Regulatory Use of (Q)SAR Models for Assessing the Safety of Known and Potential Impurities

Naomi L. 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

FDA/PQRI Conference on Evolving Product Quality – North Bethesda, MD
September 16, 2014

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

- Safety assessment of drug impurities
- Introduction to (Q)SAR modeling
  - Chemical structures
  - Endpoints
  - Algorithms
- (Q)SAR model application to drug impurities
  - ICH M7 guideline
- FDA/CDER Chemical Informatics Program
  - Computational Toxicology Consultation Service

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
- 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
  - Exposure to mutagens/carcinogens is constant (e.g., in food, environment)

Striking a Balance

- Evaluating the mutagenic potential of drug impurities is an important component of safety assessment
  - But, important to consider how much additional risk is posed by small amounts of mutagenic impurities in drugs
- 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

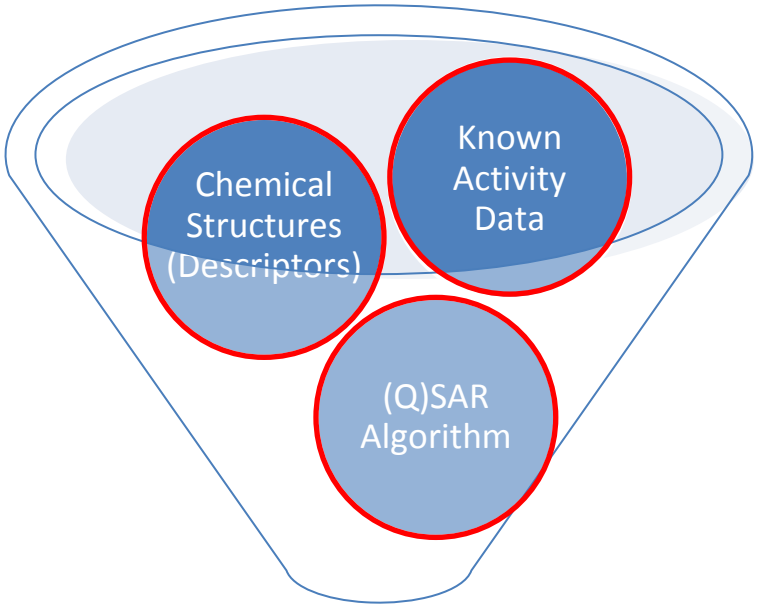
- In silico models provide the high-throughput process needed to handle a large volume of impurities
- Demonstrated to have adequate sensitivity for predicting bacterial mutagenicity (~85% depending on systems used, test sets evaluated, etc.)
  - ✓ Critical for patient safety
- For impurities:
  - Considered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

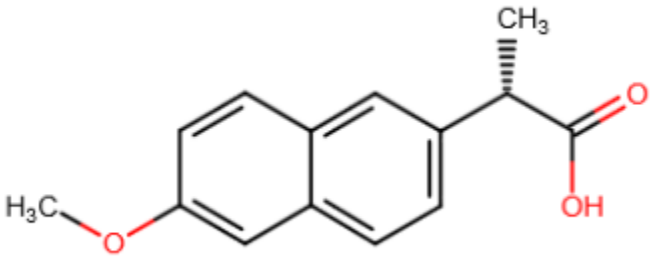
(Q)SAR Modeling: What is it?

- 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



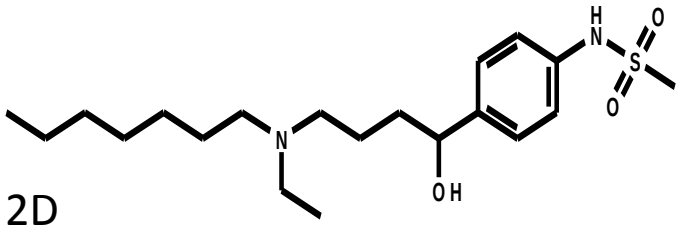


(Q)SAR Model

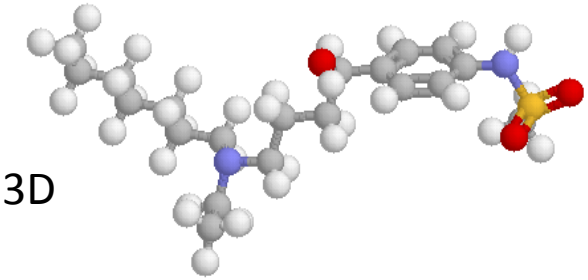
Activity Prediction

Chemical Structures

- Structural representation



2D



3D

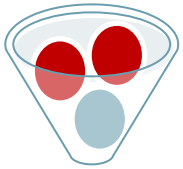
- Molecular descriptors

More Complexity

- Atom counts
- Atom connectivity
- Sub-structural fragments
- Electronic
- Surface, shape
- Quantum mechanical

