies (Powley, 2015; Barber et al., 2015).

#### 4.2.1. Case study 2: clear negative prediction from two methodologies

In Fig. 3, the depicted impurity was automatically determined to be within the applicability domain of both the expert rule-based and the statistical-based models and negative predictions were generated by both methodologies. The statistical-based model considered all atoms and bonds in the analysis (i.e., in this modelling system, no atoms or bonds appear in black) as shown in Fig. 3. A quick review of this information may be sufficient to conclude that the overall prediction for this impurity is non-mutagenic and it can be assigned to class 5.

#### 4.2.2. Case study 3: refuted negative prediction from two methodologies

O-(2-Hydroxyethyl)hydroxylamine is shown in Fig. 4 and had a negative prediction for bacterial mutagenicity using both the expert rule-based and the statistical-based models. However, there is conflicting evidence for the mutagenic response of different hydroxylamine salts in the public domain. It was therefore concluded that a potential mutagenic response on the basis of the hydroxylamine moiety should be further evaluated. O-(2-Hydroxyethyl) hydroxylamine was submitted for Ames assay testing where it induced mutations in strain TA1535 in the absence of S9.

When a negative prediction is made in only a single methodology and an inconclusive prediction or an out-of-domain assignment made in the second methodology, it may be necessary to inspect the results in more detail before generating an overall conclusion. Both situations are discussed in Sections 4.4 and 4.5.

### 4.3. Positive prediction and expert reviews

A positive prediction from either of the methodologies may lead to an overall positive prediction. Positive predictions may be refuted through an expert analysis, if appropriate. There are several issues to consider when writing an expert review refuting positive (Q)SAR results including the relevance of any alerting features or corresponding training set compounds, the ability of the chemical environment proximate to the alerting feature to mitigate the mutagenicity and information from chemical analogs (Powley, 2015; Stavitskaya et al., 2015; Barber et al., 2015). A positive assessment may be based on results from a single or multiple

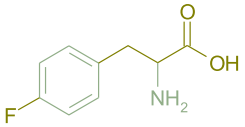



**Fig. 3.** Example 2. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

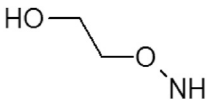



**Fig. 4.** Example 3 (O-(2-Hydroxyethyl)hydroxylamine).

models and each positive result should be individually evaluated as the underlying reasons for the positive result may be different. The following sections outline different points to be considered when refuting a positive prediction.

*4.3.1. Shared alert with known negative (ICH M7 class 4)*

The ICH M7 guideline includes the following statement: "*An impurity with a structural alert that is shared (*e.g.*, same structural alert in the same position and chemical environment) with the drug substance or related compounds can be considered as non-mutagenic … if the testing of such material in the bacterial mutagenicity assay was negative.*" (ICH M7, 2015a) The first step is to identify the structural basis for the impurity's (Q)SAR result (from the matched expert rule and/or the statistical-based model(s)). Next, a related compound with negative Ames data (such as the Active Pharmaceutical Ingredient or API, or another related impurity) is identified that also contains the same highlighted structural features ("known negative"). The following questions may then be asked:

- Are there any additional structural alerts present in the impurity that are not present in the known negative comparator compound? If so, it may not be possible to completely refute the positive (Q)SAR result and apply the class 4 argument.
- Is the alert in the same chemical environment in the impurity as in the comparator compound? Chemical reactivity of an alerting moiety may be mitigated by the presence of another feature in both molecules. Factors to consider in this comparison include (1) differences in the electron charge density (i.e. electron rich or electron deficient) around the specific alerting structure, (2) the steric environment proximal to the alerting structure, (3) the solubility or (4) the size or shape of the impurity.

*4.3.1.1. Case study 4: refuting a positive prediction based on an ICH M7 class 4 analysis.* Fig. 5 represents a series of similar impurities that were predicted to be positive in the statistical-based model. The common features responsible for the positive prediction are summarized and highlighted in red in Fig. 5. 5-strain GLP Ames data conducted according to OEDC 471 and ICH S2(R1) guidelines were generated for one structure (known negative) and were applied to other impurities where R, R1, R2 or R3 varied but without additional alerting functionality (shown in Fig. 5). The impurities in case study 4 are considered analogs of the known negative compound and all share the same highlighted positive structural features. The known negative comparator in combination with negative predictions in the expert rule-based model was sufficient to predict that the impurities would also be non-mutagenic and should be assigned to class 4.

*4.3.1.2. Case study 5: refuting a positive prediction based on an ICH M7 class 4 analysis.* Example 5 (shown in Fig. 6) was predicted to be positive by both the expert rule-based and the statistical-based models as a result of the primary aromatic amine (highlighted in red). The most relevant and structurally similar analog in the rule-based alert system, and the only analog containing a similarly substituted aniline as is present in the test structure, was experimentally negative for bacterial mutation (shown in Fig. 6) (NTP, 1980). It has been reported that the trifluoromethyl groups in the meta position to the amine are strongly deactivating for mutagenicity (Ahlberg et al., 2016). The most analogous structure from the alert in the statistical model was run in a different rule-based system and was predicted negative, since it contains a strong deactivating group. Example 5 is also fully contained within the drug substance, for which the GLP Ames assay was negative. The weight-of-evidence suggests that it is unlikely to be mutagenic and was therefore assigned to class 4. This expert review has been reviewed and accepted by a regulatory authority.

*4.3.2. An explanation of the mechanism*

A positive prediction is triggered by an alert or a significant statistical-based model feature that is present in the impurity. This fragment's associated mutagenic potential may be based on a reasonable mechanistic rationale and/or there may be sufficient positive examples matching the fragment; however, the environment around the alerting moiety within this specific impurity may preclude reaction at this site. It is possible to construct an expert review to refute the prediction (Powley, 2015; Stavitskaya et al., 2015; Barber et al., 2015). In situations where a compound is predicted negative by an expert rule-based methodology, yet predicted positive by a statistical-based methodology, it may be helpful to understand why the compound containing any highlighted group is not positive in the expert rule-based system. Does the alert definition contain any exceptions to the rule?

