Development of Pharmaceutical QSAR Models for Regulatory Purposes

Kevin P. Cross, Ph.D.
Glenn J. Myatt, Ph.D.
Leadscope, Inc.
1393 Dublin Road, Columbus, OH, USA

R. Daniel Benz, Ph.D.
Division of Drug Safety Research
Office of Testing and Research
Center for Drug Evaluation and Research
U.S. Food and Drug Administration
Silver Spring, MD, USA

Abstract

The draft regulatory guidance that provides for using QSAR models to predict the toxicity of drug impurities as an alternative to *Salmonella* mutagenicity testing represents a breakthrough in the acceptance of QSAR by the US Food and Drug Administration's Center for Drug Evaluation and Research (FDA/CDER). Such use requires validation of QSAR models for predicting the toxicities of drugs and their impurities that may or may not be drug-like. In particular, it requires the external validation of the reliability of QSAR models used for genetic toxicology predictions. This presentation focuses on our experience in developing QSAR models for regulatory assessment of drugs and drug impurities by FDA/CDER, and how this process differs from the development of QSAR models for drug discovery and development by and for the pharmaceutical industry.

Collaboration

This work was conducted under a Cooperative Research and Development Agreement (CRADA) and a Research Collaboration Agreement (RCA) between the US FDA and Leadscope Inc

Disclaimer

The views expressed in this presentation are those of the authors; this talk is not an official US FDA guidance or policy statement.

Outline

- QSAR for drug safety analysis at US FDA/CDER
  - History
  - Current practices and procedures
- QSAR prediction requirements specific to the prediction of impurities (drug-like, non-drug like, new molecular entities).
  - Draft guidance on handling impurity submissions
  - Applying QSAR models for impurity prediction

QSAR for Drug Safety Analysis at US FDA/CDER

- History of QSAR at FDA/CDER
- Current Regulatory status
  - Guidance
  - Cost/benefit analysis
  - Different criteria for acceptance for different types and uses of pharmaceuticals
- Methodology of QSAR application at FDA
  - Impurity modeling for regulatory decision making predictions
  - Consultations for Office of New Drugs
  - Explaining predictions – analogs and feature analysis
  - Transparency in predictions
  - Making chemistry from statistics
  - Consensus support – one data point amongst many

The History of Genetic Toxicology

- 1970s
  - >200 Different Tests; each group decided which to use / request
  - No universally accepted standard, proven protocols
  - Few large sets of tested chemicals at endpoints being developed
  - No harmonized regulatory guidance
- 1980s
  - Protocols standardized and many chemicals tested
  - Battery and tier approaches adopted
- 1980s – 1990s
  - Regulatory guidance appeared (EPA / OECD / FDA / ICH)
  - Battery of 3 tests with standard protocols generally adopted

QSAR Software and Models for Regulatory Purposes Should:

- Make predictions for a defined endpoint
- Take the form of an unambiguous, easily applicable algorithm
- Ideally, have a mechanistic basis
- Have a defined domain of applicability
- Have a measure of goodness-of fit (leave one / many out validation)
- Have its predictive power also assessed by using data not used in the development of the model (external validation)

How Does Regulatory QSAR Differ from Industrial QSAR for Pharmaceuticals?

- Regulatory perspective
  - Guardian of public health
  - Check and balance to industry
  - Risk identification, result confirmation
- Pharmaceutical industry perspective
  - Rejection of early prospects during drug discovery (screening)
  - Risk assessment of prospective lead candidates
  - FDA approval package construction

Software Used at FDA/CDER to Estimate the Mutagenic Potential of Diverse Chemicals

- Predictions using statistical correlations
  - *MC4PC* MultiCASE, Inc.
  - *Model Applier* Leadscope, Inc.
  - *SciQSAR* Scimatics, Inc.
  - *BioEpisteme* Prous Institute for Biomedical Research
- Predictions made with human experts rules
  - *Derek Nexus* Lhasa Limited

There are also other good companies / software programs!

Five Different Software Platforms

- Platforms selected based on two factors:
  - Different (Q)SAR methodologies
  - Predictive performance
- Complementary approaches
  - Positive prediction in one not negated by negative prediction in another
- Use same training data sets (except *Derek Nexus*)
  - Consider different structural features and descriptors/

Comparison of Software Used at FDA/CDER

	MC4PC	*Model Applier*	*SciQSAR (MDL-QSAR)*	*BioEpisteme*	*Derek Nexus*
(Q)SAR Algorithm	**Recursive Partitioning Statistics**	**Partial Logistic Regression / Expert Rules**	**Discriminant Analysis**	**Modified K-Nearest Neighbor**	**Human Expert Rules**
Molecular Structure Interpretation	**2 - 10 Atom Molecular Fragments**	**Fingerprint Molecular Scaffolds & Calculated Properties**	**Connectivity Indices (2D Descriptors)**	**2D Descriptors**	**Structural Alert (Molecular Fragment)**
Molecular Descriptors (2D / 3D)	**Limited 2D (n~6)**	**Limited 2D (n~10)**	**2D (n~200, Kier and Hall)**	**2D (n~126 volume & shape descriptors; 3D in the a future)**	**Limited 2D (n~4)**
Training Data Sets	**FDA / CDER**	**FDA / CDER**	**FDA / CDER**	**FDA / CDER and PIBR**	**Industry, Govern-ment, Literature, and FDA / CDER**
Coverage Measure	**Presence of 2 - 3 Atom Unknown Fragments**	**Presence in Molecular Feature Domain**	**Descriptor-based Membership in Class**	**None (Future: Affinity Constant Functionality)**	**None**
Operating System	**Windows Desktop / Sun Parallel Grid**	**Windows Desktop**	**Windows Desktop**	**Windows Client / Server**	**Windows Desktop**

