Using Quantitative Predictions of Continuous Toxicity Values in Read-Across

Jessica Wignall

March 14, 2017

Society of Toxicology Annual Meeting 2017

Opportunities for Read-Across Development and Application Using QSAR Approaches

Conflict of Interest

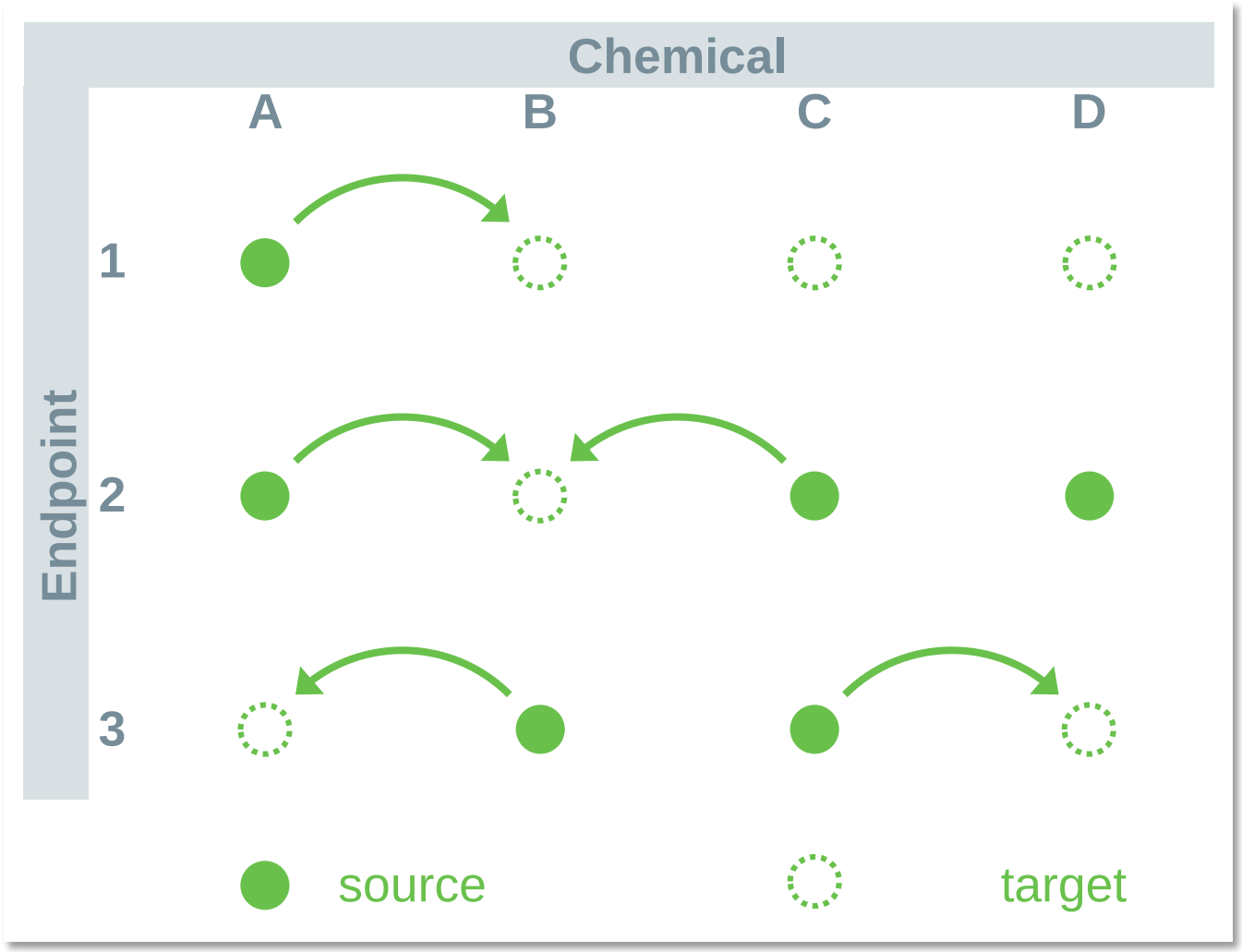
- I have no conflicts to declare.
- The work presented within was funded, in part, by grants from National Institutes of Health (P42-ES005948), U.S. Environmental Protection Agency (EPA) (STAR-RD83516601), and Oak Ridge Institute for Science and Education (ORISE).
- ICF is a contracting firm with contracts with the Federal government and industry.
- Any views presented are those of the presenter.

Acknowledgments

- **Texas A&M: Ivan Rusyn, Weihsueh Chiu**
- **UNC Chapel Hill: Alexander Tropsha, Eugene Muratov, Denis Fourches, Aleck Sedykh**
- **EPA/NCEA: Kate Guyton (now at IARC), Vincent Cogliano**
- **EPA/NCCT: Matthew Martin, Sean Watford**
- **NIEHS: Andy Shapiro**
- **California EPA: Lauren Zeise**
- **UC San Francisco: Tracey Woodruff**
- **NC State: David Reif**
- **Dow: Nicholas Ball**

