Reconciling Conflicting (Q)SAR Predictions in Impurity Evaluations

Mark W. Powley, Ph.D.
Division of Antiviral Products
Office of New Drugs
Center for Drug Evaluation and Research
US FDA
Outline

• Background
• (Q)SAR and ICH M7
• Strategies for Reconciling Discordant Predictions
• Case Studies
• CDER/OND Process for Evaluating Impurities
• Summary
• Questions

Disclaimer: The views expressed in this presentation are those of the speaker and not, necessarily, of the FDA.
Background
Why are we concerned with impurities?

- unlike API, impurities offer no direct benefit to the patient
- impurities will be present regardless of the control strategies applied
- by their nature some impurities are reactive and may possess mutagenic potential
- mutagenicity is tied to the multi-step process of carcinogenicity
  • effects will not be evident in patients for many years
  • defeats the purpose of clinical monitoring
Are we too concerned with impurities?

- lifetime risk of developing cancer in the US is ~1 in 2 for men and ~1 in 3 for women. American Cancer Society (2011) Cancer Facts and Figures
- exposure to mutagens/carcinogens is constant

Radiation

Air pollution

Diet

Less healthy Healthy
• **What does this mean from a practical standpoint?**
  
  – evaluating the mutagenic potential of drug impurities is an important component of safety assessment; however, it is important to consider how much additional risk is posed by small amounts of mutagenic impurities in drugs
  
  – a cautious approach is warranted but conducting an empirical Ames assay for every potential and known impurity is not feasible or justified
  
  – impurity evaluation process must balance the need for high-throughput with the regulatory imperative of maximizing patient safety
• **(Q)SAR**
  – provides the high-throughput process needed to handle a large volume of impurities
  – demonstrated to have adequate sensitivity (~85% depending on systems used, test sets evaluated, etc.) → critical for patient safety
  – overall, (Q)SAR is considered “fit for purpose”
  – recommended by regulatory agencies

**Assessment and Control of DNA Reactive (Mutagenic) Impurities in Pharmaceuticals to Limit Potential Carcinogenic Risk**

M7
(Q)SAR and ICH M7
ICH M7 Hazard Assessment

1. search for empirical data
2. (Q)SAR using expert/rule-based and statistically-based methodologies to predict Ames assay outcome
3. “The outcome of any computer system-based analysis should be reviewed with the use of expert knowledge in order to provide additional supportive evidence on relevance of any positive or negative prediction and to elucidate underlying reasons in case of conflicting results.”
• What are the implications of identifying a potentially mutagenic impurity by (Q)SAR?
  – not trivial
  – empirical testing to characterize mutagenicity
  – control to a “safe” level
    • default TTC: based on 1.5 µg/day or LTL adjustment
    • impurity specific TTC: based on 1/10^5 increased risk
    • PDE: for impurities with demonstrated threshold
    • other: additional cancer risk assessment methodologies as appropriate
• **What are the implications of (Q)SAR error?**
  
  – false positives: may result in empirical testing or controlling an impurity to low levels → acceptable from a safety standpoint
  
  – false negatives: may increase patient risk → less acceptable from a safety standpoint
  
  – evaluation of impurities with structural features not adequately covered by a training set: may increase patient risk → less acceptable from a safety standpoint

**Expert knowledge can be used to improve (Q)SAR performance characteristics.**
• **What is expert knowledge?**
  
  – supporting information provided by scientists with experience/expertise in genetic toxicology, chemistry, and/or computational toxicology
  
  – definition and utility are evolving; however, specific considerations may include:
    
    • relevance of any structural alerts identified
    • alert/prediction confidence
    • impact of mitigating factors
    • chemical coverage
    • empirical data for closely related analogs
• **How is expert knowledge viewed by regulators?**
  
  – from an FDA/CDER/OND pharmacology/toxicology reviewer perspective, expert knowledge can be used to:
    
    1. maximize confidence in a prediction
    2. provide rationale to supersede + or – prediction
    3. provide a basis for assessing mutagenicity in absence of a prediction

  relatively common
  
  rare – very high bar to justify regulatory decision based solely on expert knowledge
• **Additional Thoughts**
  – expert knowledge is:
    • subjective
    • variable in regards to quality
    • often not described in sufficient detail
    • not new → the concept has traditionally been a component of the regulatory decision making process
  does not preclude it’s utility
Additional Thoughts

practical experience

• Dobo et al. (2012) and Sutter et al. (2013) demonstrated that expert knowledge can be used to increase sensitivity and/or negative predictivity of (Q)SAR

• regulatory experience suggests expert knowledge is most often used to increase specificity of (Q)SAR
Strategies for Reconciling Discordant Predictions
• In a perfect world, predictions from multiple (Q)SAR systems would agree.

