



Quantitative Structure Activity Tools Used by the U.S. FDA

**Practical Decision-Making Tools for
Identifying Safer Alternatives
OEHHA-COEH Workshop October 1 - 2, 2007**

**Edwin J. Matthews, Ph.D.
Computational Toxicology Expert**

**U.S. Food and Drug Administration
Center for Drug Evaluation and Research
Office of Pharmaceutical Science
Informatics and Computational Safety Analysis Staff**

The views expressed in this presentation do not necessarily represent the official views of the FDA or the United States.

The US FDA/CDER Informatics and Computational Safety Analysis Staff (ICSAS)

Joseph F. Contrera, Ph.D.

Director

Edwin J. Matthews, Ph.D.

(Q)SAR expert

R. Daniel Benz, Ph.D.

Database manager

Naomi L. Kruhlak, Ph.D.

Chemist

Barbara L. Minnier

Quality assurance

Anna A. Frid, B.S.

(M.D.)

Outline

- **FDA decision support tools**
- **Strategy for predicting carcinogenicity**
- **Preclinical QSARs using animal data**
- **QSARs based upon human data**

**(Q)SAR Tools Used by
the FDA to Provide Fast
and Reliable Decision
Support Information**

The US FDA/CDER Informatics and Computational Safety Analysis Staff (ICSAS)

- **An applied regulatory research unit**
- **Create toxicological and clinical databases**
- **Develop rules for quantifying toxicological and clinical endpoints**
- **Evaluate predictive data mining and (Q)SAR software**
- **Develop toxicological and clinical effect prediction programs through collaborations with software companies**

Types of Decision Support Tools

Human / Mammalian Health Effects

- FDA decision support toolbox
- Global QSARs, expert systems
- Carcinogenicity, genetox, reprotox, ...
- Human organ specific adverse effects

Environmental / Non-mammalian Effects

- EPA suite of local QSARs
- OECD QSAR toolbox
- Environmental fate, aquatic toxicity, ...

Unique Features of FDA's Decision Support Tools

- All QSAR models are constructed by the FDA
- Models contain knowledge from proprietary studies in a form that can be shared
- QSAR prediction paradigms are optimized and improved to meet FDA specifications
- The same training data sets are used in multiple QSAR prediction platforms
- CRADA contributions are used to extract archive data & update QSAR models (annually)
- Professional expertise is leveraged through research collaborations with software partners

Decision Support Software

Molecular Fragment QSARs

- MC4PC MultiCASE, Inc.
- Predictive Data Miner Leadscope, Inc.

Molecular Descriptor QSARs

- MDL-QSAR Symyx - MDL, Inc.
- BioEpisteme Prous Science

Human Expert Systems, MOA

- DerekfW / Meteor Lhasa, Ltd.
- Oncologic US EPA

Critical Documentation for Tools

Molecular Fragment Global QSARs

- MultiCASE MC4PC E. Matthews; G Klopman
- LeadScope PDM C. Yang

Molecular Descriptor Global QSARs

- MDL-QSAR J. Contrera
- Prous BioEpisteme E. Matthews

Human Expert Systems, MOA

- Lhasa DerekfW C. Marchant
- US EPA Oncologic Y. Woo

Requirements for QSAR Models

Statistical Validation Criteria

- Leave-many-out (LMO): model reliability
- Leave-one-out (LOO): model stability
- External validation: balanced test sets
- Complementarity: multiple QSAR paradigms

Performance Acceptance Criteria

- High specificity: >80%
- High coverage / high applicability domain
- Multiple QSARs: identical training data sets
- Standardized weight of evidence (WOE) scoring paradigm

Decision Support Strategies

Predictions Through Global QSARs

- Utilize two or more programs
- High confidence and specificity: use consensus positives
- High performance and sensitivity: use all positives

Knowledge Through Expert Systems

- Significant structural alerts
- Reasonable mode of action
- References from the literature

**Specific Example:
FDA's Strategy for
Predicting Carcinogenic
Potential of Chemicals
in Rodents**

FDA's Rodent Carcinogenicity Database and (Q)SARs

- **24,708 study records**
- **Multiple Sources**
 - **NTP/NCI/NIEHS Technical Reports**
 - **FDA/CDER studies submitted by Pharma**
 - **Lois Gold Carcinogenicity Potency Database**
 - **IARC monographs**
 - **Literature**

ICSAS QSAR Weight of Evidence Criteria

<u>Toxic Findings</u>	<u>Potency (log Activity Units)</u>
➤ Highly Potent Toxins • <i>trans</i> -species, multiple site tumors in rodents	50 - 80
➤ Moderately Potent Toxins • <i>trans</i> -gender, single site tumors	30 - 49
➤ Marginal Findings • equivocal, weak, inconsistent findings	20 - 29
➤ Non-toxic	10 - 19

Method based upon Tennant, *Mutation Research* (1993) 286:111-118.
“Compounds that induce *trans*-species tumors present the highest degree of risk because they adversely alter mechanisms that are conserved across species.”