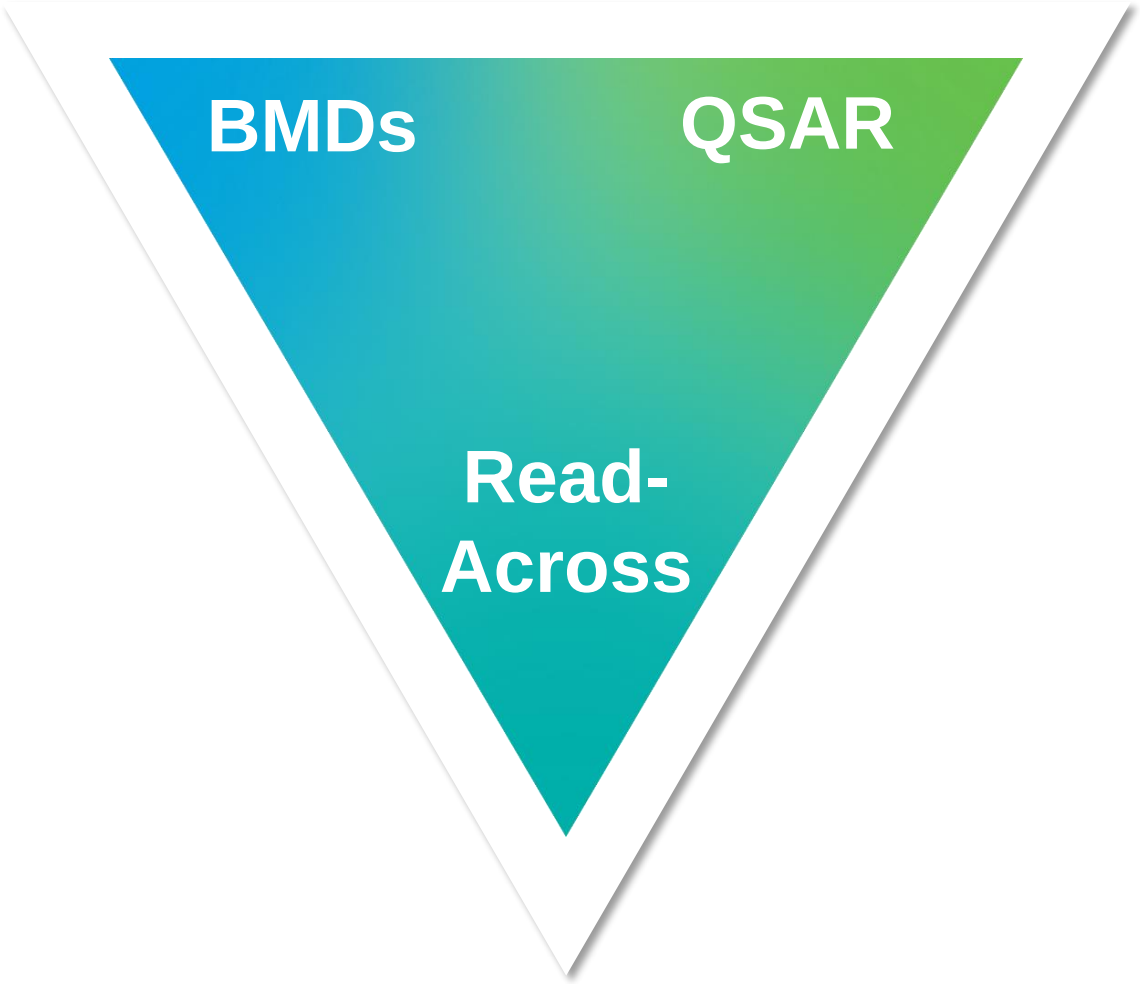
Read-Across

- **Two parts:**
  - The justification for the chemical grouping
  - The endpoint data used for target chemicals
    - –Direct use of source experimental data
    - –Translation or adaptation of source experimental data
    - –Combination of predicted and experimental data
    - –Predicted data



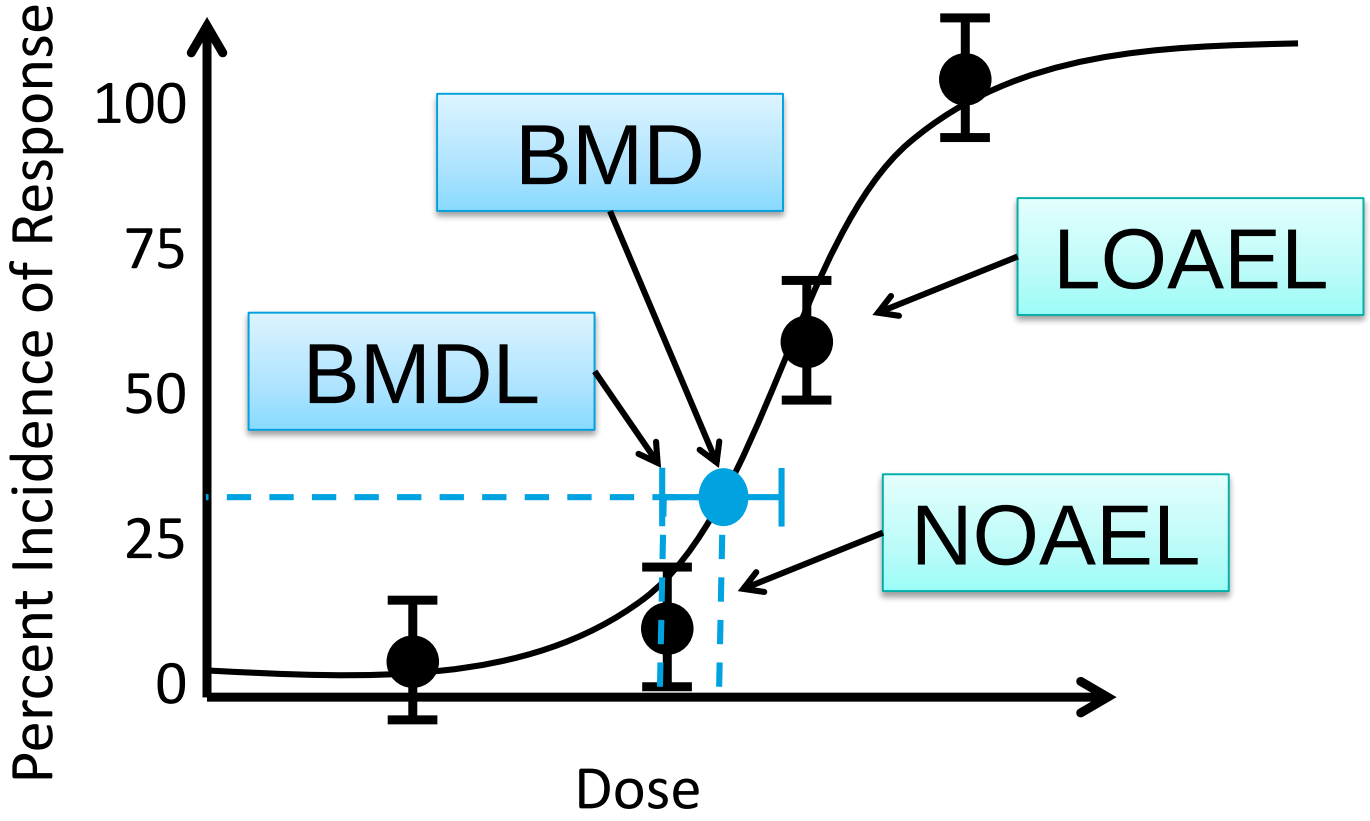
Overview

- **BMDs vs. NOAELs**
  - Using experimental data for read-across
- **QSAR**
  - Addressing the potential lack of experimental data
- **Read-Across**
  - Using QSAR methods for quantitative read-across
  - Incorporating dose-response information



