Responses from FDA Speakers

Disclaimer: The findings and conclusions in this document have not been formally disseminated by the FDA and should not be construed to represent any agency determination or policy.

1. Do FDA and EMA share a common view on whether silico analysis of PGIs is sufficient?

   Response: The FDA and EMA have previously issued regulatory recommendations that include (Q)SAR as a component of impurity evaluation. ICH M7 will further harmonize the process.

2. Does FDA provide the service described here (QSAR analyses + expert assessment) on pharma company requests?

   Response: The Computational Toxicology Service is only provided internally to FDA/CDER safety reviewers.

3. What criteria are used by FDA/CDER to identify equivocal predictions?

   Response: FDA/CDER applies definitions of equivocal predictions as recommended by each software vendor, although it is noted that in some cases these definitions were derived in collaboration with FDA/CDER based on a systematic analysis of the model's predictive performance.

4. Does the FDA use ICH M7 for parenteral drugs, or is its use limited to oral drugs?

   Response: The draft ICH M7 guideline indicates that "… approaches are applicable to all routes of administration…".

5. What is another recommended statistically-based prediction program that can be used?

   Response: To date, FDA/CDER has evaluated three statistically-based software programs that are deemed to be sufficiently predictive and interpretable for structural alert identification under ICH M7: CASE Ultra, Leadscope Model Applier, and Sarah Nexus.

6. What are some examples of mitigating factors in the expert analysis of a structural alert?

   Response: Mitigating factors are co-occurring sub-structural features or whole molecular attributes that can diminish the activity of a structural alert. Such factors include those causing steric hindrance or electronic modulation at the alerting reactive center. Specific examples include the alkyl substitution of an epoxide group, or the presence of a carboxylate group ortho to the nitrogen of an aromatic amine.

7. How often are the results from in silico analyses determined to be false positive.

Response: The combined application of an expert rule-based and a statistical-based approach without the use of expert knowledge is intended to provide high sensitivity and negative predictivity but will result in a fairly high false positive rate. However, this rate can be reduced through the use of expert knowledge, and it is a relatively common occurrence at FDA/CDER for raw positive predictions to be overturned resulting in an overall negative expert conclusion.

8. Do any of the in silico programs used by FDA/CDER assess the placement of the biophore and whether it can actually interact with DNA, i.e. steric hindrance?

   Response: Co-occurring features such as bulky substituents that are identified to be associated with a lack of mutagenic potential may be captured as part of an expert rule-based alert or identified statistically by a model building algorithm. In each case, this information may result in the mitigation of a structural alert identified in a test compound.

9. ICH M7 Note 1 states:
   "The ICH M7 guideline recommendations provide a state-of-the-art approach for assessing the potential of impurities to induce point mutations and ensure that such impurities are controlled to safe levels so that below or above the qualification threshold no further qualification for mutagenic potential is required."

   It further states:
   "In cases where the amount of the impurity exceeds 1 mg daily dose for chronic administration, evaluation of genotoxic potential as recommended in ICH Q3 A/B could be considered."

   Question: What if your impurity is over 1 mg, negative in both (Q)SARs (Derek and CASE Ultra) and still under the qualification threshold (QT) based on maximum daily dose. Should one still test according to ICH Q3 A/B, i.e., a study to detect point mutations and one to detect chromosomal aberrations?

   Response: Answer deferred to final ICH M7 guideline.

10. Can reactivity and conversion of an impurity under physiological conditions (e.g. low stomach pH) be considered a factor to support negative prediction?

    Response: Biological relevance (i.e., risk characterization) is an important consideration for decision making but mutagenic potential determined by (Q)SAR (i.e., hazard identification) is a separate issue. Therefore, (Q)SAR predictions should be evaluated independently.

    In one recent case, an Ames positive impurity was considered biologically irrelevant due to degradation in the GI tract. Based on this information, the impurity is considered non-mutagenic when present in oral drugs. The same conclusion would not be applicable to parenteral routes of administration.