*4.3.2.1. Case study 6: refuting a positive prediction based on a mechanism analysis.* In Example 6, the potential impurity was predicted to be positive by the statistical-based model but negative by the expert rule-based model. As shown in Fig. 7, the main contribution to the positive prediction by the statistical-based model was the feature highlighted in red. In reviewing the compounds supporting the alerting fragment, it was found that the alerting fragment was highly influenced by the mutagenicity data on alkyl sulfonate esters, dialkyl sulfates, or sultones (see Fig. 8), which are known alerts for mutagenicity (Ashby and Tennant, 1988; Benigni and Bossa, 2008). Example 6 is a mono-alkyl sulfate esters; these are consistently negative in the Ames assay (OECD, 2007b) and are not alkylating agents. Mono-alkyl sulfate esters are negatively charged at physiological pH and therefore are less electrophilic than their alkyl sulfonate counterparts. The mono-alkyl sulfate esters in the training set were also non-


**Fig. 5.** A series of chemicals all predicted to be positive in a statistical-based model based on the feature highlighted in red. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

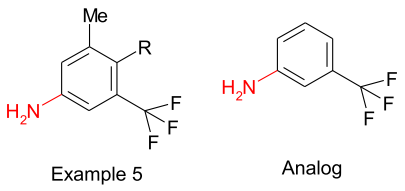



**Fig. 6.** Example 5 and analog. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

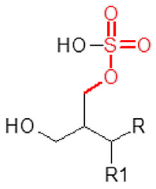



**Fig. 7.** Example 6 predicted to be positive in the statistical-based methodology, primarily based on the feature highlighted in red. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

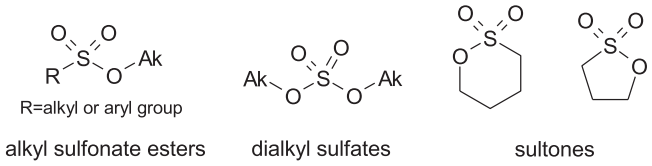



**Fig. 8.** Structural definitions for alkyl sulfonate esters, dialkyl sulfates, and sultones.

mutagenic except for the mono-alkyl sulfate esters of known mutagenic poly aromatic hydrocarbons such as benz(a)anthracene and chrysene. Therefore, Example 6 is predicted to be non-mutagenic.

### 4.3.3. *The relevance of features from statistical-based methodologies*

A positive prediction from a statistical-based model may be refuted if the structural features that are the basis for the "alert" in the model ("positive contributing features") are not relevant, illustrated as follows:

- **Coincidental features:** Structural features are identified through machine-learning when building statistical-based models. Positive features are identified when present in a group of predominantly mutagenic training set compounds. However, these mutagenic compounds could also contain other structural features that better represent the actual moiety responsible for the observed mutagenicity. In these cases, the statistical model has identified coincidental features. If the positive prediction was based primarily upon these coincidental features then an expert analysis refuting the prediction may be made (Powley, 2015; Barber et al., 2015). One example of such a situation is where an amine oxide is flagged in a set of aromatic nitro compounds.
- **Mitigating features**: A positive prediction may be refuted if the positive model features are mitigated by negative features present at or proximal to the same reaction center.
- **Limited training set examples**: It is possible that a positive model feature was derived from a small number of examples. An expert analysis may refute a positive prediction made primarily using such features.
- **No significant positive model features**: The positive prediction may result from very small contributions from many unrelated or unconnected positive model features.
- **Irrelevant training set examples**: It is possible that a positive model feature was derived from a set of compounds covering multiple structural classes. It is also possible that some of these structural classes do not apply to the specific impurity (they are part of a different chemical series) and an expert review to refute the positive prediction may be an option if the impurity is within one of the non-mutagenic chemical classes.
- **Underlying data are incorrect or not adequate**: It may be possible to identify model features based on data that are not correct as a result of certain experimental conditions, such as an interaction between a test material containing an acid halide or sulfonyl halide and DMSO in the Ames test (Amberg et al., 2015).

*4.3.3.1. Case study 7: refuting a positive prediction based on a mechanism and coincidental features.* In this example the potential impurity was predicted to be positive by the statistical-based model but negative by the expert rule-based model. Example 7 is an *N*-oxide of a non-aromatic amine bearing a phenyl/aryl group. The major contributing features are highlighted in red in Fig. 9. Firstly, the training sets inadequately represent *N*-oxide of a non-aromatic amine bearing a phenyl/aryl group, whose predicted mutagenic activity was influenced by other co-occurring alerting features. Secondly, the literature indicates that the tertiary alkyl amine *N*-oxides are non-mutagenic. Finally, a structural analysis was performed for mutagenicity on nitrogenous aryl compounds and their corresponding *N*-oxides using $TRP^{+}$ reversion in *E. coli* (Pai et al., 1978). This structural analysis included 10 tertiary aryl amines and their corresponding *N*-oxides. As part of the weight of the evidence, it was concluded that primary aromatic amines and their corresponding hydroxylamines, and *N*-hydroxycarbamates were mutagenic, but not the tertiary aryl amines or their corresponding tertiary *N*-oxides, as shown in Fig. 10 (Pai et al., 1978). Hypothetically, dealkylation of the amine to yield a primary aromatic amine is a potential mechanism of mutagenicity; however, in case study 7, this would yield aniline, which is known to lack mutagenic potential. The lack of mutagenicity following dealkylation is further supported by the observation that the parent drug substance (API) structure contains the corresponding primary aromatic amine and was negative in the bacterial reverse mutation assay. Based on analysis of the training sets, a negative expert rule-based prediction, literature analysis for tertiary amine *N*-oxides, and its structural similarity to the drug substance, Example 7 is predicted to be non-mutagenic.