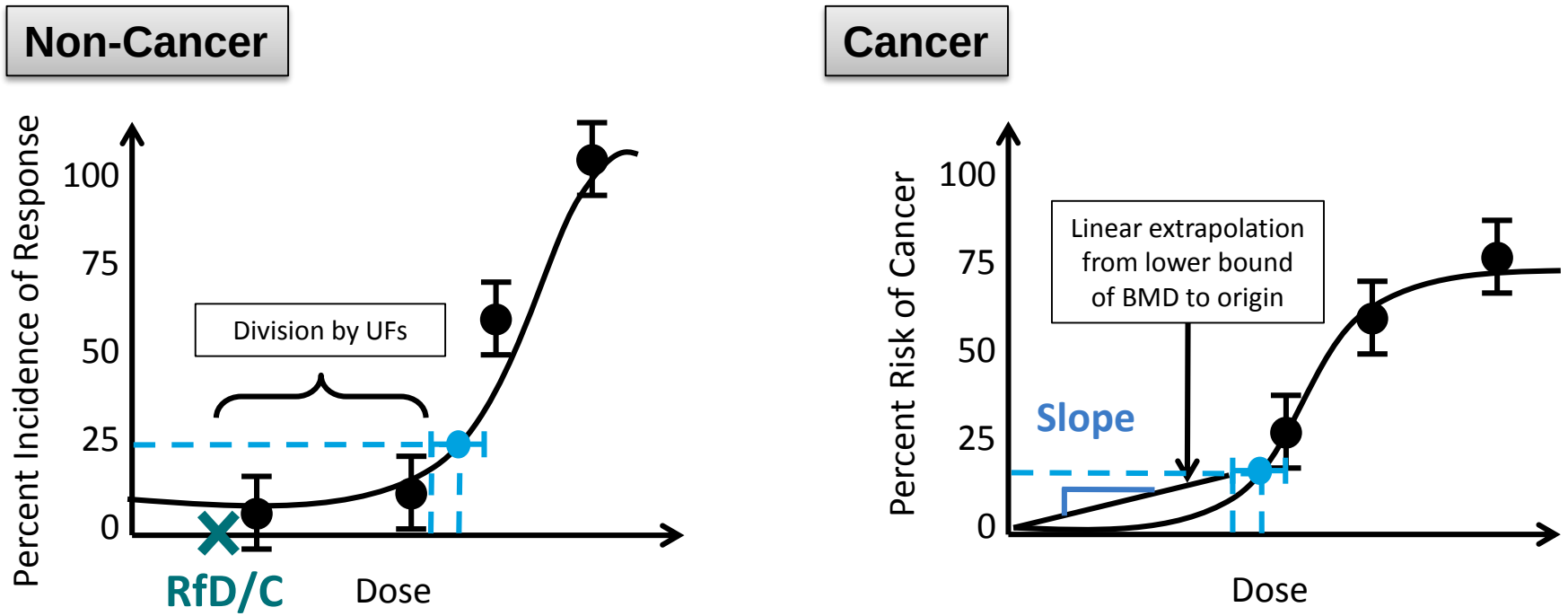
Points of Departures Can be Used in Decision Making

Quantitative Dose-Response Assessment



NOAEL = No observed adverse effect level
LOAEL = Lowest observed adverse effect level
BMD = Benchmark dose
BMDL = Benchmark dose lower confidence limit

Benchmark Dose: A Data-Driven POD



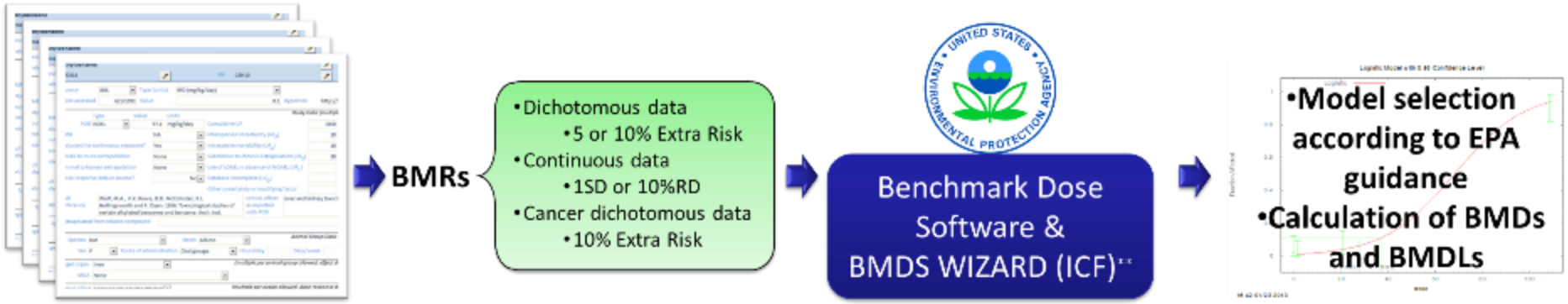
But there are limitations to BMDs:

- **Time-intensive**
- **Complex**
- **Not all data amenable to modeling**

Standardized Calculation of BMDs and BMDLs for a Large Number of Chemicals

Collected* **880** dose-response datasets for

352 unique chemicals with Toxicity Value(s) (e.g., RfD, OSF)



- **~75% of collected datasets <u>can be modeled</u> with BMDS**
- **Batch-calculated BMD/Ls available for over 300 chemicals**

See Wignall et al., 2014

*Under contract with SRC, Inc.

**Available for download from: https://www.epa.gov/bmds/download-benchmark-dose-software-bmds and https://www.icf.com/solutions-and-apps/bmds-wizard

New! BMDS Python Interface and Web Server

- If you have your own dataset(s) of interest…
- Batch process dose-response data using the U.S. EPA BMDS software:
  - The BMDS python interface package you can use on your own computer to run your datasets
  - The BMDS webserver package you can use to create automated pipelines and integrate with other websites
- Standard Python package designed to integrate into workflows for ease of use
- Validated with Wignall et al. dataset

BMD Correlation	BMDL Correlation
0.9965	0.9843

Google “python” + “BMDS”

For Python developers, "pip install bmds"

https://github.com/shapiromatron/bmds/blob/master/notebooks/2014-wignall-ehp-rerun.ipynb

Part #1: parse excel and get data in usable format

The data for this paper is available in an Excel file. Dose-response data were collapsed into different columns and semicolon delimited.

We load the Excel file into the Python, and then create datasets from each row, only including cases where there are three or more dose-groups.