FDA's Rodent Carcinogenicity Database and (Q)SARs

- **1,572 QSAR chemicals**
- **7 Models, 4 QSAR Programs**
 - Male Mouse, Female Mouse, Mouse Composite
 - Male Rat, Female Rat, Rat composite
 - Rodent composite
- **2 Expert System Programs**

Prediction Of Carcinogenicity Using All Positives From 1 or 2 QSAR Programs

<u>Statistic</u>	<u>any1</u>	<u>any2</u>
Specificity	84.5	79.5
Sensitivity	46.2	62.9
ROC (Se/FP)	3.17	3.14
Chi-square	157	266
Coverage	96.1	100

Prediction of Carcinogenicity Using Consensus Positives from 1 to 4 QSAR Programs

<u>Statistic</u>	<u>any1</u>	<u>any2</u>	<u>any 3</u>	<u>all 4</u>
Specificity	52.5	84.1	95.8	99.2
Sensitivity	81.3	58.5	36.5	15.8
ROC (Se/FP)	1.71	3.83	8.67	21.9
Chi-square	169	278	235	111

Correlation of Consensus Positive QSAR Predictions of Carcinogenicity, Derek Mode of Action, and Genetox Data

MOA: Thiouracil Analogue

Genetox	Non-genotoxic
Carcinogenicity	3/3 chemicals carcinogenic
QSAR Programs	3/3 chemicals predicted +

MOA: Azirine / Aziridine

Genetox	Genotoxic
Carcinogenicity	14/14 chemicals carcinogenic
QSAR Programs	14/14 chemicals predicted +

Correlation of Consensus Positive QSAR Predictions of Carcinogenicity, Derek Mode of Action, and Genetox Data

Derek Structure Alert & Mode of Action	Chemicals Tested	Genetic Toxicity Call	Carcinogenicity			(Q)SAR Program Predictions								
			Call	No.	%	Carcinogens			Noncarcinogens					
						2	1	0	2	1	0			
030, Thiouracil analogue	3	NG	C	3	100.0	3								
075, Azirine or aziridine	14	G	C	14	100.0	14								
078, Allylbenzene derivative	3	NG	C	3	100.0	3								
086, Pyrrolidine ester	6	G	C	6	100.0	6								
101, Nitrogen or sulphur mustard	23	G	C	23	100.0	23								
366, Estradiol analogue	11	NG	C	11	100.0	11								
500, N-Polyhaloalkylthio	3	G	C	3	100.0	3								
070, N-Nitro or N-nitroso	97	G	C	89	91.8	89						8		
032, Thiourea	6	NG	C/NC	5	83.3	4			1		1			
074, Mono- or di-alkylhydrazine	53	G/NG	C/NC	43	81.1	42			1		9	1		
079, Glycidyl ether ...	5	G	C/NC	4	80.0	4					1			
512, Retinoid analogue	5	NG	C/NC	4	80.0	3	1				1			
123, Halogenated alkene	10	G/NG	C/NC	7	70.0	7					3			
076, beta-Lactone	3	G	C/NC	2	66.6	2					1			
<u>072, Epoxide</u>	<u>17</u>	<u>G/NG</u>	<u>C/NC</u>	<u>9</u>	<u>52.9</u>	<u>9</u>					<u>6</u>			<u>2</u>
	259			226		223	1	2			30	1		2

QSAR Toolbox For Animal and *In Vitro* Toxicological Endpoints

Genetic Toxicology

- 27,498 study records
- 5,880 QSAR chemicals
- 3 QSAR Programs
- 21 Models:

**Coming Soon: QSARs
also based upon
EPA PMN Data**

- Microbial composite
- *Salmonella* composite
- *E. coli* composite
- *E. coli* WP strains
- Fungal composite
- Yeast *Saccharomyces*
- *Drosophila* composite
- *Drosophila* sex-linked
- *Drosophila* heritable translocation
- Mammalian *in vivo* mutation
- CHO & V79 *in vitro* mutation
- Dominant lethal
- Micronucleus *in vivo* composite
- Micronucleus *in vivo* mouse
- Micronucleus *in vivo* undefined species
- Chromosome aberrations *in vivo* composite
- Chromosome aberrations *in vivo* mouse
- Unscheduled DNA synthesis composite
- UDS rat hepatocytes
- UDS human fibroblasts
- Unscheduled DNA synthesis other cells