*4.3.3.2. Case study 8: refuting a positive prediction based on coincidental features.* Example 8 is shown in Fig. 11 and was predicted to be positive by the statistical-based model and negative by the expert rule-based model. An expert review of Example 8 described below concluded that the probability of mutagenicity is low based on a review of the training set and a comparison with the drug substance, which was negative in the bacterial reverse mutation assay. The most relevant model features were evaluated and found to contain examples of another alert more likely to be responsible for the positive prediction (see supplemental material for more details). These features included a planar anthracene-like tricyclic aromatic core; however, the polycyclic core of Example 8 is puckered, due to the presence of sp3 carbon atoms, with C—H bonds almost orthogonal to either plane defined by any two fused rings, hence making the structure non-planar. Example 8 was therefore predicted to be non-mutagenic.

*4.3.3.3. Expert reviews based on chemical analogs from public or in-house sources.* Experimental Ames data for structural analogs can

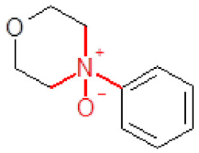



**Fig. 9.** Example 7 with features contributing to the positive prediction highlighted in red. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

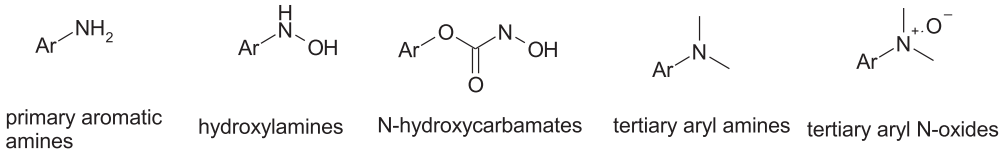



**Fig. 10.** Structural definitions for primary aromatic amines and their corresponding hydroxylamines, and N-hydroxycarbamates as well as tertiary aryl amines and corresponding tertiary N-oxides.

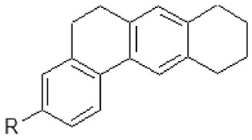


**Fig. 11.** Example 8.

also be used when the training sets do not contain suitable numbers of related structures (Powley, 2015; Barber et al., 2015; Stavitskaya et al., 2015). Sufficiently similar analogs from the literature, public databases or in-house information may be used to provide justification for refuting a positive or overruling an inconclusive prediction. The number of analogs and the degree of structural similarity needs to be assessed on a case-by-case basis (Powley, 2015).

*4.3.3.4. Case study 9: refuting a positive prediction using data from chemical analogs.* Example 9 was predicted to be positive by the statistical-based methodology and negative by the expert rule-based methodology. A database search identified a number of close analogs as shown in Fig. 12. All analogs were experimentally non-mutagenic in the Ames assay. The extension of the carbon side chain of Example 9 should not increase its reactivity compared to the analogs. Example 9 is therefore predicted to be non-mutagenic. This compound, in fact, has been shown to be experimentally negative for bacterial mutagenicity (Carmellino, 1993). This example is used to illustrate the concept of an analog search and, as part of this analysis, it is necessary to assess the adequacy of the underlying Ames data.

### 4.4. Expert reviews for inconclusive (Q)SAR results

Inconclusive predictions are generated when there is not enough evidence to make a mutagenic or non-mutagenic prediction with adequate confidence. In general, all approaches discussed earlier to refute a positive or negative prediction can reasonably be applied to an inconclusive prediction for a covered (i.e., within the applicability domain) compound in an attempt to resolve the prediction and generate a negative or positive overall conclusion. The following outlines several potential approaches to assessing the results as part of an expert review to reach a conclusion that the impurity is likely mutagenic or non-mutagenic.

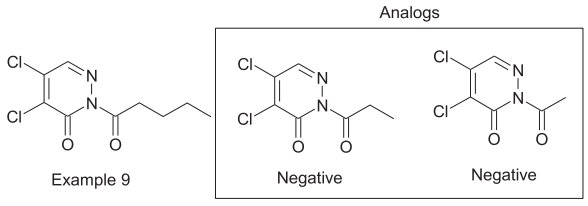



**Fig. 12.** Example 9 with analogs.

- **Visual inspection by an expert**: One approach to assess inconclusive predictions is for a chemist or toxicologist to visually inspect the results to verify there are no valid alerts for mutagenicity with a plausible mechanism. For example, the chemist or toxicologist who is visually inspecting the results may have knowledge of mutagenicity alerts and/or mechanisms derived from proprietary data not built into the (Q)SAR models. It may be important to consider portions of the molecule (e.g. functional groups) not represented in the (Q)SAR models. Systematic substructural searching of functional groups not considered by the models may also support the identification of features that are positively associated with bacterial mutagenicity data (i.e. there is a statistically significantly greater number of positive examples than would be expected by chance). The expert may also consider whether the structural features highlighted by the statistical-based models show significant association with bacterial mutagenicity.
- **Strength of a single prediction**: Where only a single methodology has generated a prediction, an assessment of the strength of this prediction may be made to determine whether it is sufficient as the basis of an overall conclusion.

*4.4.1. Case study 10: assessing an inconclusive prediction using the literature*

Example 10 (Fig. 13) was predicted to be negative by the expert rule-based methodology and inconclusive by the statistical-based methodology; in the latter the most significant contribution was from the primary aromatic amine. As discussed in Ahlberg et al. (2016), primary aromatic amines are mutagenic only in the presence of an activating functional group. Both functional groups (the bromo group in the para position and the carboxylate in the ortho position) are not activating according to Ahlberg et al. (2016) (based on an analysis of primary aromatic amine data from public and proprietary databases) and therefore Example 10 was predicted to be non-mutagenic. This compound has been tested in a standard Ames assay using 5 strains and is non-mutagenic (Greene et al., 2015).