Use the Results of More than One Computational Toxicology Program

- None of the programs have 100% coverage, sensitivity, and specificity
- All of the programs have distinct approaches for making predictions (the same answer should not be expected!) and are thus complementary and can be used for consensus prediction strategies
- For technical reasons, current FDA/CDER models were biased for specificity; they predict a positive result only if they are “really sure” the chemical is positive
- Use the results from multiple software programs and related models to boost sensitivity: Call the overall result positive if any one of the programs or related endpoints gives a high specificity positive prediction

External Validation of Currently Used (Q)SAR Models for Salmonella Mutation

Training set n = 3575; External test set n = 2571	Lhasa *Derek for Windows*	Leadscope *Model Applier*	MultiCASE *MC4PC*	SciMatics *SciQSAR*	Any one positive ≡ positive
Coverage	undefined	83%	87%	98%	97%
Sensitivity	72%	82%	59%	64%	91%
Specificity	undefined	74%	78%	71%	52%
Concordance	undefined	78%	70%	68%	70%
- Predictivity	undefined	83%	72%	71%	87%
+ Predictivity	71%	74%	67%	65%	62%

Current FDA / CDER Procedure

- For every computational toxicology consultation request:
  - Check that the chemical structures are correct (ChemID+; crosscheck with molecular weight and molecular formula)
  - Run 5 software programs with their *Salmonella* mutagenicity models and obtain the predictions
  - Report that a chemical is positive if any one of the programs give a positive prediction
- When requested initially or as follow-up:
  - Examine the reasoning for the predictions and decide if they are credible
  - Check the testing results for chemicals with similar structures
  - Check predictions for related endpoints

"Require Expert Input to Assess Relevance" - Nigel Greene, April 7, 2011

FDA/CDER Salmonella Mutagenicity QSAR Prediction Sample

Chem No	Chemical Name	*Salmonella* Mutagenicity				Overall *Salmonella* Call
		DfW	LMA	MC	SQ	
1	Chemical 1	NSA	-	-	-	-
2	Chemical 2	NSA	-	-	-	-
3	Chemical 3	+	-	NC	-	+
4	Chemical 4	NSA	-	-	-	-
5	Chemical 5	+	+	-	-	+
6	Chemical 6	NSA	NC	NC	-	NC
7	Chemical 7	NSA	+	+	+	+

DfW = Lhasa Limited *Derek for Windows*
LMA = Leadscope *Model Applier*
MC = MultiCASE *MC4PC*
SQ = SciMatics *SciQSAR*
\+ = positive
− = negative
Eqv = equivocal
NSA = no structural alerts are identified by DfW
NC = test chemical features are not adequately represented in the model training data set, leading to no call
N/A = no available model
A = active/positive in actual laboratory experiment(s)
M = marginal/equivocal in actual laboratory experiment(s)
I = inactive/negative in actual laboratory experiment(s)

U.S. FDA Models Available for QSAR Prediction

- Non-clinical effect models
  - 6 Carcinogenicity
  - 4+9 Genetic toxicity
  - 11 Reproductive, Developmental, and Behavioral toxicity
  - 1 Phospholipidosis
  - 8 Quantitative MTD
- Clinical effect models
  - 6 Renal / Bladder
  - 5 Hepatobiliary
  - 13 Cardiological
  - 22 Pulmonary (under development)
  - 19 Immunological (under development
  - 2 Quantitative MRDD

