Use of (Q)SAR to Evaluate Potential Genotoxic Impurities

Naomi Kruhlak, PhD
Lead, Chemical Informatics Program
Division of Applied Regulatory Science
Office of Clinical Pharmacology
Office of Translational Sciences
FDA Center for Drug Evaluation and Research

GPhA Fall Technical Conference
November 4, 2015
Disclaimer

The findings and conclusions in this presentation reflect the views of the authors and should not be construed to represent FDA’s views or policies.

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.
Overview

- Why use (Q)SAR for drug impurities?
- Introduction to (Q)SAR modeling
  - Underlying principles
  - Modeling methodologies
  - Structural alerts
- How to use (Q)SAR models for assessing impurities under ICH M7
  - Typical workflow
  - Application of expert knowledge
  - Documentation of analysis
Drug Impurities

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
Striking a Balance

- Evaluating the mutagenic potential of drug impurities is an important component of safety assessment

- From a practical standpoint:
  - 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

- In silico models provide the high-throughput process needed to handle a large volume of impurities
- Demonstrated to have adequate sensitivity for predicting bacterial mutagenicity (~85% depending on systems used, test sets evaluated, etc.)
  - Critical for patient safety
- For impurities:
  - Considered “fit for purpose”
  - Recommended by regulatory agencies
  - State-of-the-art approach for assessing mutagenicity
(Q)SAR Modeling: What is it?

- Identifies correlations between chemical structural features and biological activity
- Uses the results of actual laboratory testing or clinical outcomes
  - General assumption: Similar molecules exhibit similar physicochemical and biological properties
- Make prediction of a compound’s biological activity based on its chemical structure
  - rapidly
  - consistently

QSAR – quantitative – statistically-derived model
SAR – qualitative – expert rule-based model

(Q)SAR
Building a (Q)SAR Model

Chemical Structures (Descriptors) → (Q)SAR Algorithm → Known Activity Data → (Q)SAR Model → Activity Prediction
(Q)SAR Methodologies

- Statistically-derived models
  - E.g., partial least squares regression analysis (PLS), support vector machines (SVM), discriminant analysis, k-nearest neighbors (kNN)
  - Use a classic training set
  - Rapid to build
  - Vary in interpretability

- Expert rule-based models
  - Capture human expert-derived correlations
  - Often supported by mechanistic information, citations
  - Highly interpretable
  - Anonymously capture knowledge from proprietary data
  - Time-consuming to build
Example: Commercial Fragment-based (Q)SAR Tool

Model Construction:

1. Reduce structures to fragments

2. Identify structures primarily associated with active molecules (structural alerts)

3. Identify modulators of activity

Test Chemical Prediction:

1. Reduce structure to fragments

2. Compare fragments to a list of structural alerts and modulators

3. Provide warnings if fragments are unknown

Predicted activity score

Step 10
**Statistically Identified Structural Alerts**

- **Bacterial mutation**

<table>
<thead>
<tr>
<th>Structure</th>
<th>Total Weight</th>
<th>Salmonella Mut</th>
<th>Salmonella Mut,Salmonella Mut,probability</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Structure" /></td>
<td>0.603</td>
<td></td>
<td></td>
</tr>
<tr>
<td><img src="image2" alt="Structure" /></td>
<td>0.537</td>
<td></td>
<td></td>
</tr>
<tr>
<td><img src="image3" alt="Structure" /></td>
<td>0.494</td>
<td></td>
<td></td>
</tr>
<tr>
<td><img src="image4" alt="Structure" /></td>
<td>0.469</td>
<td></td>
<td></td>
</tr>
<tr>
<td><img src="image5" alt="Structure" /></td>
<td>0.449</td>
<td></td>
<td></td>
</tr>
<tr>
<td><img src="image6" alt="Structure" /></td>
<td>0.449</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Statistical algorithm can identify biologically meaningful fragments.
Why use a computer?

Why not simply use visual inspection?

- Highly complex associations can be captured by a model
  - Published alerts are quite general. A model can identify regions within alert space where the alert is less reliable
    - Mitigating features
    - Activity cliffs
  - Can consider the effect of multiple factors simultaneously
  - Can calculate statistics (e.g., positive predictivity) for alerts
- Consistent
- Rapid – screen multiple chemicals against multiple associations
- Inexpensive
Chemical Informatics Program

- An applied regulatory research group that:
  - Creates chemical structure-linked toxicological and clinical effect databases
  - Develops rules for quantifying in vitro, animal and human endpoint data
  - Evaluates data-mining and (Q)SAR software
  - Develops toxicological and clinical effect prediction models through collaborations with software companies

- Computational toxicology consultations that:
  - Provide (Q)SAR evaluations for drugs, metabolites, contaminants, degradants, etc. to FDA/CDER safety reviewers
  - Perform structure-similarity searching for read-across purposes
  - Provide expert interpretation of (Q)SAR data submitted to FDA/CDER
ASSESSMENT AND CONTROL OF DNA REACTIVE (MUTAGENIC) IMPURITIES IN PHARMACEUTICALS TO LIMIT POTENTIAL CARCINOGENIC RISK

Section 6:
“A computational toxicology assessment should be performed using (Q)SAR methodologies that predict the outcome of a bacterial mutagenicity assay (Ref. 6). Two (Q)SAR prediction methodologies that complement each other should be applied. One methodology should be expert rule-based and the second methodology should be statistical-based. (Q)SAR models utilizing these prediction methodologies should follow the general validation principles set forth by the Organisation for Economic Co-operation and Development (OECD).