*4.4.2. Case study 11: assessing an inconclusive prediction using analogs*

Example 11 (shown in Fig. 14) was predicted to be negative by the expert rule-based model and inconclusive by the statistical-based model. Since Example 11 contains a hydrazine substructure and specific classes of hydrazines are known to be mutagenic, an analysis based on the evaluation of published Ames assay data for

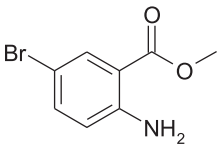


**Fig. 13.** Example 10.

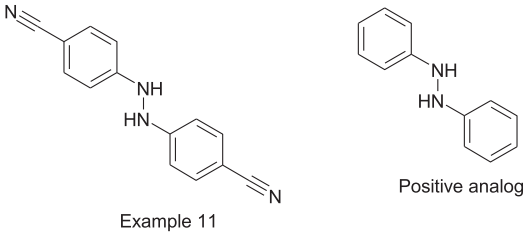



**Fig. 14.** Example 11 with analog.

structural analogs was performed. This assessment led to the identification of numerous structural analogs tested in the Ames assay including the analog shown in Fig. 14 that was reported to be mutagenic. Hence, Example 11 was predicted to be mutagenic.

### 4.5. Expert reviews for "out of domain" statements

When an impurity is presented to a model that is sufficiently different from the types of chemicals used in the reference/training set, the model should not make a prediction, in accordance with OECD validation principle #3. These out-of-domain results, however, may also be assessed as part of an expert review. As background to the analysis, it may be helpful to understand why the model was unable to make a prediction for this specific impurity. In a similar manner to inconclusive results that were discussed earlier, it may be possible to generate an expert review for an out-of-domain result based on: (1) a visual inspection by an expert chemist or toxicologist, (2) an assessment of the strength of a prediction by a single methodology, (3) an understanding of relevant mutagenic mechanisms, and (4) data for structural analogs. Another approach that may be helpful in assessing this type of result is to investigate whether the out-of-domain result is attributable to the addition of a non-reactive group. The first step as part of this assessment is to determine if there are any similar chemicals that were predicted negative or where there is a negative experimental result. If the only difference from the out-of-domain structure is the addition of a non-reactive group (e.g. an amine protected by two tert-butoxycarbonyl (Boc) groups or other non-alerting fragment) and as long as this group could not cause an additional functional group to become an activated alert, then this scenario may be used to address an out-of-domain situation.

Running another model is also an option to address an out-of-domain or indeterminate (Q)SAR prediction; however, it should be noted that running a third model is not required by ICH M7. Similar to the first two models, the third model should also follow the OECD (2007a) (Q)SAR validation principles to ensure that one is simply not running models until one with a less stringent applicability domain calculation is found.

#### 4.5.1. Case study 12: assessing an out-of-domain response based on the mechanism

Example 12 is a large compound containing greater than 30 non-hydrogen atoms (Fig. 15). Example 12 was determined to be out-of-domain by the statistical-based model. The example also contains an aromatic amine moiety which is structurally alerting.

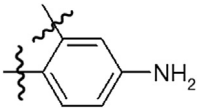



**Fig. 15.** Example 12 (>30 non-hydrogen atoms).

The mechanism of mutagenicity associated with aromatic amines requires oxidation by cytochrome P450 to hydroxylamines and then further activation by *O*-acylation. The *O*-acylated *N*-arylhydroxylamine is converted to a highly electrophilic nitrenium ion, which then reacts with DNA (Benigni and Bossa, 2011). Aromatic amines found within pharmaceutical intermediates are more likely to be negative in the bacterial mutation assay than those that have data available in the public literature, according to an analysis of in-house databases (McCarren et al., 2011). This has been attributed to the bias towards larger molecular weight compounds in drug development with increased steric hindrance to formation of the reactive mutagenic metabolite or decreased ability of the metabolite to cross bacterial cell walls (Glende et al., 2002; Hatch et al., 2001; Benigni, 2005). For example, it has been reported that the addition of bulky alkyl groups away from the amino group changes a mutagenic aromatic amine to a non-mutagenic species (Glende et al., 2002). Hydrolysis or metabolism to generate a small aromatic amine that may be mutagenic is not possible in Example 12. Therefore, Example 12 is predicted to be non-mutagenic due to the size of the compound which results in a potential lower bioavailability, and inhibited formation of the putative reactive nitrenium metabolite.

#### 4.5.2. Case study 13: assessing an out-of-domain prediction using a similar analog

Example 13 was out-of-domain by the statistical-based models and predicted to be negative by the expert rule-based model. Example 13 (shown in Fig. 16) is very similar to the drug substance, which was also out-of-domain for the statistical based models. The change in position was concluded not to change the potential for mutagenic reactivity, since there were no alerting features on the drug substance or the impurity (based on the expert rule-based model). Therefore, based on its structural similarity to the drug substance (which was negative for bacterial mutagenicity in the Ames assay), Example 13 was predicted to be non-mutagenic.

Case study 13 illustrates an expert analysis based on a change of a substituent position. Changes in the position of heteroatoms within the ring can also be important to consider. For example, 3-Aminoisoxazole is non-mutagenic and 5-amino-4-chloro-3-methylisoxazole is mutagenic, as shown in Fig. 17. Both are examples of primary aromatic amines, where the aromatic system is a 5-membered heterocycle and both rings contain a single nitrogen and oxygen; however, the position of these heteroatoms is different in the two compounds relative to the primary aromatic amine. These compounds, along with an analysis of the structure-activity relationship, are discussed in Ahlberg et al. (2016).