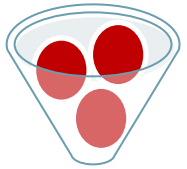
Less Interpretability

Known Activity Data



- Discrete endpoints
  - Binary (dichotomous) activity scores
  - E.g., mutagenicity, carcinogenicity
  - Easier to model
  - Less informative
  - Difficult to characterize some endpoints this way
- Continuous endpoints
  - Range of numerical values
  - E.g., logP, hERG IC_{50}, MTD
  - More informative, if predictive
  - Harder to model
  - More dependent on consistent data sets

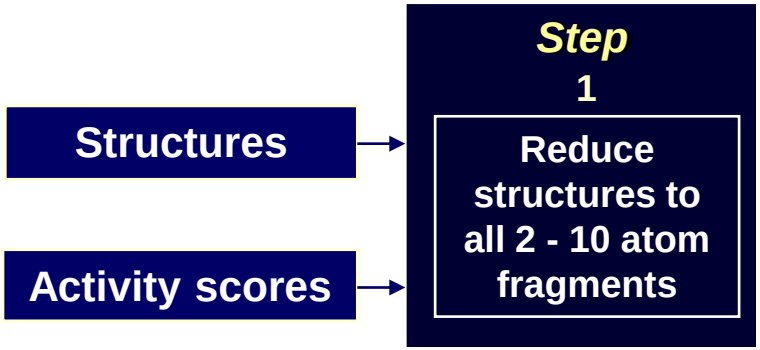
(Q)SAR Algorithms



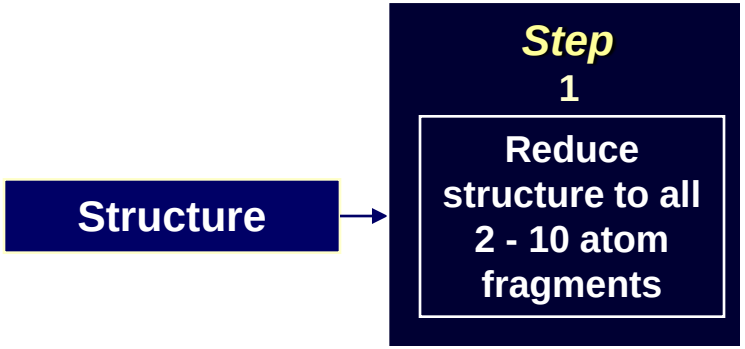
- Statistically-derived models
  - E.g., partial least squares regression analysis (PLS), support vector machines (SVM), discriminant analysis, k-nearest neighbors (kNN)
  - Use a classic training set
  - Rapid to build
  - Vary in interpretability
- Expert rule-based models
  - Capture human expert-derived correlations
  - Often supported by mechanistic information, citations
  - Highly interpretable
  - Anonymously capture knowledge from proprietary data
  - Time-consuming to build

Example: Commercial Fragment-based QSAR Tool

Model Construction:

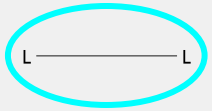
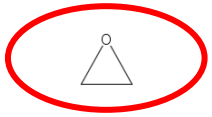
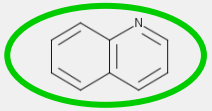
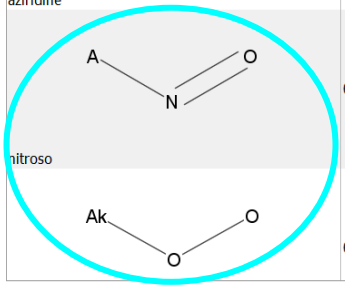


Test Chemical Prediction:



Statistically Identified Structural Alerts

- **Bacterial mutation**

Structure	Total Weight	Salmonella Mut	Salmonella Mut.Salmonella Mut.probability
 GTSA(PA)_24_unsubst het-het [Salmonella]	0.603		
 unhindered epoxide	0.537		
 quinoline	0.494		
 aziridine	0.469		
 nitroso	0.449		
Ak O O	0.449		

Statistical algorithm can identify biologically meaningful fragments

Chemical Informatics Program

- An applied regulatory research group that:
  - Creates 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
- The Computational Toxicology Consultation Service that:
  - Provides (Q)SAR evaluations for drugs, metabolites, contaminants, degradants, etc. to FDA/CDER safety reviewers
  - Performs structure-similarity searching for read-across purposes
  - Provides expert interpretation of (Q)SAR data submitted to FDA/CDER

The ICH M7 (Step 4) Guideline

ASSESSMENT AND CONTR GENIC) IMPURITIES IN PHARMACEUTICALS TO LI C RISK

Section 6:

"A computational toxicolo rmed using (Q)SAR methodologies that predi utagenicity assay (Ref. 6). Two (Q)SAR prediction me each other should be applied. One methodolog nd the second methodology should be st utilizing these prediction methodologies should foll iples set forth by the Organisation for Economi nt (OECD).