Content of Computational QSAR Model for Salmonella Assay

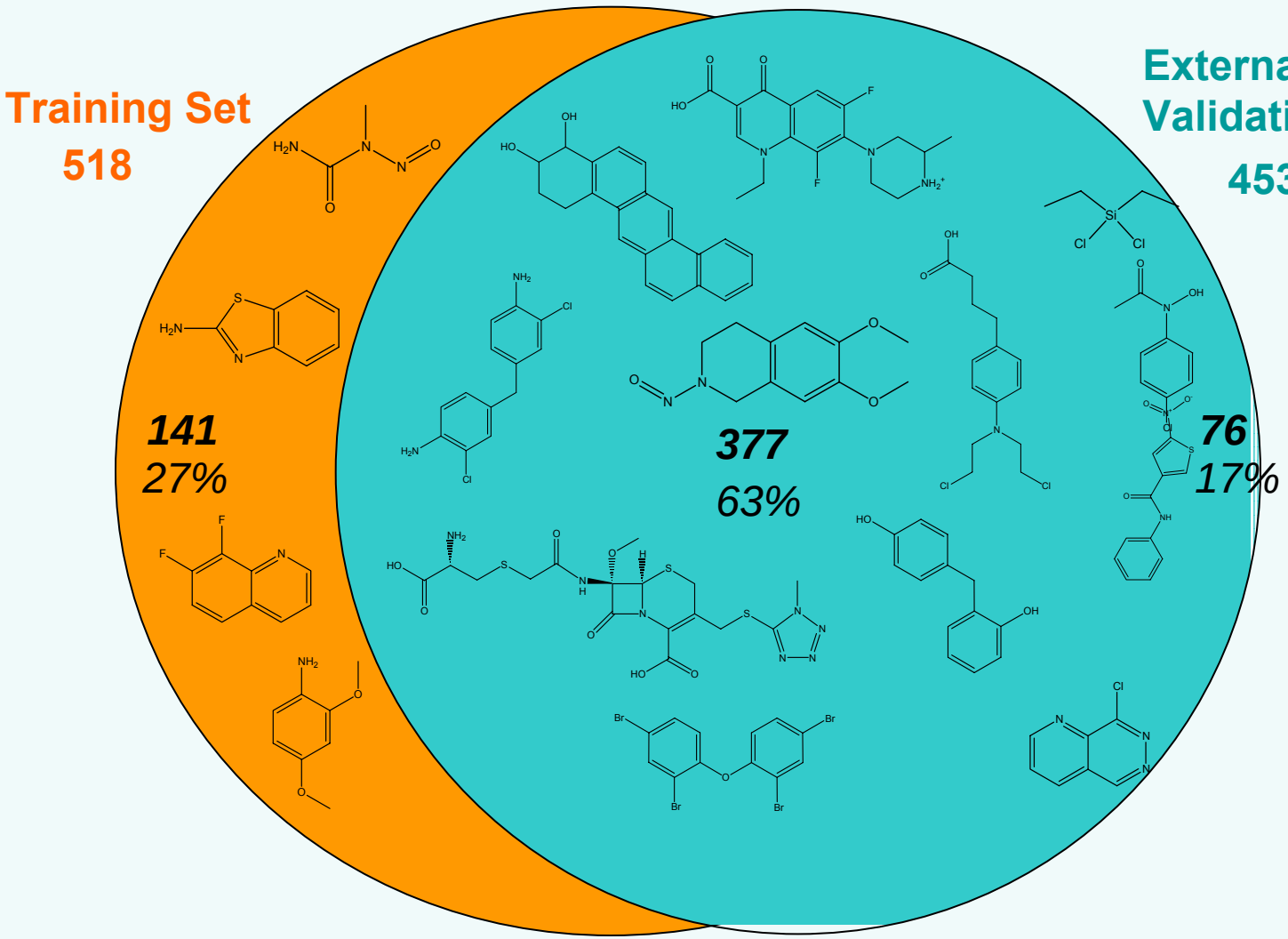
- 3575 molecular structures (Public)
- 1591 Ames positive (44%), 1984 Ames negative (56%)
- 94% calculated with drug-like properties
- Study records with *Salmonella t.*± S9, TA100, TA1535, TA1537, TA98, TA97, TA1538, TA1536
- Data sources: Drugs@FDA, CFSAN PAFA database, NIH/NLM Genetox database, NIEH Genetox summary reports, EPA/OPP, and public structural alerts from Leadscope; *Reg. Toxicol. Pharmacol 2005. 43:313-323*
- Computational tools: Leadscope *Model Applier* and *Enterprise* software
- Generates 422 clusters based on structural fingerprints

External Validation of the New Salmonella QSAR Model

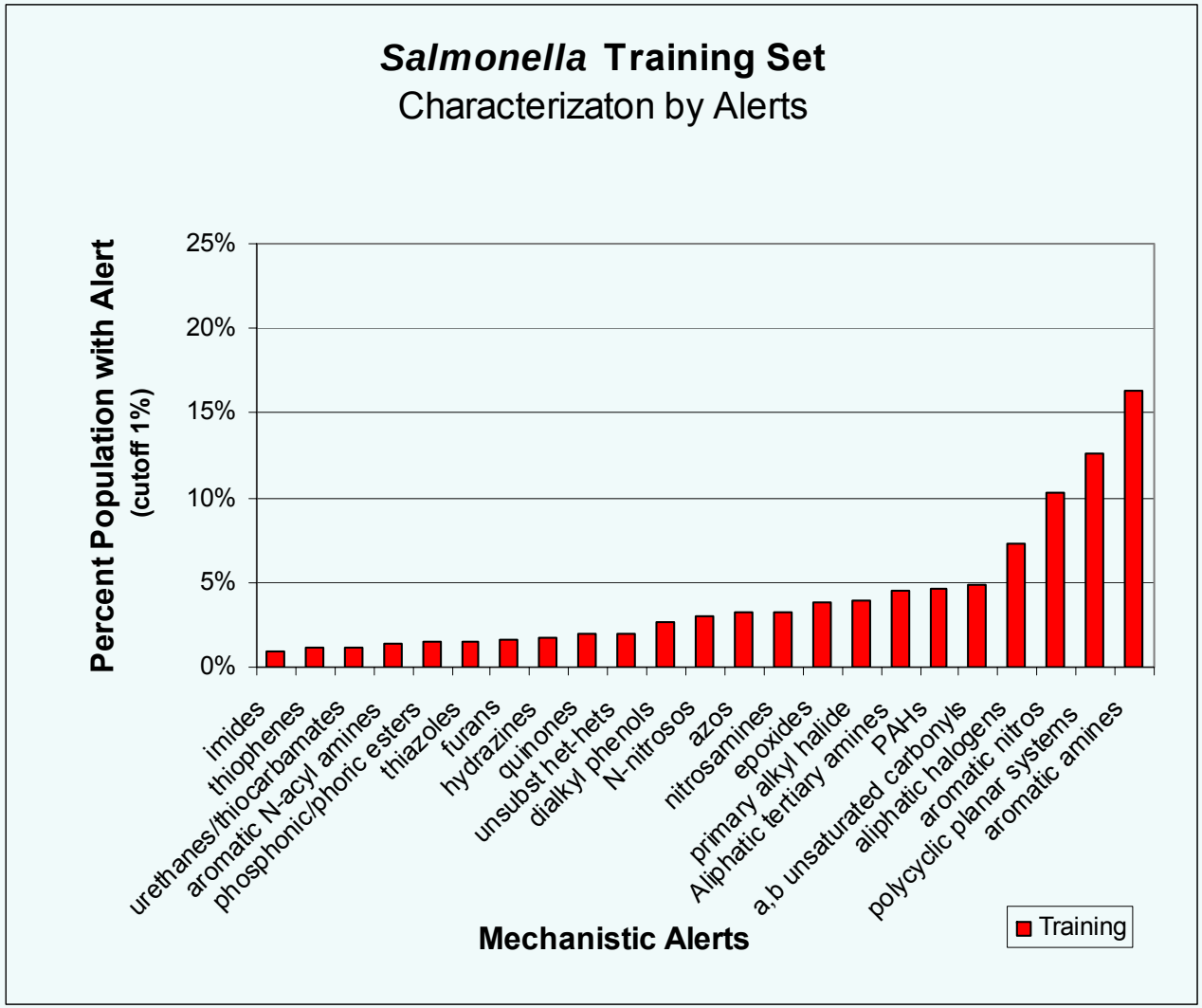
- 2572 Chemicals foreign to the model
- Structure clustering of the validation set
  - 453 structure clusters
  - 17% not shared with model
  - 83% shared with model