We'll end up with three different lists, one for each data type:

- Continuous data
- Dichotomous data
- Dichotomous cancer data

In [ ]:
```
fn = './data/BMD_Results_2014-06-17.xlsx'
assert os.path.exists(fn)

df = pd.read_excel(fn)
df.head()
```

In [ ]:
```
continuous = df[(df['DRType'] == 'Continuous') & (df['#Doses']>=3)]
continuous = continuous[['Index', '#Doses', 'Doses', 'Mean Response',
                         'SD of response', 'Total Number of Animals']]

def continuous_dictify(d):
    try:
        return dict(
              id=d.Index,
              doses=list(map(float, d.Doses.split(';'))),
              ns=list(map(int, d['Total Number of Animals'].split(';'))),
              means=list(map(float, d['Mean Response'].split(';'))),
              stdevs=list(map(float, d['SD of response'].split(';'))),
        )
    except:
        print('Row {} not included'.format(d.Index))
        return None

continuous_datasets = [
    d for d in continuous.apply(continuous_dictify, axis=1)
    if d is not None
]
```

Lessons Learned

- **BMD/Ls are useful as points of departure**
- **BMD/Ls can be calculated in a standardized way VERY quickly**
- **These batch-calculated BMD/Ls can be used for many purposes, including:**
  - Evaluating weight of evidence for a chemical, such as across multiple studies or multiple effects
  - Interpreting or using high throughput data, including screening assays or transcriptomics
  - Serving as datasets for QSAR modeling or read-across evaluations

Data Exist for Many Types of Toxicity Values

Toxicity value type	Toxicity value name	Number of compounds with a toxicity value
Oral exposure non-cancer	Reference Dose (RfD)	668
	No Observed Adverse Effect Level (NOAEL)	487
	Benchmark Dose (BMD)	136
	Benchmark Dose Lower Level (BMDL)	136
Oral exposure cancer	Oral Slope Factor (OSF)	300
	Cancer Potency Value (CPV)	223
Inhalation exposure (non-cancer and cancer)	Reference Concentration (RfC)	149
	Inhalation Unit Risk (IUR)	148

Sources: Integrated Risk Information System; Office of Pesticide Programs; Provisional Peer-Reviewed Toxicity Values; Agency for Toxic Substances and Disease Registry; California EPA; Health Effects Assessment Summary Tables (EPA)

Considerations when Evaluating QSAR Performance

- **Model performance should be calculated based on external datasets as much as possible (Tropsha et al., 2003)**
- **Model performance is limited by how "good" the experimental data is (Lo Piparo et al., 2014)**
  - "Prediction errors cannot be better than experimental variability"
- **Model performance is improved by using both larger datasets and closely related datasets (McLellan et al., 2011)**
  - These considerations have implications for predicting *in vivo* outcomes for environmental chemicals, where data is limited and variable.

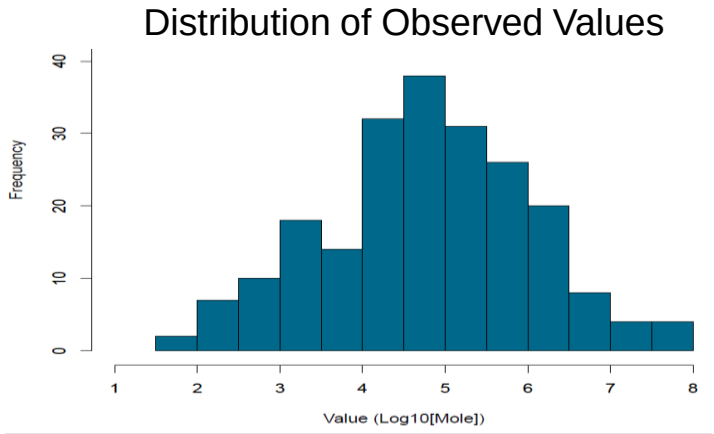
Objectives to Build Useful and Predictive Models

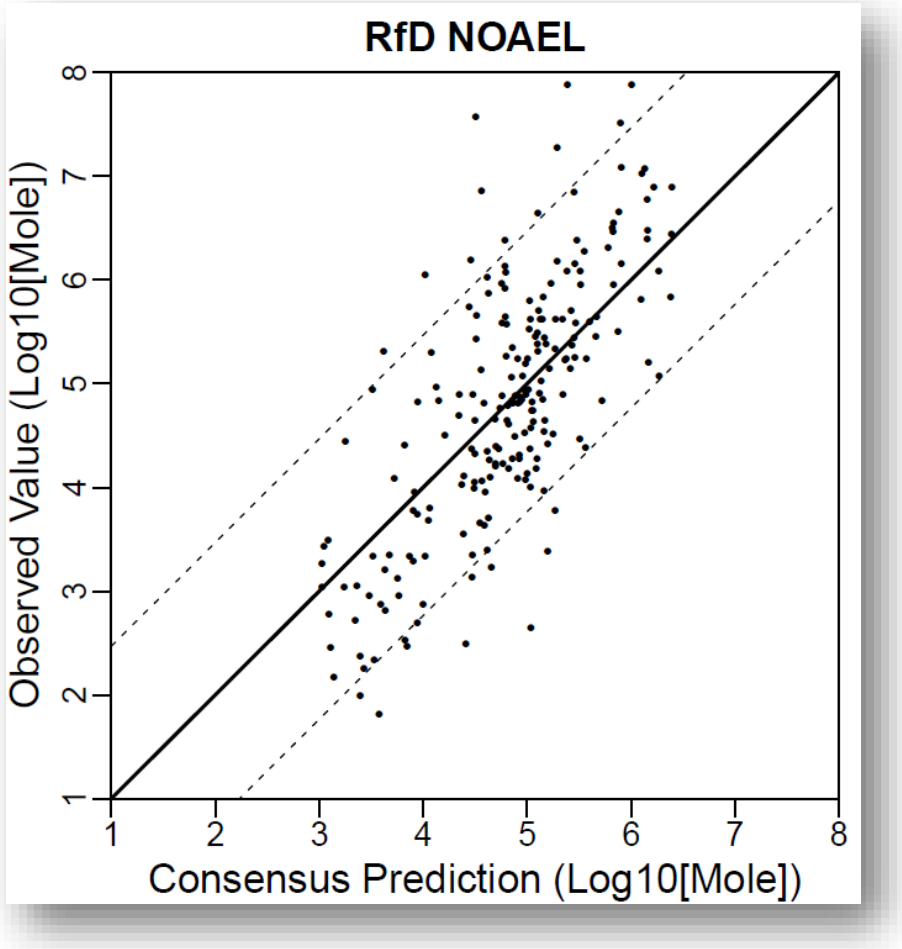
1. **Predict continuous outcomes that are of use to decision makers, including PODs.**
   - Used RfD; NOAEL; BMD; BMDL; OSF; CPV; RfC; and IUR data
2. **Facilitate transparency and communication by using publicly available chemical descriptors, easy to understand algorithms, and external validation**
   - Descriptor types: cdk + ISIDA → Consensus model
   - Algorithm: Random Forest in Python
   - Validation: 5-fold external cross-validation
3. **Provide data through accessible online portals**
   - Used online portal ChemBench* to build models
   - Models and predictions available through ToxValue.org

Carolina Cheminformatics Workbench, developed by the Carolina Exploratory Center for Cheminformatics Research (CECCR); https://chembench.mml.unc.edu/home