<table>
<thead>
<tr>
<th>Structure</th>
<th>Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>(Q)SAR System #1</td>
</tr>
<tr>
<td>(Q)SAR System #2</td>
<td>Negative</td>
</tr>
<tr>
<td>Impurity X</td>
<td>Integrated</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Structure</th>
<th>Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>(Q)SAR System #1</td>
</tr>
<tr>
<td>(Q)SAR System #2</td>
<td>Positive</td>
</tr>
<tr>
<td>Impurity X</td>
<td>Integrated</td>
</tr>
</tbody>
</table>

Note - While the predictions agree, expert knowledge could still influence the integrated conclusion.
In the real world, predictions from multiple (Q)SAR systems don’t always agree.

<table>
<thead>
<tr>
<th>Structure</th>
<th>Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>(Q)SAR System #1 Positive</td>
</tr>
<tr>
<td></td>
<td>(Q)SAR System #2 Negative</td>
</tr>
<tr>
<td>Impurity X</td>
<td>Integrated</td>
</tr>
<tr>
<td></td>
<td>?</td>
</tr>
</tbody>
</table>

In each case, expert knowledge may help reconcile the discordant predictions.
• **Reconciling Discordant Predictions**
  – evaluating a positive prediction
    • Are the training set structures used to derive the alert relevant to the impurity?
    • Is the mechanism of mutagenicity associated with the alert applicable for the impurity?
    • Are there public or proprietary structures that are clearly appropriate and sufficient to characterize mutagenic potential?
    • How “strong” is the alert?
    • other
• **Reconciling Discordant Predictions**
  – evaluating a negative prediction
    • If a suspect reactive group is present,
      o was the expected reactivity influenced by an irrelevant mitigating factor?
      o was it adequately represented in the training set?
    • Are there public or proprietary structures that are *clearly* appropriate and sufficient to characterize mutagenic potential?
    • other
• **Reconciling Discordant Predictions**
  – evaluating an out of domain prediction
    • Are there public or proprietary structures that are *clearly* appropriate and sufficient to characterize mutagenic potential?
    • other
• **Reconciling Discordant Predictions**
  – overall
  • If a positive result cannot be superseded through consideration of expert knowledge, the integrated prediction should be positive.
  • Using expert knowledge to supersede an out of domain result comes with a heavy burden.
  • An adequate description of expert knowledge, including any supporting structures, is a critical consideration for regulatory decision making.
Case Studies
**Example – Positive Prediction**

- impurity predicted positive by (Q)SAR
- alert derived from training set molecules with other obvious reactive groups
- the alert was deemed irrelevant based on differences between training set molecules with the alert and the impurity structure
- conclusion → alert irrelevance used to refute the positive prediction
Example – Negative Prediction

• impurity predictive negative by (Q)SAR; however, a reviewer identified potentially reactive group
• review of model training set identified only 2 molecules containing suspect reactive group
• review of published literature suggested neither molecule was adequately tested in the Ames assay
• conclusion → suspect reactive group was not well represented in the model training set and was not appropriately characterized; therefore, based on a lack of coverage, (Q)SAR should not have been used to qualify this impurity
Example – Negative Prediction

• impurity predictive negative by (Q)SAR; however, structure possessed a well-known reactive group
• review of data found that the aromatic nitro group was identified as an alert; however, the molecule was predicted negative based on impact of a mitigating factor
• mitigating factor was present in 29 Ames negative molecules → 0 of the 29 molecules possessed an aromatic nitro group
• conclusion → negative prediction deemed inappropriate due to influence of a questionable mitigating factor
Example – Out of Domain

- (Q)SAR did not provide a prediction for the impurity due to lack of coverage
- Ames negative structural analog was identified
  - the difference between the 2 molecules was limited to the presence of a known reactive group in the analog compared to a –OH group in the same location of the impurity
- conclusion → although the impurity structure was not covered by (Q)SAR, empirical data from closely related structural analog was deemed applicable for the impurity
CDER/OND Process for Evaluating the Mutagenic Potential of Impurities
• **Who is involved?**

  - **Pharmacology/Toxicology Reviewer**
    - evaluate Sponsor (Q)SAR assessment
    - provide/assess expert knowledge
    - submit structures to consultation service as appropriate
    - consults other internal resources as appropriate
    - make final recommendation
• **Who is involved?**
  – Chemistry Reviewer
    • provide/assess expert knowledge
  – Computational Toxicology
    • performs (Q)SAR evaluation
    • provide/assess expert knowledge
    • evaluate novel (Q)SAR approaches
Summary
• (Q)SAR
  – balances high-throughput with patient safety
  – like all assays (Q)SAR is not perfect, but it is an appropriate tool for evaluating the mutagenic potential of impurities (i.e., fit for purpose)

• Expert Knowledge
  – has always been a part of impurity evaluations; however, the concept has recently received increased attention
  – has limitations but is still useful for improving (Q)SAR performance characteristics (e.g., sensitivity and specificity) and identifying cases where (Q)SAR is not an appropriate qualification tool
• **Regulatory Decision Making**
  – decision making integrates input from multiple disciplines
  – extensive expert knowledge is not always necessary
  – complete and transparent reporting of (Q)SAR results as well as expert knowledge is key to improving the regulatory process
    • inadequate description of expert knowledge will require clarification → potential delays