Intersection of Clustered Compound Spaces

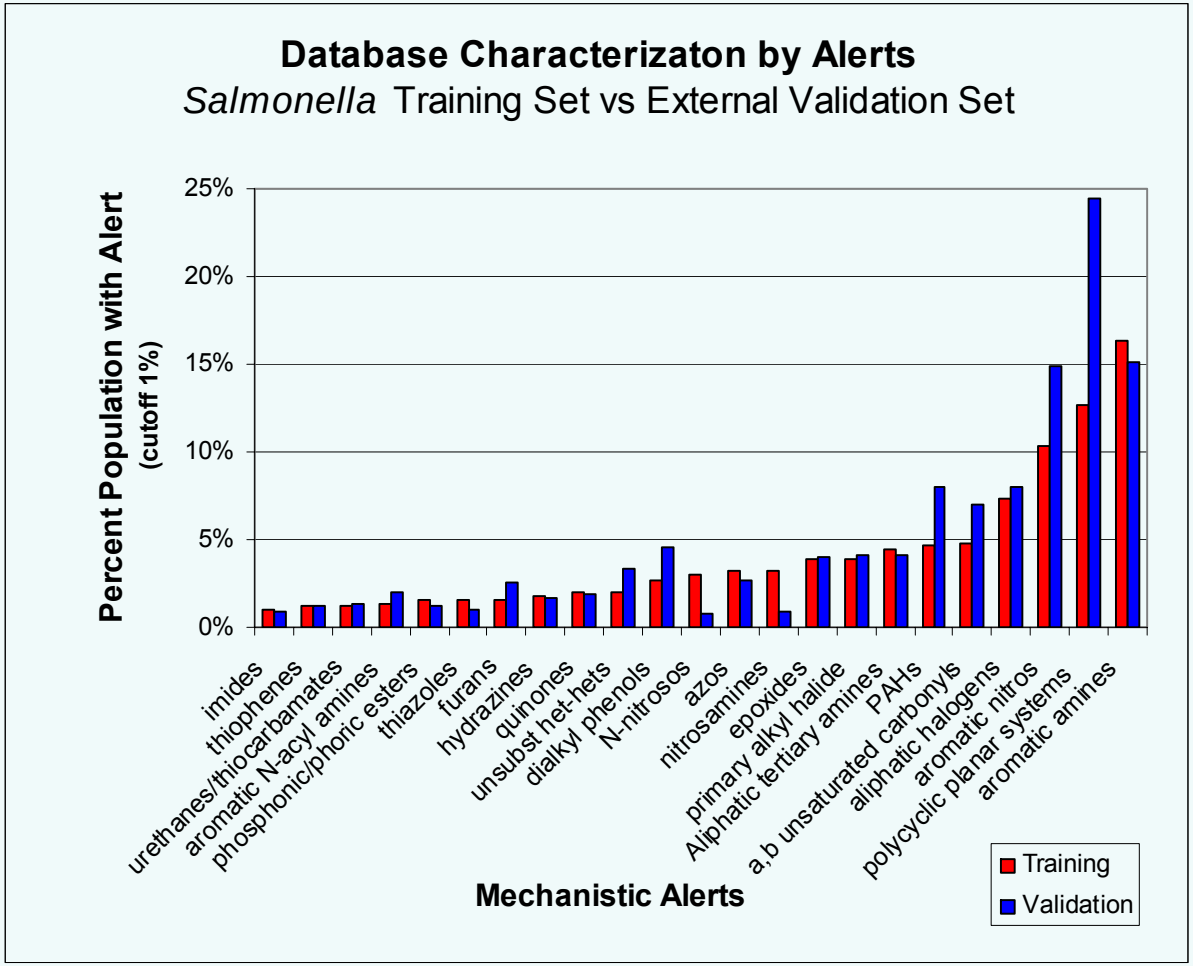
594 clusters total



Toxicity Fingerprint of the Salmonella Training Set



Toxicity Fingerprint of the External Validation Set vs. the Salmonella Training Set



Model Validation Results

Rigorous: Based on 2572 chemicals foreign to the model

External cross validation results

81% Sensitivity
73% Specificity
77% Concordance
89% Coverage

Internal cross validation results

77% Sensitivity
88% Specificity
83% Concordance
89% Coverage

Sensitivity = known positives that are correctly predicted = TP / (TP+FN)
Specificity = known negatives that are correctly predicted = TN / (TN+FP)
Concordance = correct predictions for known positives and negatives
Coverage = percent of test set that is in the applicability domain of the model

Can the New Salmonella QSAR Model Cover the Chemical Space of Known Drug Impurities?

- ICH Q3A Impurities in New Drug Substances classification of impurities
  - Organic
  - Inorganic
  - Residual solvents

Can the New Salmonella QSAR Model Cover the Chemical Space of Known Drug Impurities?