The absence of structural alerts from two complementary (Q)SAR methodologies (expert rule-based and statistical-based) is sufficient to conclude that the impurity is of no mutagenic concern, and no further testing is recommended (Class 5 in Table 1).”
The ICH M7 (Step 4) Guideline

**Model output** “... can be reviewed with the use of expert knowledge in order to provide additional supportive evidence on relevance of any positive, negative, conflicting or inconclusive prediction and provide a rationale to support the final conclusion.”

For example:

- Understand the reasoning for a prediction
- Consider data from relevant, structurally similar compounds (analogs) not used to construct model
Expert Knowledge

- Although the use of expert knowledge tends to be subjective, its application can enhance the overall accuracy of predictions by providing a rationale to supersede a positive or a negative prediction and maximize confidence in the overall prediction.
- Particularly useful for resolving ambiguous (Q)SAR outcomes (e.g., equivocal, out of domain).
(Q)SAR Software Used by FDA/CDER

- Statistically-Derived Models
  - **CASE Ultra/MC4PC** MultiCASE, Inc.
  - **Model Applier - Statistical Models** Leadscope, Inc.
  - **Sarah Nexus** Lhasa Limited

- Expert Rule-Based Models
  - **Derek Nexus** Lhasa Limited
  - **Model Applier - Expert Alerts** Leadscope, Inc.
  - **CASE Ultra - Expert Alerts** MultiCASE, Inc.

All software above are used by FDA/CDER under Research Collaboration Agreements (RCAs)
(Q)SAR Software Selection Criteria

- Different methodologies can yield different predictions
  - Predictions are complementary
  - Yields higher sensitivity and negative predictivity
  - Second statistical system improves coverage

- Predictions are chemically meaningful and transparent
  - Structural alerts and associated training set structures can be identified to explain why a prediction was made
  - Application of expert knowledge is facilitated

- Software and models are publicly available
  - Our results are reproducible by pharmaceutical sponsors and others
(Q)SAR Software Acceptability

- Under the ICH M7 guideline, sponsors may submit (Q)SAR analyses performed using models that are fit-for-purpose
  - Commercially available
  - Freely available
  - Constructed in-house

- CDER has prior knowledge of several commercial and freely available (Q)SAR software

- For software that CDER has no prior knowledge, supporting documentation demonstrating that a model is fit-for-purpose is desirable
  - 2 models: expert rule-based and statistical-based
  - Predict bacterial (Ames) mutagenicity
  - Consistent with OECD Validation Principles
OECD Validation Principles

- To facilitate the consideration of a (Q)SAR model for regulatory purposes, it should be associated with the following information:
  1) a defined endpoint
  2) an unambiguous algorithm
  3) a defined domain of applicability
  4) appropriate measures of goodness-of-fit, robustness and predictivity
  5) a mechanistic interpretation, if possible

For every (Q)SAR analysis:

1) Check that the impurity structure is correct (e.g., crosscheck with molecular weight and molecular formula)
2) Check for experimental Ames data
3) Generate predictions for the impurity structure
   - Individual model outcomes: positive, negative, equivocal, or out-of-domain
   - Generate an overall conclusion
4) Determine the credibility of the reasoning for the predictions, e.g.,
   - identify alerting portion of the molecule, compare to API
   - assess training set structures used to support a prediction
   - evaluate confidence scores
   - confirm structure is within each model’s domain of applicability
5) Check for experimental data for chemicals with similar structures (analogs)
6) Report overall expert conclusions
Relevant Information for Reporting

- **Materials and methods**
  - Name and version of software and (Q)SAR models used
  - Prediction classification criteria, such as the cutoff or threshold values to define a positive/negative/equivocal result

- **Results and Conclusions**
  - Summary of each prediction, as well as the overall conclusion
  - Confirmation that the impurity is within the model’s domain of applicability
  - Description of any confirmatory application of expert knowledge, including analogs (where appropriate)
  - Rationale for superseding any prediction

- **Appendix**
  - Raw (Q)SAR outputs
  - Ames data for structurally related compounds used to confirm or refute a prediction
Concluding Remarks

- (Q)SAR models provide a high-throughput means to assess mutagenic potential of impurities
  - Models are deemed “fit-for-purpose” under ICH M7
  - At CDER, expert knowledge is routinely applied to (Q)SAR predictions

- Prediction transparency and interpretability are key
  - ICH M7 guideline is not software-specific
  - Choice of software and models impacted by model interpretability
  - Facilitates application of expert knowledge

- Comprehensive reporting reduces the need for follow-up clarification
  - Documentation of software and model names and versions
  - Summary of results and conclusions
  - Additional detail if model predictions are overruled based on expert knowledge
The Chemical Informatics Program Team

Naomi Kruhlak
Lidiya Stavitskaya
Barbara Minnier
Andy Fant
Dongyu Guo
Neil Hartman
Mark Powley
Jae Wook Yoo
(Marlene Kim)

Support:
• Critical Path Initiative
• ORISE
• RCA partners