Model Performance Varies Across Toxicity Value Type

Toxicity value (# of compounds)	Consensus model Q^2
RfD (668)	0.48
NOAEL (487)	0.51
BMD Non-Cancer (136)	0.34
BMDL Non-Cancer (136)	0.26
OSF (300)	0.43
CPV (223)	0.38
RfC (149)	0.55
IUR (148)	0.38

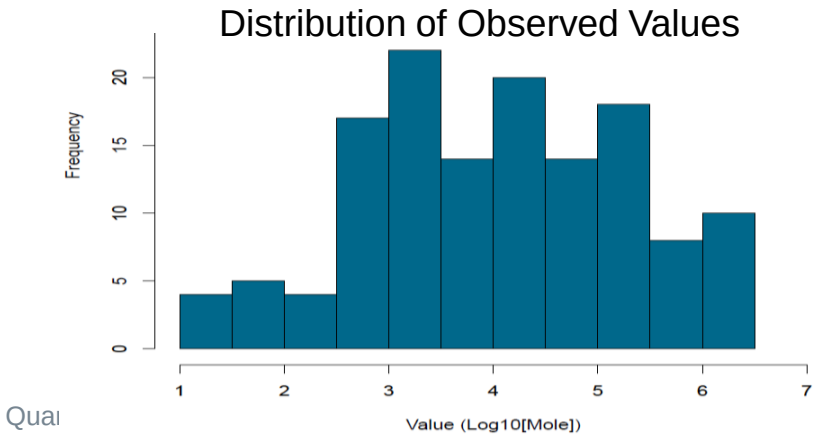




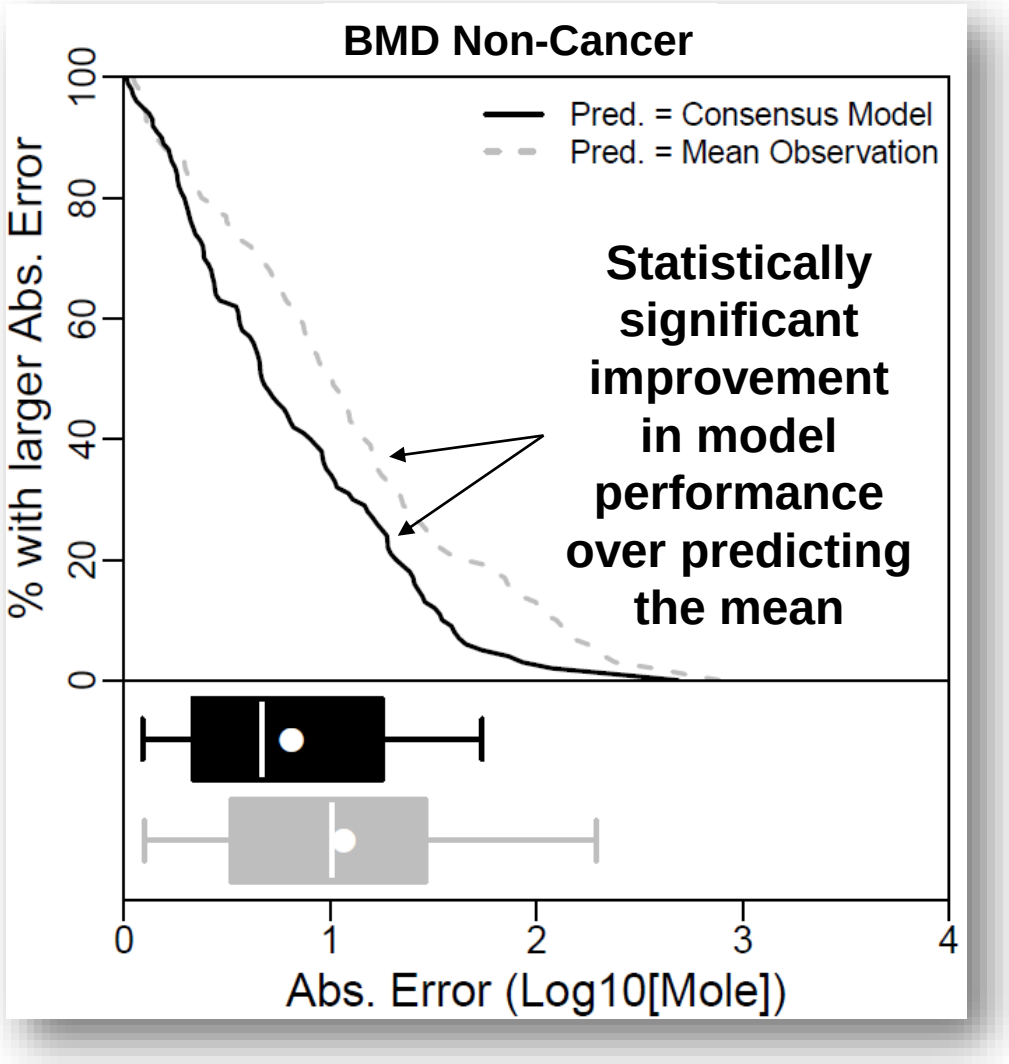
*All models were shown to perform significantly better than chance

Even Models with Low Predictivity Provide Information

Toxicity value (# of compounds)	Consensus model Q^2	p-value for improvement over average
RfD (668)	0.48	< 0.0001
NOAEL (487)	0.51	< 0.001
BMD NC (136)	0.34	0.014
BMDL NC (136)	0.26	0.12
OSF (300)	0.43	< 0.0001
CPV (223)	0.38	< 0.0001
RfC (149)	0.55	< 0.001
IUR (148)	0.38	< 0.001