Reproductive and Developmental Toxicology

- 51,724 study records
- 2,115 QSAR chemicals
- 3 QSAR Programs
- 21 Models:
 - Reproductive toxicity in adult male ▪ rodent ▪ rat ▪ mouse
 - Reproductive toxicity in adult female ▪ rodent* ▪ rat ▪ mouse
 - Sperm toxicity in ▪ rodent ▪ rat ▪ mouse
 - Fetal dysmorphogenesis in ▪ rodent ▪ rat ▪ mouse ▪ rabbit
 - Behavioral toxicity in ▪ rodent ▪ rat ▪ mouse

Maximum Tolerated Dose: (Lifetime exposure)

- 3,925 study records
- 1,266 QSAR chemicals
- 1 QSAR Program (MC4PC)
- 8 Models:
 - Male rat ▪ low / high toxicity
 - Female rat ▪ low / high toxicity
 - Male mouse ▪ low / high toxicity
 - Female mouse ▪ low / high toxicity

Acute Toxicity

- 1,668 study records
- 1,273 QSAR Chemicals
- 2 QSAR Programs (MC4PC & MDL-QSAR)
- 4 QSAR Models
 - Rat
 - low / high toxicity
 - Mouse
 - low / high toxicity

QSAR Toolbox for Predicting Adverse Effects of Chemicals in Humans Using Human Data

Maximum Recommended Daily Dose and No Effect Level in Humans

- 1,309 study records
- 1,309 QSAR chemicals
- 2 QSAR Programs (MC4PC & MDL-QSAR)
- 2 Models:
 - Low toxicity
 - High toxicity

Hepatobiliary Effects In Humans

- **120,419 Study Records**
- **1,660 QSAR Chemicals**
- **4 QSAR Programs**
- **5 QSAR Models:**
 - **Bile Duct Disorders**
 - **Cholestasis and Jaundice**
 - **Cytotoxic Injury**
 - **Gall Bladder Disorders**
 - **Liver Enzyme Disorders**

Urinary Tract Effects in Humans

- **214,563 study records**
- **1,660 QSAR chemicals**
- **4 QSAR Programs**
- **6 Models:**
 - **Acute Renal Disorders**
 - **Bladder Disorders**
 - **Blood in Urine**
 - **Kidney Function Tests**
 - **Nephropathies**
 - **Urolithiases**

Cardiological Effects in Humans

- **396,985 study records**
- **1,660 QSAR chemicals**
- **4 QSAR Programs (when complete)**
- **8 Models:**
 - **Conduction Disorders**
 - **Coronary Artery Disorders**
 - **Electrocardiogram Disorders**
 - **Heart Failure**
 - **Myocardial Disorders**
 - **Palpitations**
 - **Rate Rhythm Disorders**
 - **Valve Disorders**

Carcinogenicity References

- QSAR modeling of carcinogenic risk using discriminant analysis and topological molecular descriptors. Contrera *et al.* *Curr. Drug Discov. Technol.* 2:55-67, 2005
- Predicting the carcinogenic potential of pharmaceuticals in rodents using molecular structural similarity and E-state indices. Contrera *et al.* *Reg. Toxicol. Pharmacol.* 38:243-259, 2003
- A new highly specific method for predicting the carcinogenic potential of pharmaceuticals in rodents using enhanced MCASE QSAR-ES software. Matthews and Contrera. *Reg. Toxicol. Pharmacol.* 28:242-264, 1998

Genetic Toxicology References

- An analysis of genetic toxicity, reproductive and developmental toxicity, and carcinogenicity data: I. Identification of carcinogens using surrogate endpoints. Matthews *et al. Reg. Toxicol. Pharmacol.* 42:83-96, 2006
- An analysis of genetic toxicity, reproductive and developmental toxicity, and carcinogenicity data: II. Identification of genotoxicants, reprotoxicants, and carcinogens using *in silico* methods. Matthews *et al. Reg. Toxicol. Pharmacol.* 42:97-110, 2006
- *In silico* screening of chemicals for bacterial mutagenicity using electrotopological E-state indices and MDL-QSAR software. Contrera *et al. Reg. Toxicol. Pharmacol.* 43:313-323, 2005

General References

- Toxicology and clinical computational analysis and the US FDA/CDER. Benz. *Expert Opin. Drug Metab. Toxicol.* 3:109-124, 2007
- *In silico* approaches to explore toxicity end points: issues and concerns for estimating human health effects. Matthews and Contrera. *Expert Opin. Drug Metab. Toxicol.* 3:125-134, 2007
- Progress in QSAR Toxicity Screening of Pharmaceutical Impurities and Other FDA Regulated Products. Kruhlak *et al.* *Advances in Drug Delivery Reviews* (in press) 2007
- Use of toxicological information in drug design. Matthews *et al.* *J. Mol. Graphics Mod.* 18:605-615, 2000

[http://www.fda.gov/cder/
Offices/OPS_IO/default.htm](http://www.fda.gov/cder/Offices/OPS_IO/default.htm) ↓