11. Can the agency detail their Derek settings?

Response: FDA/CDER uses the current off-the-shelf version of Derek without any custom alerts. For interpretation of Derek confidence terms, see the response to Question 12, below.

12. Derek will make classifications such as 'plausible" as opposed to positive/negative. How does the agency weight this type of classification?

    Response: For bacterial mutagenicity, "Plausible" and higher confidence terms are considered positive, "Equivocal" is considered equivocal, and all other terms are considered negative.

13. Why use 2 (Q)SARs if these are as expensive as an Ames test?

    Response: There are examples of regulatory submissions with a large number of impurities (e.g., ≥ 50). There are also impurities that may be difficult to synthesize or isolate. (Q)SAR provides the high-throughput needed to handle large volumes of impurities and, in some cases, may be the only feasible method for evaluating mutagenic potential. However, the Sponsor has the option of conducting an Ames assay instead of performing a (Q)SAR assessment.

    In addition, the more chemicals analyzed, the lower the cost per chemical by (Q)SAR under an annual license. For institutions with the need for only a few (Q)SAR analyses per year, many (Q)SAR software vendors offer single prediction pricing without the need for an annual software license.

14. What sources do you search to determine if there is known data for a chemical?

    Response: Searching common sources such as PubMed, TOXNET, NTP database, etc. can provide publicly available data. There are also tools available for a cost that allows searching of certain proprietary data sets.

15. When a sponsor submits a compound that has been evaluated in e.g. Derek and Model Applier, do you automatically re-do the analysis if the request is put in internally from the reviewer, or do you review what the sponsor provided first, then see if there are gaps necessitating a re-do? Specifically, do you evaluate the quality of the sponsor submission before you decide to re-do it yourself?

    Response: The process is currently being refined; however, well-performed assessments will not necessarily be repeated by FDA/CDER.

16. Do you know if there will be any language in ICH M7 that defines what an expert is? If not, what does FDA/CDER consider an expert to be?

    Response: The draft ICH M7 guideline does not define an expert. The quality of information submitted to support a regulatory decision is the key consideration, not the credentials of the individual providing the information.

17. Do you limit your interpretation of compound data to genotoxicity or do you take into consideration human exposure via diet or environment?

    Response: Biological relevance can be a component of the risk characterization process.

18. What would be the conclusion if the prediction was positive in one system and negative in another and there were negative non-GLP 2 strain Ames data?

    Response: In some cases, a non-standard Ames assay may be sufficient to refute a positive (Q)SAR prediction. These situations are viewed on a case-by-case basis.

19. When a raw prediction is overruled, is that information provided to software vendors for incorporation into their software.

    Response: If non-proprietary knowledge can be extracted from a prediction without revealing the underlying proprietary information about an impurity's chemical structure, then this information may be shared with software vendors through formal collaborative agreements.

20. FDA/CDER is currently evaluating additional software. Is the recommendation for the number and/or type of programs likely to change based on the ongoing evaluation?

    Response: FDA/CDER is reviewing additional software to assess its suitability and performance for application under ICH M7. At this time, the requirement for two methodologies under ICH M7 is likely to remain unchanged.

21. What is FDA/CDER's position or action plan for impurities when the API is Ames positive?

    Response: These situations are viewed on a case-by-case basis; however, mutagenic impurities are not typically an important consideration if the API itself is mutagenic.

22. What if the impurity and Ames positive API give different alerting structures by (Q)SAR?

    Response: See response for Question 21.

23. What format should be used to present the expert knowledge in a submission?

    Response: Expert knowledge should be fully described in a written narrative including any chemical structures and empirical data used to support the conclusion.

24. Can you elaborate on the slide that described the application of novel approaches? What approaches are those?

    Response: "Novel approaches" refers to in-house models, often constructed by Sponsors using their own data (proprietary or non-proprietary) and/or using their

own statistical algorithms. Many of the same considerations also apply to the less-widely-used commercial and open source tools available in the public domain. FDA/CDER is in the process of performing benchmarking assessments of some of these tools to determine their suitability under ICH M7.