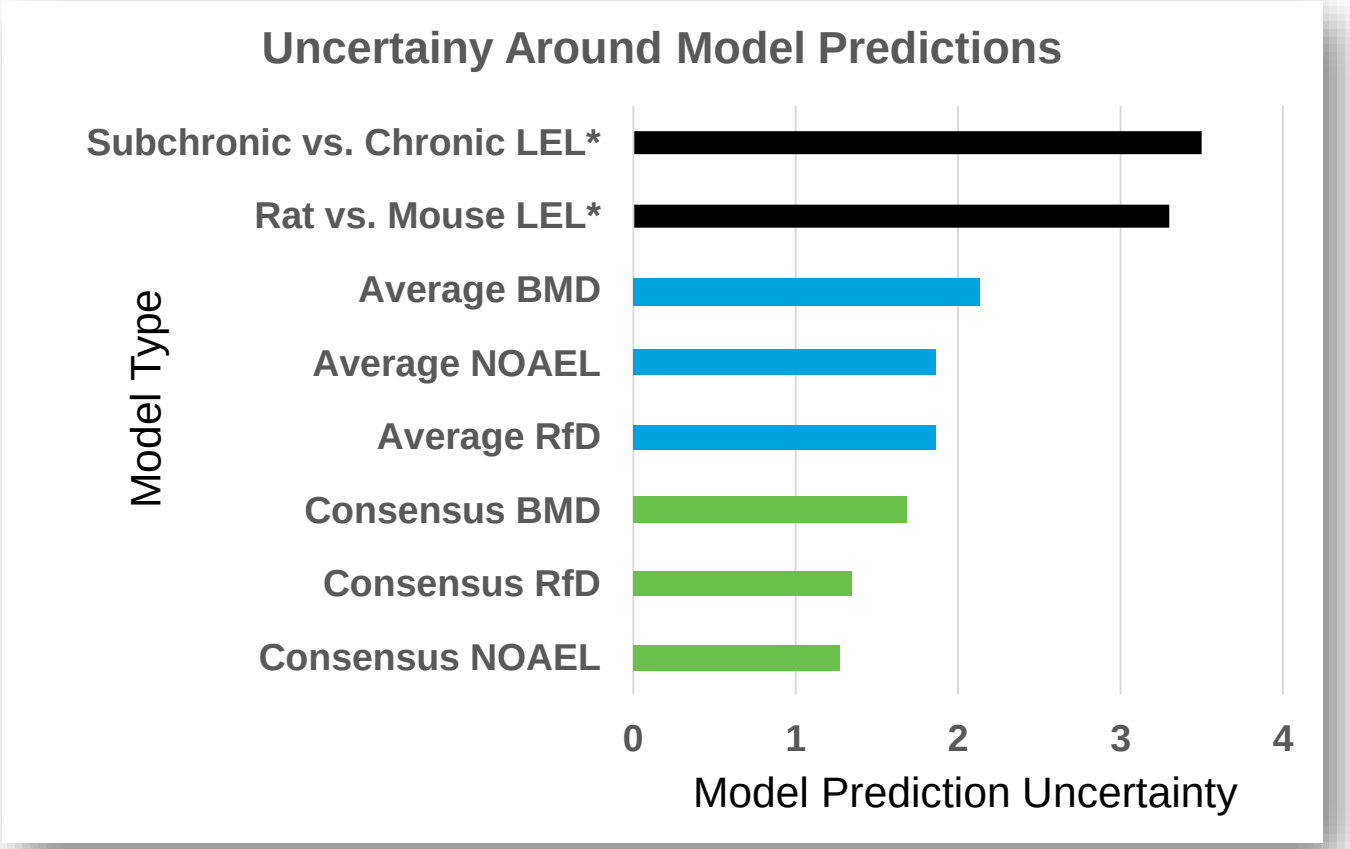


Quai



QSAR Models In the Context of Baseline Expectations of Model Uncertainty

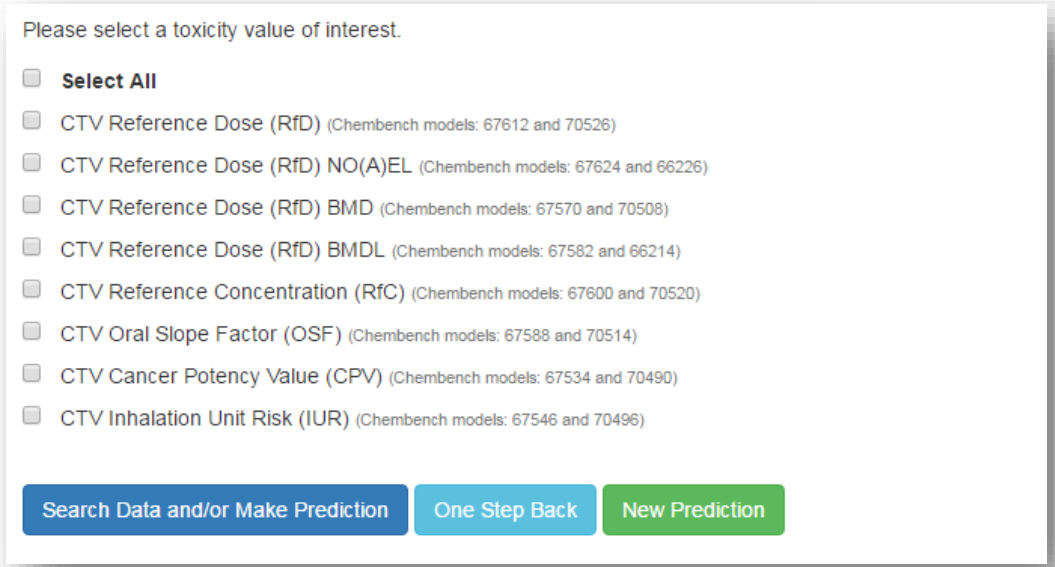
- **Uncertainty around model predictions can be benchmarked against ability to predict rat chronic lowest effect levels (LEL) from rat subchronic LELs or other models.**
- **Consensus models reduce uncertainty around predictions compared to other model types.**

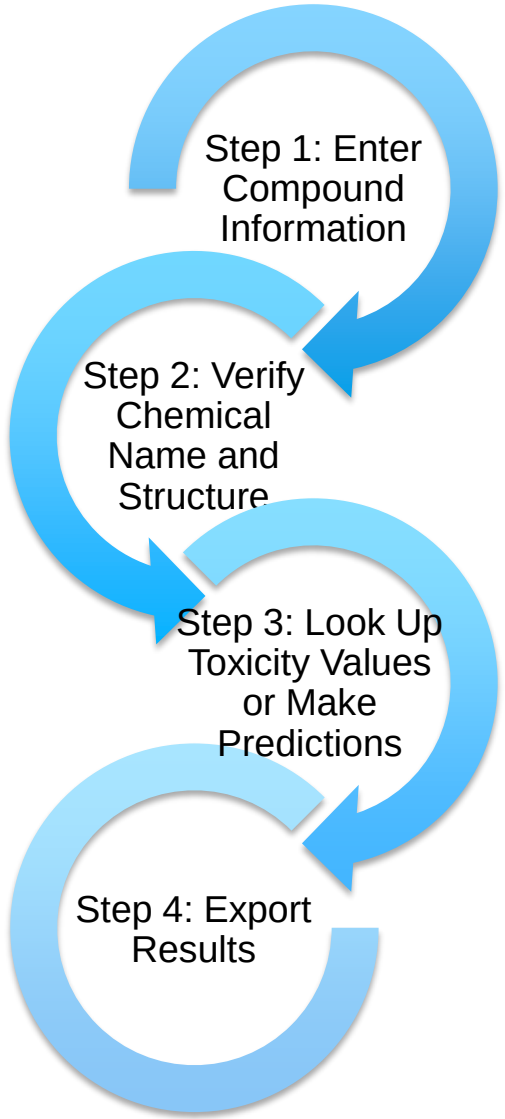


*As reported in previous analyses, source: Matt Martin, Personal Communication

Online Portal for QSAR Predictions

ToxValue.org



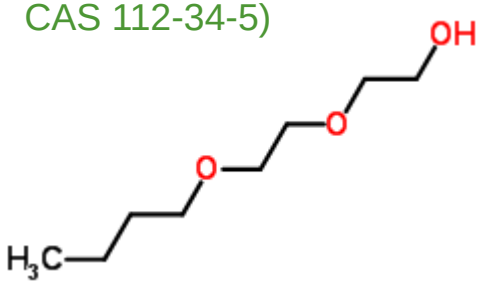
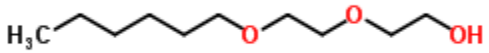
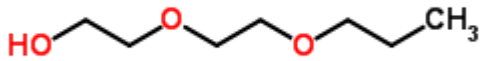