- Organic impurities
  - Starting materials
  - By-products
  - Intermediates
  - Degradation products
  - Reagents, ligands, and catalysts

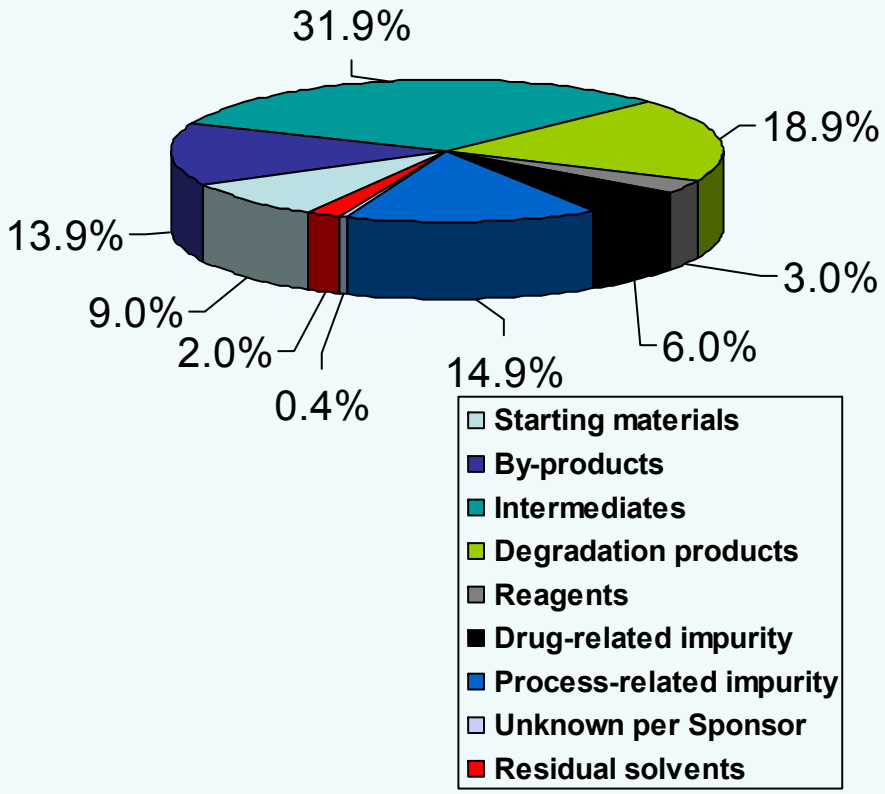
*Data Mining**

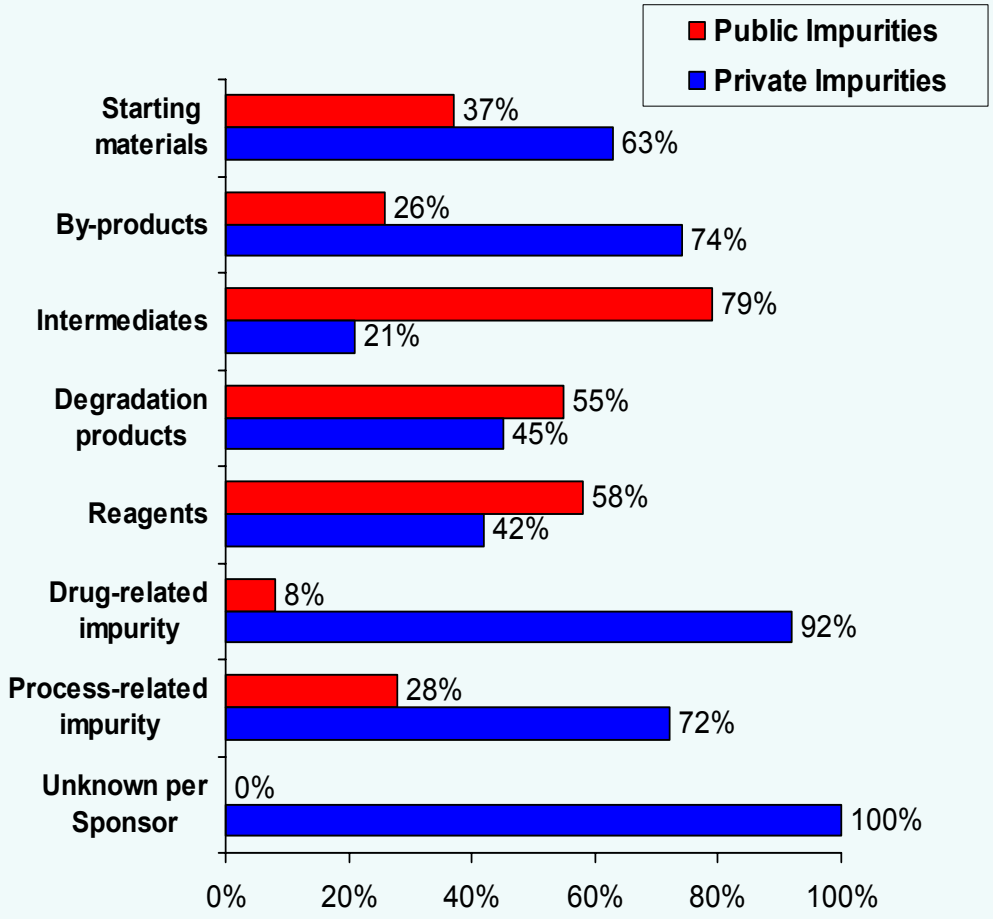
- FDA/CDER INDs and NDAs plus public records for modelable drug impurities
- Only known drug impurities with structure identified were included; random
- Impurities present in drug products, cutting across any therapeutic area and stage of drug development
- Data were transformed to enable *in silico* analysis

*Drug Impurities Database and Analysis from:
Valerio, Luis; Cross, Kevin; SOT 2011