The absence of structural y (Q)SAR methodologies (expert rule-based and sta e that the impurity is of no mutagenic concern, and n ed (Class 5 in Table 1)."

INTERNATIONAL CONFERENCE ON HARMONISATION OF TECHNICAL REQUIREMENTS FOR REGISTRATION OF PHARMACEUTICALS FOR HUMAN USE

ICH HARMONISED TRIPARTITE GUIDELINE

ASSESSMENT AND CONTROL OF DNA REACTIVE (MUTAGENIC) IMPURITIES IN PHARMACEUTICALS TO LIMIT POTENTIAL CARCINOGENIC RISK

M7

Current *Step 4* version

dated 23 June 2014

This Guideline has been developed by the appropriate ICH Expert Working Group and has been subject to consultation by the regulatory parties, in accordance with the ICH Process. At Step 4 of the Process the final draft is recommended for adoption to the regulatory bodies of the European Union, Japan and USA.

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."

Definition of (Q)SAR and SAR:

"In the context of this guideline, refers to the relationship between the molecular (sub) structure of a compound and its mutagenic activity using (Quantitative) Structure-Activity Relationships derived from experimental data."

➔ **Structural alert**

(Q)SAR Software Used by FDA/CDER

- Statistically-Derived Models
  - *CASE Ultra/MC4PC* MultiCASE, Inc.
  - *Model Applier* Leadscope, Inc.
- Expert Rule-Based Models
  - *Derek Nexus* Lhasa Limited

All software above are used by FDA/CDER under Research Collaboration Agreements (RCAs)

(Q)SAR Software Selection Criteria

- Software utilize different prediction methodologies
  - Prediction algorithms
  - Chemical structural descriptors
- Predictions are complementary
  - What one software program misses another may pick up
  - Discordant predictions are acceptable
- Predictions are chemically meaningful and transparent
  - Structural alerts and associated training set structures can be identified to explain why a prediction was made
- Software and models are publicly available
  - Our results are reproducible by pharmaceutical sponsors and others

External Validation of 3 Ames Models

Training set n = 3979; External test set n = 3700	Lhasa *Derek Nexus*	Leadscope *Model Applier*	MultiCASE CASE Ultra
Coverage	[98%]	87%	88%
Sensitivity	75%	82%	82%
Specificity	[69%]	68%	58%
Concordance	[73%]	76%	72%
- Predictivity	[67%]	73%	71%
+ Predictivity	77%	78%	72%

Combined performances are based on the rule that a positive overall call is made if any one platform makes a positive prediction

Current Procedure

- For every computational toxicology consultation request:
  - Check that the chemical structures are correct (*e.g.*, crosscheck with molecular weight and molecular formula)
  - Check for experimental data
  - Generate predictions for the requested structures and API, if appropriate, with three software programs for endpoints requested
  - Determine the credibility of the reasoning for the predictions
    - Identify alerting portion of the molecule
    - Evaluate statistics
    - Assess training set structures used to derive an alert
    - Determine whether structure is within each model's domain of applicability
  - Check for experimental data for chemicals with similar structures

Bacterial Mutation Prediction Example

Chem. No.	Chemical Name	*Salmonella* Mutagenicity			Overall Software Prediction	Overall Expert Prediction
		DX	LMA	CU		
1	Chemical 1	-	-	-	-	-
2	Chemical 2	-	-	+	+	-
3	Chemical 3	+	-	NC	+	+
4	Chemical 4	-	NC	-	-	-
5	Chemical 5	+	+	-	+	+
6	Chemical 6	-	NC	NC	NC	NC
7	Chemical 7	-	+	+	+	+

DX = Lhasa Limited Derek Nexus
LMA = Leadscope Model Applier
CU = MultiCASE CASE Ultra
\+ = positive
− = negative
NC = test chemical features are not adequately represented in the model training data set, leading to no call

- Expert interpretation of the model output improves predictive performance

Concluding Remarks

- (Q)SAR models provide a high-throughput means to assess genotoxic potential of impurities
  - Models are deemed "fit for purpose"
  - Supported by new ICH regulatory guideline
  - Results are routinely reviewed with the use of expert knowledge
- Prediction transparency and interpretability are key
  - Identification of structural alerts is important
  - Impacts choice of software and models
  - ICH M7 guideline is not software-specific
- Role of FDA/CDER Computational Toxicology Consultation Service evolving under ICH M7:

Running entire analysis → Filling prediction and/or data gaps → Interpreting sponsor submissions

Acknowledgements

Chemical Informatics Team Members:

Naomi Kruhlak
Lidiya Stavitskaya
Barbara Minnier
Dongyu Guo
Mark Powley
Neil Hartman
Jae Yoo
Kurt Hewes

Support:

Critical Path Initiative
MCM Initiative
ORISE
RCA partners

Contact Information: Naomi.Kruhlak@fda.hhs.gov
Chemical Informatics Program
FDA/CDER/OTS/OCP/DARS