#### 4.5.3. Case study 14: assessing an out-of-domain prediction using public analogs

Aminoacetonitrile (Example 14) was out-of-domain for the statistical-based models and predicted to be negative by the expert rule-based model. No standardized Ames testing has been performed with aminoacetonitrile. However, data from structurally

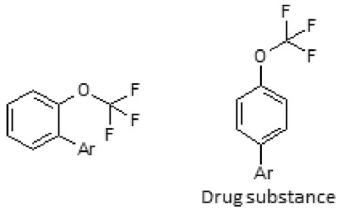



**Fig. 16.** Example 13 alongside the drug substance which is negative in the Ames assay.

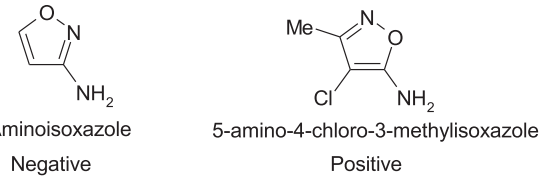



**Fig. 17.** Examples of how the position of heteroatoms may influence mutagenicity.

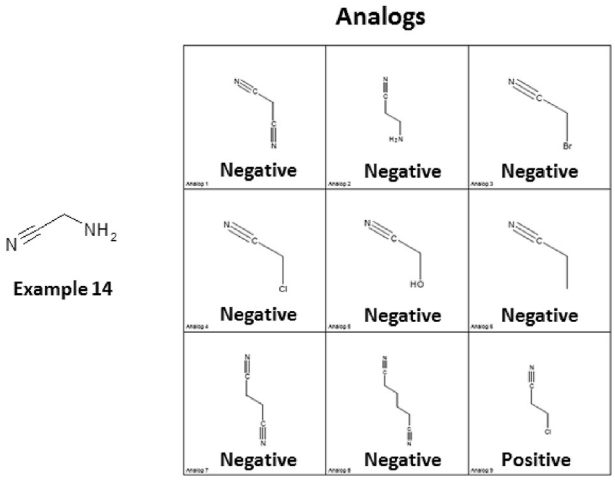



**Fig. 18.** Example 14 with analogs (including Ames results).

similar compounds suggest that it is non-mutagenic (see Fig. 18). This included 3-aminopropionitrile (Analog 2) that was tested negative in TA98, TA100, TA1535, TA1537, and TA1538 with and without metabolic activation. (CCRIS 3-aminopropionitrile) The single analog that was mutagenic (3-chloropropionitrile) contains an additional alerting structure (monofunctional alkyl chloride) not shared with aminoacetonitrile. In addition to these nearest neighbors, aminoacetonitrile is also structurally similar to cyanamide which is also non-mutagenic in a 5-strain Ames assay with *E. coli* (FIOSH, 2014). Therefore, Example 14 was predicted to be non-mutagenic.

#### 4.5.4. Case study 15: assessing an out-of-domain based on the addition of a non-reactive group

Example 15 is the Boc protected form of Compound Y (shown in Fig. 19). Example 15 was predicted to be negative by the expert rule-based methodology but out-of-domain for the statistical-based methodology. Compound Y was predicted to be negative in the statistical-based methodology. Boc protection is used to prevent chemical reactivity of the secondary amine and can be cleaved under acidic conditions (Schelhass and Waldmann, 1996). Therefore, Example 15 is also not predicted to be mutagenic given its similarity and reduced chemical reactivity compared to Compound Y.

Situations can arise where it is not possible to generate a (Q)SAR prediction with either methodology due to the impurity being out-of-domain, or both methodologies returning inconclusive predictions. When no model is able to generate a prediction, a pragmatic approach would be to either perform an Ames test or assign the impurity to class 3. However, in situations when no experimental data are generated, expert knowledge could be used to supersede even these predicted outcomes, with the caveat that it should include justifiable scientific evidence for regulatory acceptance.

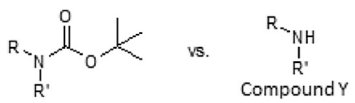



**Fig. 19.** Example 15 with analog Y (predicted negative).

#### 4.5.5. Case study 16: assessing an out-of-domain result from two methodologies

Example 16 (shown in Fig. 20) was concluded to be out-of-domain by both the expert rule-based and the statistical-based models as a result of the novelty of the R-group. Example 16 is similar to the drug substance; the only difference is that the primary amine group of the drug substance has been has been converted to the bis-boc imide, shown in Fig. 20, through Boc protection of the primary amine. The drug substance is also concluded to be out-of-domain by both (Q)SAR methodologies; however, it has been tested and is non-mutagenic in the standard 5-strain Ames test. Since there is no expected reactivity from the bis-boc functionality, Example 16 is predicted to be non-mutagenic (which was confirmed experimentally in a standard 5 strain Ames assay). As in Case Study 14, given the bis-boc protection serves to reduce reactivity, it could be reasonable to classify this as a non-mutagenic compound despite the lack of predictions in both methodologies.

## 5. Reporting

The final report may include a description of the methodologies used, a summary of the results along with any expert reviews that should be transparent and "*include supporting information to arrive at the overall conclusion for Class 4 and Class 5 impurities*" (ICH M7, 2015a). The selection of the impurities to be reported is dependent on the stage of development, as shown in Table 3, which presents a summary from the ICH M7 guideline.