Drug Impurities Database

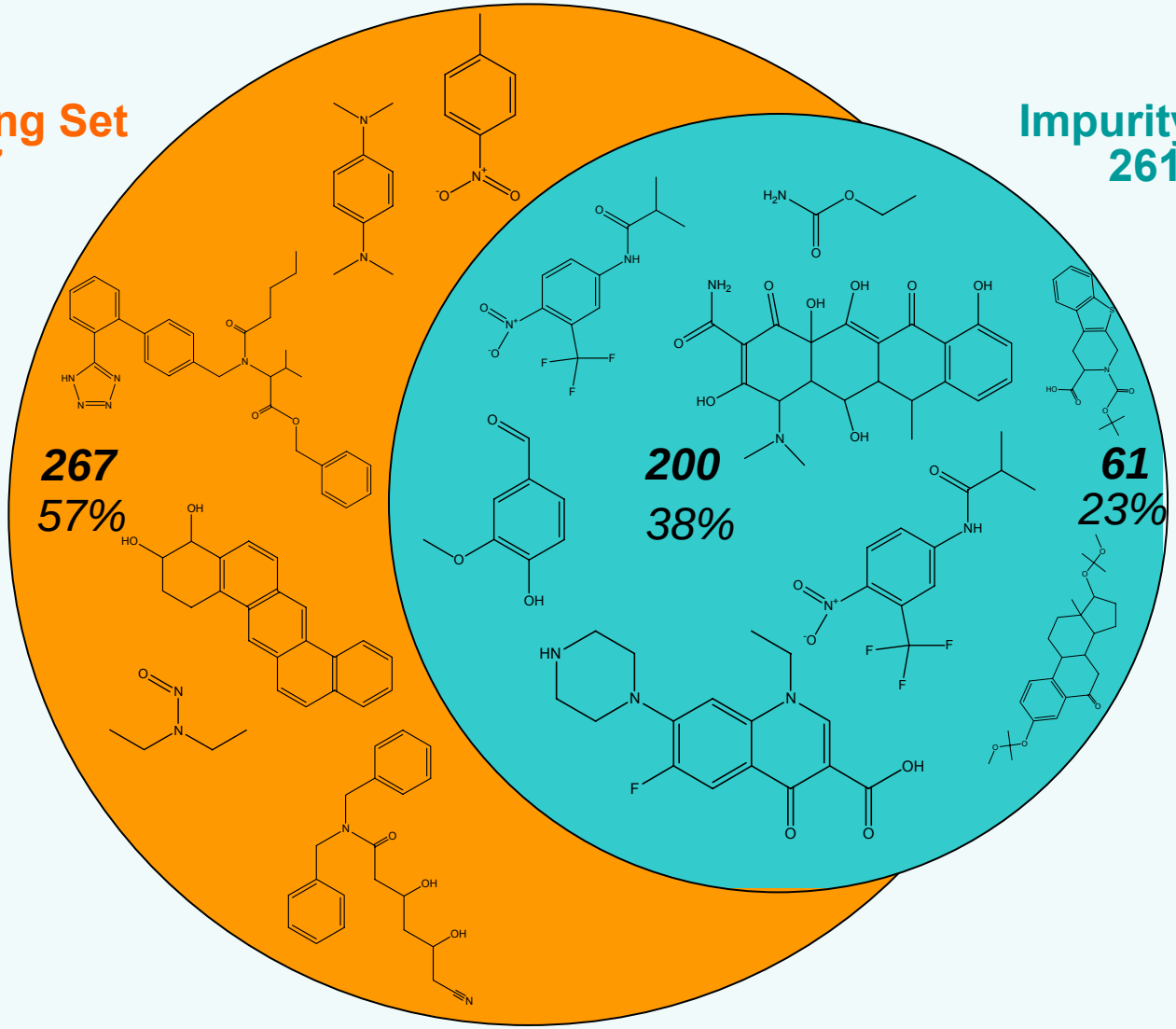
Total 1094 molecules