Questions? conditionaltoxvalue@gmail.com

Read-Across Example

Diethylene glycol ethers (Di EGEs)

Experimental values

Chemical	Diethylene glycol ethyl ether (DGEE, CAS 111-90-0)	Diethylene glycol monobutyl ether (DEGBE, CAS 112-34-5)	Diethylene glycol hexyl ether (DGHE, CAS No. 112-59-4)	Diethylene glycol propyl ether (DGPE, CAS 6881-94-3)
				
Critical No Effect Level	NOAEL: 167 mg/kg-day based on kidney and liver effects in pigs	NOAEL: 50 mg/kg-day for anemia in rats	?	?

Dose	Incidence
0	0/3
167	0/3
500	1/2
1117	1/1

Dose	#	Mean	SD
0	10	9.27	0.35
50	10	9.13	0.22
250	10	8.94	0.34
1000	10	8.53	0.31

	50 mg/kg-day ↕ 167 mg/kg-day	50 mg/kg-day ↕ 167 mg/kg-day

Quantitative Predictions for Read-Across

Read-Across Example

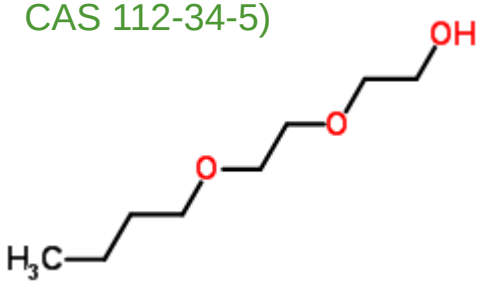
Diethylene glycol ethers (Di EGEs)

Experimental values
Predicted values

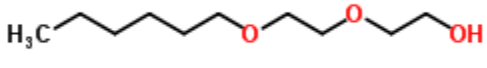
Chemical

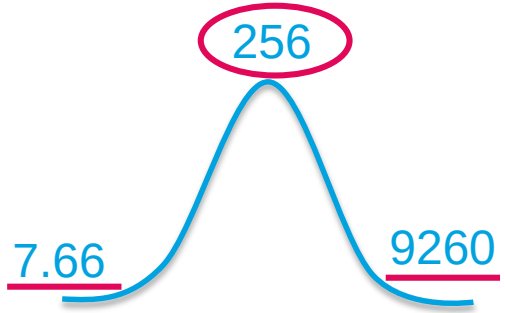
Diethylene glycol ethyl ether (DGEE, CAS 111-90-0)

Diethylene glycol monobutyl ether (DEGBE, CAS 112-34-5)

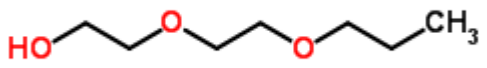


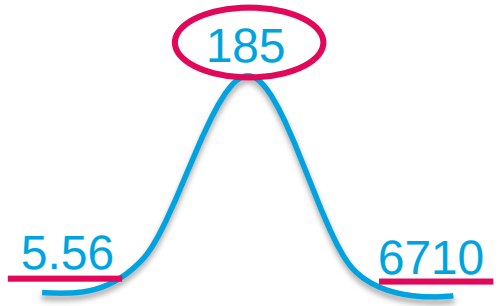
Diethylene glycol hexyl ether (DGHE, CAS No. 112-59-4)





Diethylene glycol propyl ether (DGPE, CAS 6881-94-3)





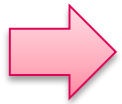
Critical Point of Departure

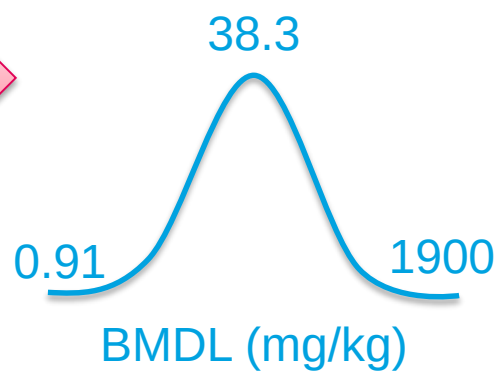
NOAEL: 167 mg/kg-day based on kidney and liver effects in pigs

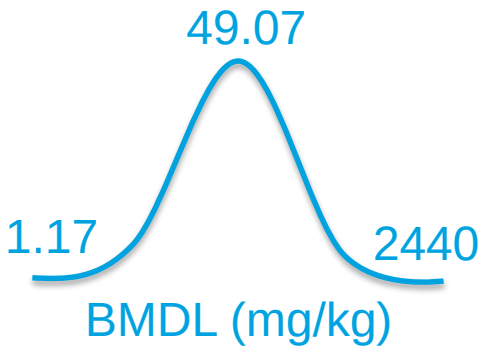
BMD 443 mg/kg-day
BMDL 45.2 mg/kg-day

NOAEL: 50 mg/kg-day for anemia in rats

BMD 222 mg/kg-day
BMDL 81.4 mg/kg-day



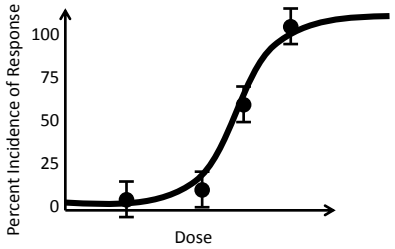




Additional BMD Data Coming…

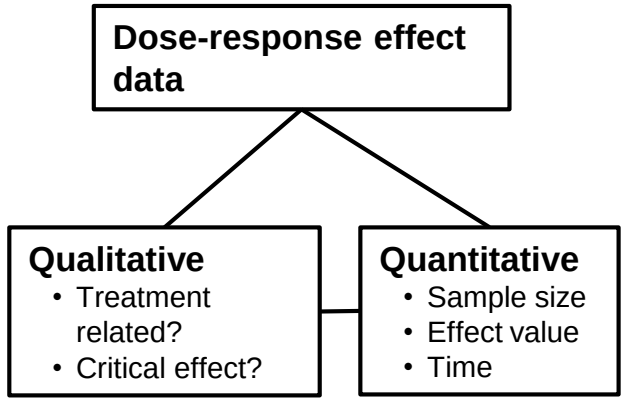
- **Assumptions inherent in aggregating various systemic toxicity endpoints into one dataset**
  - The more the homogenous the better, but balanced against need for robust training sets
- **Limited in vivo data for model building**
  - However, efforts underway to extract additional quantitative dose-response data from ToxRefDB animal studies