The following elements may be included in the report of a (Q)SAR assessment consistent with ICH M7 with the level of detail dependent on the stage of development:

1. Materials and methods
   - Software, models and databases used, along with version numbers and parameters set
2. Summary of the results and conclusions
   - Chemical structure of the impurity that may include highlighting to illustrate what the software has identified as structural features associated with or not associated with positive bacterial mutagenicity data (when this highlighting can be generated automatically by the system)
   - Experimental data and/or (Q)SAR results from both methodologies (the experimental and (Q)SAR results may be in different tables or sections)
   - Overall conclusion based on the prediction results and any expert review (i.e., mutagenic or non-mutagenic) along with class 1–5 assignment

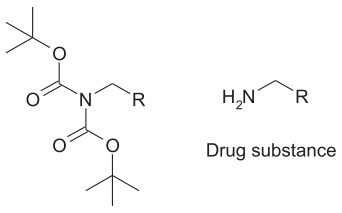



**Fig. 20.** Example 16 alongside the non-mutagenic drug substance.

**Table 3**
Reporting requirements at each development phase.

| Development phase | Reporting requirements |
|---|---|
| Phase 1 clinical trials of 14 days or less | Class 1 and class 2 impurities; cohorts of concern |
| Phase 1 clinical greater than 14 days or Phase 2a clinical trials | Class 1, class 2 and class 3 impurities; cohorts of concern |
| Phase 2b clinical trials or Phase 3 clinical trials | List of actual/potential impurities assessed by (Q)SAR, Class 1, class 2 and class 3 impurities; plan for control, bacterial mutagenicity test results. |
| Common Technical Document (Marketing Application) | List of actual/potential impurities assessed by (Q)SAR, Class 1, class 2, class 3, class 4, and class 5 impurities; supporting information, plan for control, bacterial mutagenicity study reports. |

- Summary of any supporting expert reviews or remarks

3. Supporting information
   - Expert review(s) supporting or refuting the (Q)SAR result, along with examples and references to illustrate
4. References, especially those used to support an expert review, if applicable
5. Appendices
   - Complete bacterial mutagenicity study reports at the time of marketing application may be included in the appendices or cross-referenced or hyperlinked from another section

When models are used that are not familiar to regulatory agencies, it will be necessary to provide additional documentation showing how these models are consistent with the OECD (Q)SAR validation principles (OECD, 2007a). The supplemental information contains an example of a regulatory submission using examples described above.

## 6. Conclusions

The ICH M7 guideline provides a framework for assessing DNA reactive impurities and describes how these impurities may be controlled. This framework is currently being implemented across the pharmaceutical industry and international regulatory agencies. An important component of this guideline is the use of (Q)SAR as an alternative to conventional testing for the assessment of the mutagenic potential of drug substance impurities. (Q)SAR models represent a state-of-the art approach to predicting mutagenicity that balances the need for high-throughput while maximizing patient safety.

This paper has outlined a number of practical principles and procedures to be considered when conducting a (Q)SAR analysis consistent with the ICH M7 guideline. This includes, in the absence of adequate experimental data, how to combine the results from the recommended (Q)SAR models, when to consider generation of a detailed expert review, and what such a review may contain. The contents of a full report for inclusion as part of a regulatory submission have been outlined. Through adoption of common principles and procedures, the practical implementation of a (Q)SAR analysis consistent with the ICH M7 guideline will become more standardized, consistent, and transparent. Additionally, the generation and review of these reports should become more streamlined over time for both pharmaceutical manufacturers and regulatory agencies.

## Acknowledgments

We acknowledge Maurizio Franzini for helping with some of the expert review examples.

## Transparency document

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.yrtph.2016.02.004.

## Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.yrtph.2016.02.004.

## References

Ahlberg, E., Amberg, A., Beilke, L.D., Bower, D., Cross, K.P., Custer, L., Dobo, K., Ford, K.A., VanGompel, J., Harvey, J., Honma, M., Jolly, R., Joossens, E., Kemper, R., Kenyon, M., Kruhlak, N., Kuhnke, L., Leavitt, ., Neilan, C., Naven, R., Quigley, D.P., Shuey, D., Spirkl, H.P., Stavitskaya, L., Teasdale, A., White, A., Wichard, J., Zwickl, C., Myatt, G.J., 2016. Extending (Q)SARs to incorporate proprietary knowledge for regulatory purposes: a case study using aromatic amine mutagenicity. Regul. Toxicol. Pharmacol. 77, 1–12.

Amberg, A., Harvey, J.S., Czich, A., Spirkl, H.-P., Robinson, S., White, A., Elder, D.P., 2015. Do Carboxylic/Sulfonic acid halides really present a mutagenic and carcinogenic risk as impurities in final drug products? Org. Process Res. Dev. 19, 1495–1506.

Anderson, C., Copestake, P., Robinson, L., 2000. A specialist toxicity database (TRACE) is more effective than its larger, commercially available counterparts. Toxicology 151, 37–43.

Ashby, J., Tennant, R.W., 1988. Chemical structure, Salmonella mutagenicity and extent of carcinogenicity as indicators of genotoxic carcinogenesis among 222 chemicals tested in rodents by the U.S. NCI/NTP. Mutat. Research Genetic Toxicol. 204, 17–115.

ATSDR, 2015. http://www.atsdr.cdc.gov/.

Barber, C., Amberg, A., Custer, L., Dobo, K.L., Glowienke, S., Gompel, J.V., Gutsell, S., Harvey, J., Honma, M., Kenyon, M.O., Kruhlak, N., Muster, W., Stavitskaya, L., Teasdale, A., Vessey, J., Wichard, J., 2015. Establishing best practise in the application of expert review of mutagenicity under ICH M7. Regul. Toxicol. Pharmacol. 73, 367–377.

Benigni, R., 2005. Structure–activity relationship studies of chemical mutagens and carcinogens: mechanistic investigations and prediction approaches. Chem. Rev. 105, 1767–1800.

Benigni, R., Bossa, C., Richard, A.M., Yang, C., 2008. A novel approach: chemical relational databases, and the role of the ISSCAN database on assessing chemical carcinogenicity. Ann. Ist. Super. Sanita 44, 48–56.

Benigni, R., Bossa, C., 2008. Structure alerts for carcinogenicity, and the Salmonella assay system: a novel insight through the chemical relational databases technology. Mutat. Res. 659, 248–261.

Benigni, R., Bossa, C., 2011. Mechanisms of chemical carcinogenicity and mutagenicity: a review with implications for predictive toxicology. Chem. Rev. 111, 2507–2536.