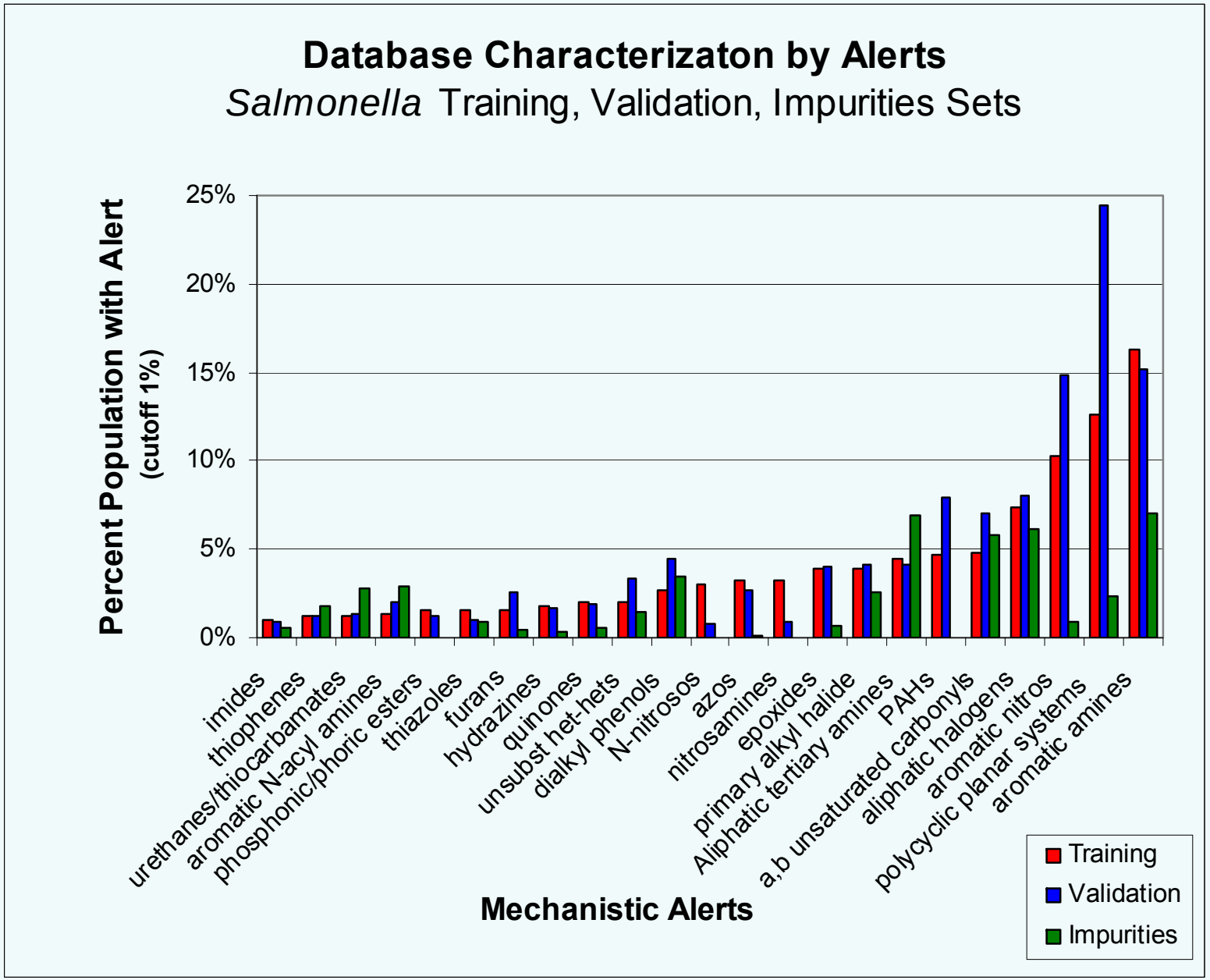




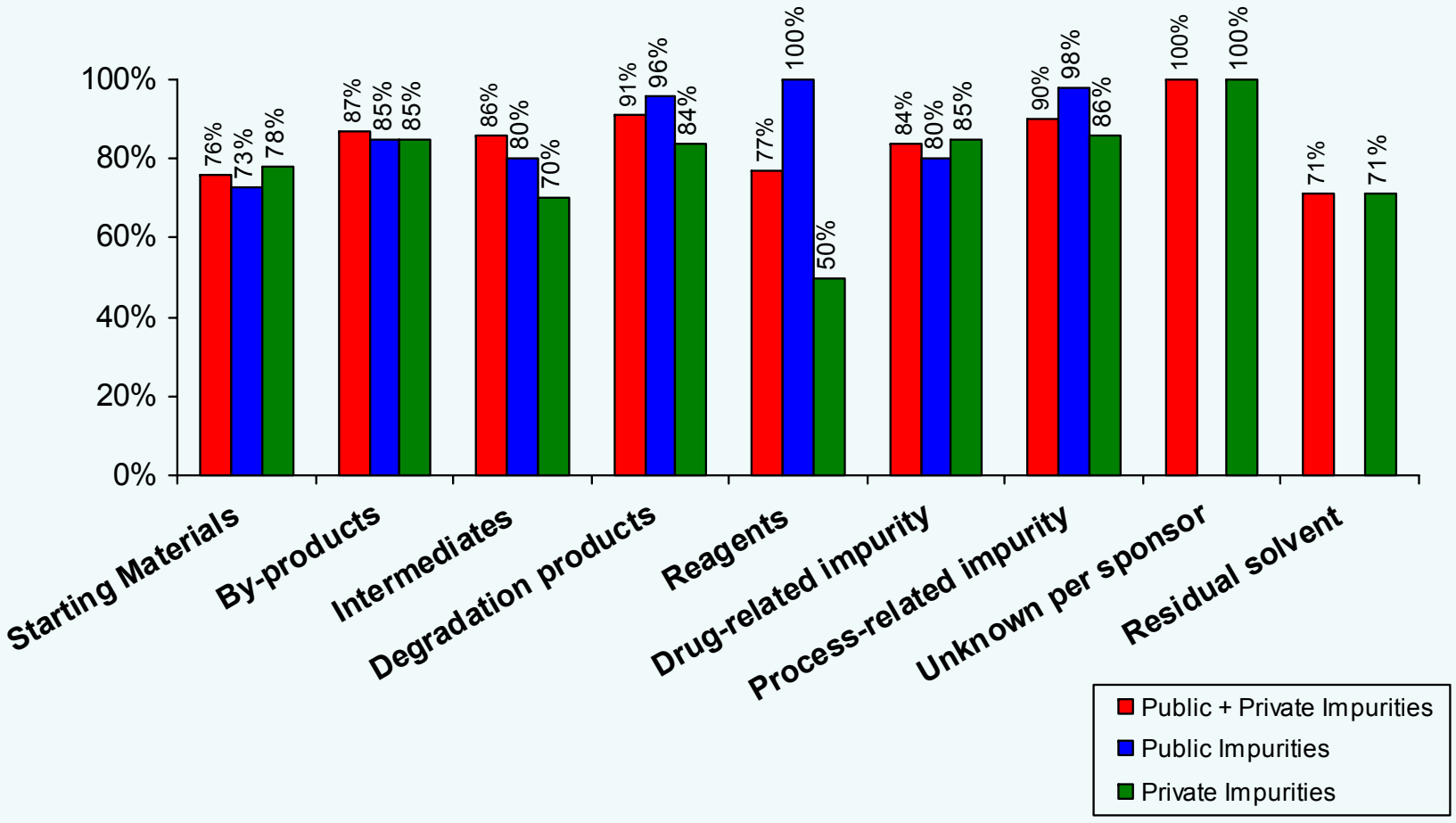
Intersection of Clustered Compound Spaces