25. Is it possible to remove non-relevant training set examples to run a second prediction?

    Response: All global models by their nature contain non-relevant training set compounds for a given query structure. This is intended to provide a broader applicability domain over which they can provide predictions. While removing selected compounds and rebuilding a model is technically feasible, it results in the generation of a custom model which is then subject to greater scrutiny with respect to its performance characteristics before it can be accepted. A more rational approach is to use the standard, fully-validated model and then provide a justification why a specific prediction should be overruled, if necessary, based on the irrelevance of particular training set structures supporting an alert.

26. Does the Model Applier from Leadscope remove training set examples that have non-relevant alerts?

    Response: See response to Question 27. In cases where an alert is supported by a set of training set structures containing other coincidental structural alerts that are more likely to explain their empirical activity, this information can be conveyed in a written narrative and used as justification to overrule a positive prediction.

27. How would you address scenarios where there is an equivocal prediction (e.g. Derek -ve prediction / Mcase +ve) where the applicant provides an expert view that the MCase prediction is not relevant and hence concludes the compound is overall predicted negative? What happens if the Sponsor and Regulator disagree? What options then exist for the applicant? Can they appeal or discuss?

    Response: It is not uncommon for a positive (Q)SAR prediction to be overturned through consideration of expert knowledge (see Question 7). Should disagreements occur, the Sponsor should engage the appropriate FDA/CDER review division to discuss.

    Of further note, it is important to differentiate between an "equivocal" prediction from a single model and conflicting predictions from multiple models, which are not typically referred to as "equivocal" in this context. In the case of conflicting predictions, the positive prediction carries greater weight to increase overall sensitivity and since the methodologies are complementary: what one model misses, another may pick up.

28. In cases where one system gives a "negative" prediction, a second gives an "out of domain" result, and there is no information for an expert to evaluate or supersede the "out of domain" result, would the overall call be considered negative?

    Response: In this specific case, the (Q)SAR approach would be considered insufficient to characterize mutagenic potential. An initial follow-up might be to obtain a valid

prediction from a third system. If a definitive answer cannot be obtained, the Sponsor could engage the FDA/CDER review division to discuss additional follow-up.

29. We would like to know why three software platforms are run routinely by FDA/CDER while ICH M7 requires only two.

    Response: Due to the broad range of structures processed by the Computational Toxicology Consultation Service, we use three software platforms to increase coverage (i.e. the rate of in-domain predictions) since the models we use are complementary for this parameter.

30. What if in the training set 9 compounds were not relevant but 1 is similar to the test structure. How do you resolve this? This brings up a larger issue of differences in interpretation between the regulator and applicant and how this is resolved at a later stage.

    Response: It depends on whether or not the one similar training set structure is sufficient to raise alarm. Read-across could be used to add perspective to the analysis. Inclusion of a narrative describing the rationale for the conclusions and providing other supporting data will reduce the likelihood of differing views between Regulators and Sponsors. However, if the conclusion is still in dispute, the Sponsor should engage the appropriate review division to discuss.

31. What is LTL adjustment and how does it work?

    Response: The less than lifetime adjustment of acceptable daily intake considers both an impurity's duration of exposure and dose. As duration decreases the dose of a mutagenic impurity can be increased without significantly impacting lifetime risk of developing cancer. See the draft ICH M7 guideline for detailed discussion.

32. Is there a learning space on the web to develop expert knowledge?

    Response: Expert knowledge covers a multitude of considerations and assessments, based predominantly on chemistry and toxicology. Most commercial (Q)SAR software vendors can provide detailed information on how to perform expert analysis of predictions obtained with their own software.

33. Have you compared raw predictions to raw predictions + expert knowledge in terms of their correlation with the results of the Ames test?

    Response: A recently published white paper (Sutter et al., 2013) describes this type of analysis using data sets from several large pharmaceutical companies and concludes that raw predictions plus expert knowledge yields improved negative predictivity over raw predictions alone

    Sutter et al, 2013. Use of in silico systems and expert knowledge for structure-based assessment of potentially mutagenic impurities. Regul. Toxicol. Pharmacol.67:39-52.