BIBRA, 2015. http://www.bibra-information.co.uk/trace-unique-chemical-database/.

Carmellino, M.L., Massolini, G., Pagani, G., Zanif, F., Leri, G., 1993. Fungicidal and genotoxic activity of new substituted pyridazinones. Framaca 48, 1427–1438.

CCRIS, 2011. http://toxnet.nlm.nih.gov/newtoxnet/ccris.htm.

CPDB, 2011. http://toxnet.nlm.nih.gov/newtoxnet/cpdb.htm.

CCRIS, 3-aminopropionitrile. http://toxnet.nlm.nih.gov/cgi-bin/sis/search2/r?dbs+ccris:@term+@rn+151-18-8.

Diehl, M.S., Willaby, S.L., Snyder, R.D., 2000. Comparison of the results of a modified miniscreen and the standard bacterial reverse mutation assays. Environ. Mol. Mutagen. 36, 72–77.

Dobo, K.L., Greene, N., Fred, C., Glowienke, S., Harvey, J.S., Hasselgren, C., Jolly, R., Kenyon, M.O., Munzner, J.B., Muster, W., Neft, R., Reddy, M.V., White, A.T., Weiner, S., 2012. In silico methods combined with expert knowledge rule out mutagenic potential of pharmaceutical impurities: an industry survey. Regul. Toxicol. Pharmacol. 62, 449–455.

DSSTox-Archive, 2012. ftp://ftp.epa.gov/dsstoxftp/DSSTox_Archive_20120411/.

ECHA, 2015. http://echa.europa.eu/.

EMA, 2006. Guideline on the Limits of Genotoxic Impurities. CPMP/SWP/5199/02 EMEA/CHMP/QWP/251344/2006. http://www.ema.europa.eu/ema/pages/includes/document/open_document.jsp?webContentId=WC500002903.

EMA, 2010. Questions and Answers on the 'Guideline on the Limits of Genotoxic Impurities'. http://www.ema.europa.eu/ema/pages/includes/document/open_document.jsp?webContentId=WC500002907.

ExPub, 2015. http://www.expub.com/index.html.

FDA, 2008. Draft guidance for industry – genotoxic and carcinogenic impurities in