528 clusters total

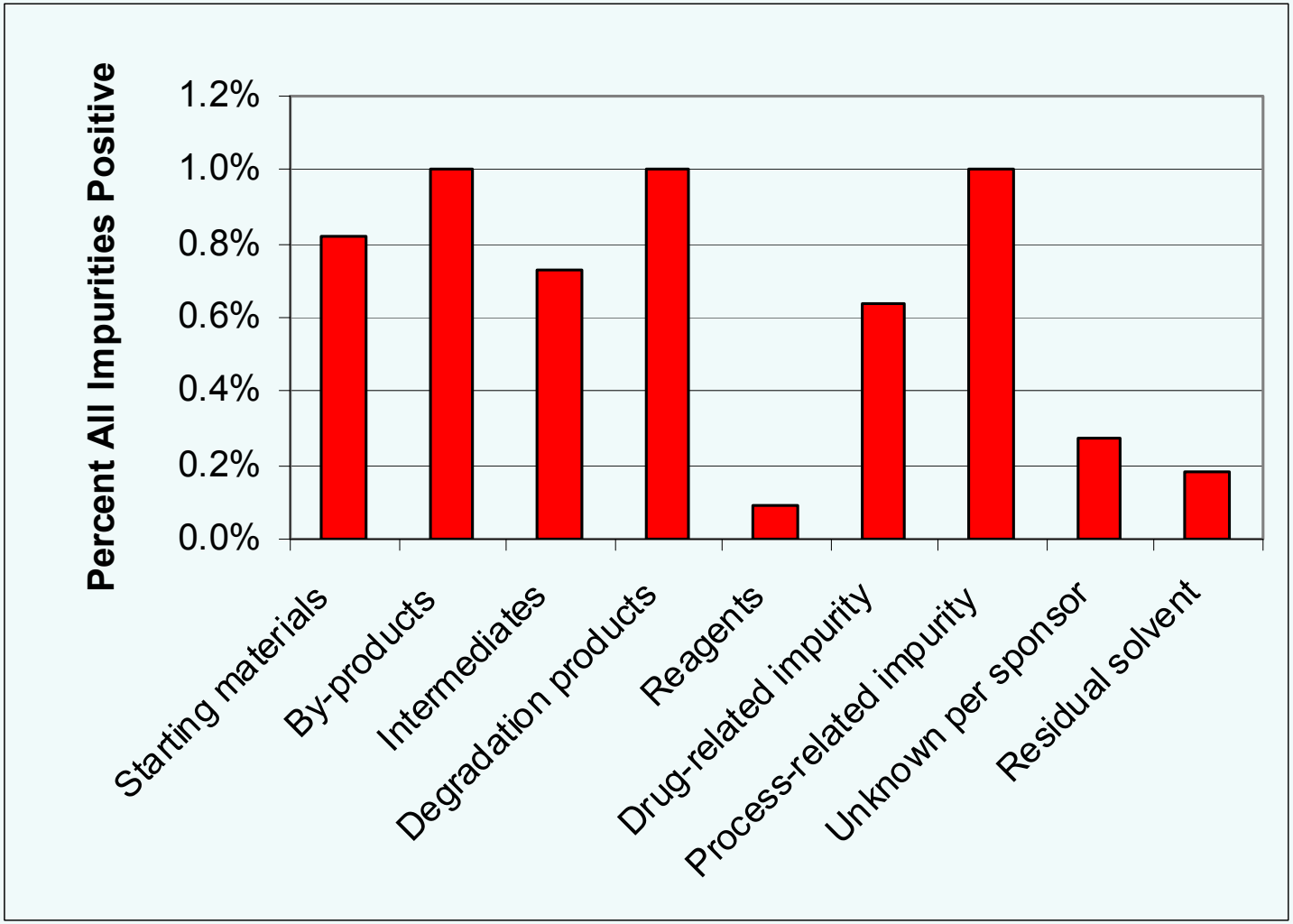



Database Characterizaton by Alerts
Salmonella Training, Validation, Impurities Sets
Percent Population with Alert
(cutoff 1%)
25%
20%
15%
10%
5%
0%
imides
thiophenes
urethanes/thiocarbamates
aromatic N-acyl amines
phosphonic/phoric esters
thiazoles
furans
hydrazines
quinones
unsubst het-hets
dialkyl phenols
N-nitrosos
azos
nitrosamines
epoxides
primary alkyl halide
Aliphatic tertiary amines
PAHs
a,b unsaturated carbonyls
aliphatic halogens
aromatic nitros
polycyclic planar systems
aromatic amines
Mechanistic Alerts
Training
Validation
Impurities

Origin of the Impurities Covered by the Salmonella Training Set



QSAR Positive Predictions of Set of 1094 Impurities



Conclusions

- QSAR modeling for regulatory purposes has different constraints than other applications
  - Pharmaceutical compounds including impurities
  - Large domain required for new molecular entities
  - Models built to confirm submission findings
- External validation by known toxicity features and mechanisms is important
- Models built for pharmaceuticals also work for assessing impurities

Acknowledgements

- Dr. Luis Valerio – Office of Pharmaceutical Science/Immediate Office, CDER, US FDA
- Dr. Sarvin Moghaddam – Office of Pharmaceutical Science/Immediate Office, CDER, US FDA

Thank You!