drug substances and products: recommended approaches.
FIOSH, 2014 Federal Institute for Occupational Safety and Health, 2014. CLH Report: Proposal for Harmonised Classification and Labelling. Substance Name: Cyanamide. Federal Office for Chemicals, Dortmund, Germany.
GENE-TOX, 1998. http://toxnet.nlm.nih.gov/newtoxnet/genetox.htm.
Glende, C., Klein, M., Schmitt, H., Erdinger, L., Boche, G., 2002. Transformation of mutagenic aromatic amines into non-mutagenic species by alkyl substituents. Part II: alkylation far away from the amino function. Mutat. Res. 515, 15–38 [Research Support, Non-U.S. Gov't].
Gold, L.S., Slone, T.H., Ames, B.N., Manley, N.B., Garfinkel, G.B., Rohrbach, L., 1997. Chapter 1: carcinogenic potency database. In: Gold, L.S., Zeiger, E. (Eds.), Handbook of Carcinogenic Potency and Genotoxicity Databases. CRC Press, Boca Raton, FL, pp. 1–605.
Gold, L.S., Manley, N.B., Slone, T.H., Ward, J.M., 2001. Compendium of chemical carcinogens by target organ: results of chronic bioassays in rats, mice, hamsters, dogs and monkeys. Toxicol. Pathol. 29, 639–652.
Gold, L.S., Manley, N.B., Slone, T.H., Rohrbach, L., Garfinkel, G.B., 2005. Supplement to the carcinogenic potency database (CPDB): results of animal bioassays published in the general literature through 1997 and by the national toxicology program in 1997 and 1998. Toxicol. Sci. 85, 747–800.
Greene, N., Dobo, K.L., Kenyon, M.O., Cheung, J., Munzner, J., Sobol, Z., Sluggett, G., Zelesky, T., Sutter, A., Wichard, J., 2015. A practical application of two in silico systems for identification of potentially mutagenic impurities. Regul. Toxicol. Pharmacol. 72, 335–349.
Hatch, F.T., Knize, M.G., Colvin, M.E., 2001. Extended quantitative structure-activity relationships for 80 aromatic and heterocyclic amines: structural, electronic, and hydropathic factors affecting mutagenic potency. Environ. Mol. Mutagen. 38, 268–291.
IARC, 2015. http://Monographs.iarc.fr/ENG/Classification/index.php.
ICH M7, 2015a. ICH M7 – assessment and control of DNA reactive (mutagenic) impurities in pharmaceuticals to limit potential carcinogenic risk. http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM347725.
ICH M7, 2015b. Addendum to ICH M7: assessment and control of DNA reactive (mutagenic) impurities in pharmaceuticals to limit potential carcinogenic risk: application of the principles of the ICH M7 guideline to calculation of compound-specific acceptable intakes M7 (R1). http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Multidisciplinary/M7/M7_Addendum_Step_2.pdf.
INCHEM, 2015. http://www.inchem.org.
IRIS, 2015. http://www.epa.gov/iris/.
JECDB, 2015. http://www.echemportal.org/echemportal/.
Kasper, P., Müller, L., 2015. Genotoxic impurities in pharmaceuticals. In: Graziano, M.J., Jacobson-Kram, D. (Eds.), Genotoxicity and Carcinogenicity Testing in Pharmaceuticals. Springer, pp. 55–74.
Klimisch, H.-J., Andreae, M., Tillmann, U., 1997. A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data. Regul. Toxicol. Pharmacol. 25, 1–5.
Kruhlak, N.L., Benz, R.D., Zhou, H., Colatsky, T.J., 2012. (Q)SAR modeling and safety assessment in regulatory review. Clin Pharmacol Ther. 91, 529–534.
Leadscope, 2015. http://www.leadscope.com/toxicity_databases/.
McCarren, P., Springer, C., Whitehead, L., 2011. An investigation into pharmaceutically relevant mutagenicity data and the influence on Ames predictive potential. J. Cheminformatics 3, 51.
Müller, L., Mauthe, R.J., Riley, C.M., Andino, M.M., Antonis, D.D., Beels, C., Degeorge, J., Knaep, A.G.D., Ellison, D., Fagerland, J.A., Frank, R., Fritschel, B., Galloway, S., Harpur, E., Humfrey, C.D., Jacks, A.S., Jagota, N., Mackinnon, J., Mohan, G., Ness, D.K., O'Donovan, M.R., Smith, M.D., Vudathala, G., Yotti, L., 2006. A rationale for determining, testing, and controlling specific impurities in pharmaceuticals that possess potential for genotoxicity. Regul. Toxicol. Pharmacol. 44, 198–211.
MultiCASE, 2015. http://www.multicase.com/case-ultra-models#bacterial_muta_bundle.
Naven, R.T., Greene, N., Williams, R.V., 2012. Latest advances in computational genotoxicity prediction. Expert Opin. Drug Metab. Toxicol. 8, 1579–1587.
NTP, 1980. http://tools.niehs.nih.gov/cebs3/ntpViews/?activeTab=summary&studyNumber=758178.
NTP, 2015. http://tools.niehs.nih.gov/cebs3/ui/.
OECD, 1997. Test 471. Bacterial Reverse Mutation Test OECD Guideline for Testing of Chemicals Section 4 1997 July. http://www.oecd-ilibrary.org/environment/test-no-471-bacterial-reverse-mutation-test_9789264071247-en.
OECD, Organisation for Economic Co-operation and Development, 2007a. Guidance Document on the Validation of (Quantitative) Structure-activity Relationships [(Q)SAR] Models, OECD Environment Health and Safety Publications Series on Testing and Assessment No. 69 (ENV/JM/MONO(2007)2).
OECD, Organisation for Economic Co-operation and Development, 2007b. SIDS Initial Assessment Report for SIAM 25. Category of Alkyl Sulfates, Alkane Sulfonates and α-Olefin Sulfonates. 16–19 October 2007.
OECD 490, 2015. http://www.oecd-ilibrary.org/environment/test-no-490-in-vitro-mammalian-cell-gene-mutation-tests-using-the-thymidine-kinase-gene_9789264242241-en.
Pai, V., Bloomfield, S.F., Jones, J., Gorrod, J.W., 1978. Mutagenicity testing of nitrogenous compounds and their N-oxidised products using TRP+ reversion in E. Coli. In: Gorrod, J.W. (Ed.), Biological Oxidation of Nitrogen. Elsevier/North-Holland Biomedical Press, pp. 375–383.
PAN, 2014. http://www.pesticideinfo.org/.
Pharmapendium, 2015. http://www.pharmapendium.com/.
Powley, M.W., 2015. (Q)SAR assessments of potentially mutagenic impurities: a regulatory perspective on the utility of expert knowledge and data submission. Regul. Toxicol. Pharmacol. 71, 295–300.
Robinson, L., Mcilwaine, I., Copestake, P., Anderson, C., 2000. Comparative evaluation of the performance of online databases in answering toxicology queries. Int. J. Inf. Manag. 20, 79–87.
RTECS, 2015. http://www.cdc.gov/niosh/rtecs/default.html.
Schelhaas, M., Waldmann, H., 1996. Protecting group strategies in organic synthesis. Angew. Chem. Int. Ed. Engl. 35, 2056–2083.
Stavitskaya, L., Aubrecht, J., Kruhlak, N.L., 2015. Chemical structure-based and toxicogenomics models. In: Graziano, M.J., Jacobson-Kram, D. (Eds.), Genotoxicity and Carcinogenicity Testing in Pharmaceuticals. Springer, pp. 13–34.
Stavitskaya, L., Minnier, B.L., Benz, R.D., Kruhlak, N.L., 2013. FDA Center for drug evaluation and research,"Development of improved Salmonella mutagenicity QSAR models using structural fingerprints of known toxicophores". In: Society of Toxicology, 52nd Annual Meeting. Retrieved from. https://www.leadscope.com/media/SOT%202013%20Stavitskaya%20CDER%20poster.pdf.
Sutter, A., Amberg, A., Boyer, S., Brigo, A., Contrera, J.F., Custer, L.L., Dobo, K.L., Gervais, V., Glowienke, S., Gompel, J.V., Greene, N., Muster, W., Nicolette, J., Reddy, M.V., Thybaud, V., Vock, E., White, A.T., Müller, L., 2013. Use of in silico systems and expert knowledge for structure-based assessment of potentially mutagenic impurities. Regul. Toxicol. Pharmacol. 67, 39–52.
Sweet, D.V., Anderson, V.P., Fang, J., 1999. An overview of the registry of toxic effects of chemical substances (RTECS): critical information on chemical hazards. Chem. Health Saf. 6, 12–16.
Tennant, R.W., 1991. The genetic toxicity database of the national toxicology program: evaluation of the relationships between genetic toxicity and carcinogenicity. Environ. Health Perspect. 96, 47–51.
ToxNet, 2015. http://toxnet.nlm.nih.gov/.
VITIC, 2015. http://www.lhasalimited.org/products/vitic-nexus.htm.
Wexler, P., 2001. TOXNET: an evolving web resource for toxicology and environmental health information. Toxicology 157, 3–10.
Young, R.R., 2002. Genetic toxicology: web resources. Toxicology 173, 103–121.
Zeiger, E., Risko, K.J., Margolin, B.H., 1985. Strategies to reduce the cost of mutagenicity screening with the Salmonella assay. Environ. Mutagen. 7, 901–911.