NOAEL = 1000 mg/kg



ToxRefDB 2.0 so far…

- **Processed over 2100 chronic and subchronic studies**
  - Includes Office of Pesticide Programs Data Evaluation Records and National Toxicology Program studies
- **200K quantitative data points**

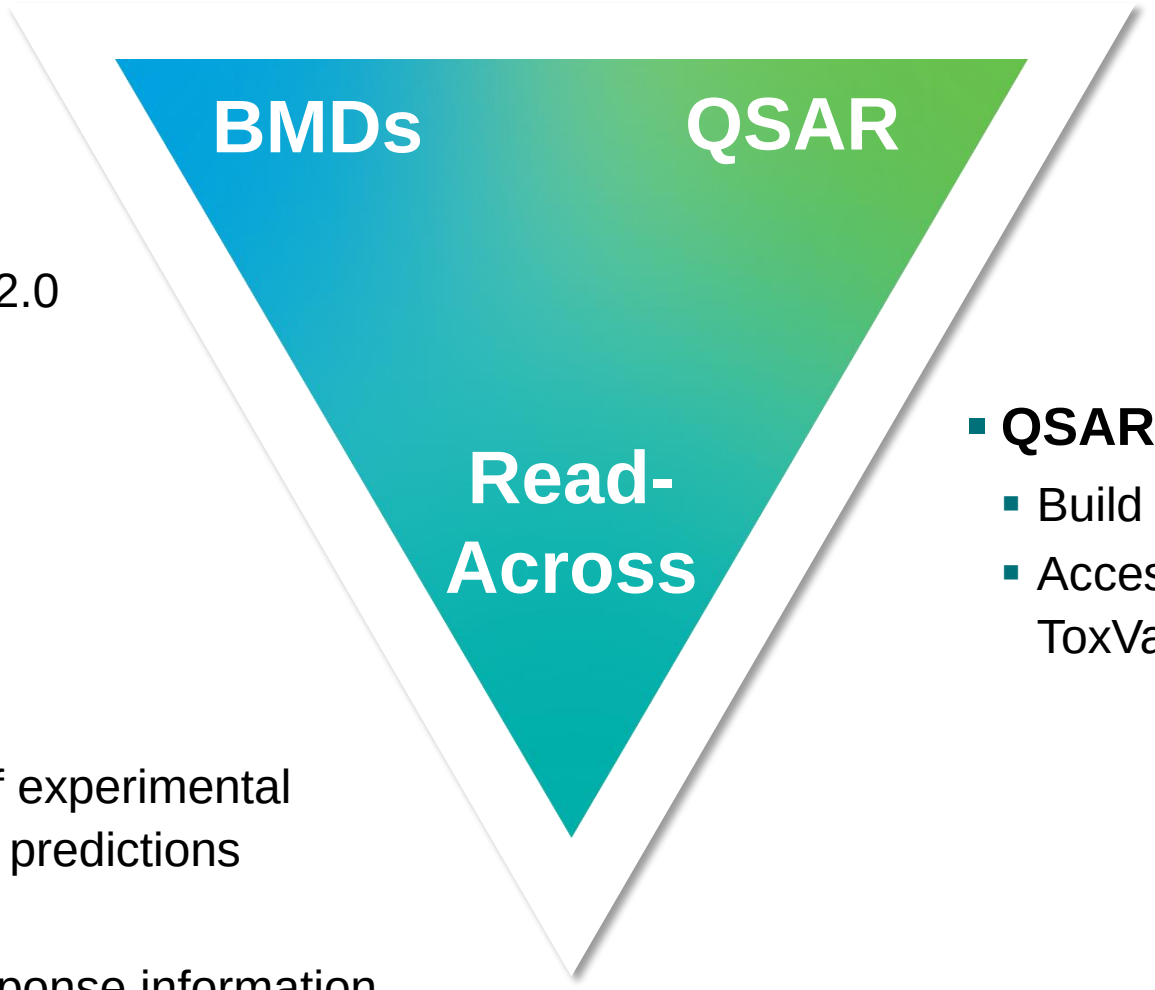


Dose	Sample Size	Mean	Standard Deviation
0	25	2.61	0.81
10	25	2.81	1.19
50	24	2.96	1.37
150	24	4.66	1.72
400	24	11.23	2.84

As reported in Abstract #1386, Watford et al.

Summary

- **BMDs**
  - Use when data are available
  - More coming from ToxRefDB 2.0
  - Open access Python tools



- **QSAR**
  - Build predictive models
  - Access existing models on ToxValue.org

- **Read-Across**
  - Use a combination of experimental (when available) and predictions
  - Quantify uncertainty
  - Incorporate dose-response information

Thank you!

References

- https://github.com/shapiromatron/bmds/blob/master/notebooks/2014-wignall-ehp-rerun.ipynb
- https://github.com/shapiromatron/bmds
- **Models referenced in this presentation can be found at the following sites**
  - http://www.toxvalue.org/
  - https://chembench.mml.unc.edu/home
- **Cdk descriptors: http://wiki.qspr-thesaurus.eu/w/CDK**
- **Ferrari T., Gini G. 2010. An open source multistep model to predict mutagenicity from statistical analysis and relevant structural alerts. Chemistry Central Journal, 4(Suppl 1):S2.**
- **ISIDA descriptors: http://infochim.u-strasbg.fr/spip.php?rubrique49**
- **Lo Piparo E, Maunz A, Helma C, Vorgrimmler D, Schilter B. 2014. Automated and reproducible read-across like models for predicting carcinogenic potency. Regul Toxicol Pharmacol. 70(1):370-8. doi: 10.1016/j.yrtph.2014.07.010. Epub 2014 Jul 15.**
- **McLellan, M. R., Ryan, M. D., & Breneman, C. M. 2011. Rank Order Entropy: why one metric is not enough. Journal of Chemical Information and Modeling, 51(9), 2302–2319. http://doi.org/10.1021/ci200170k**
- **Politi R, Rusyn I, Tropsha A. 2014. Prediction of binding affinity and efficacy of thyroid hormone receptor ligands using QSAR and structure-based modeling methods, Toxicology and Applied Pharmacology, 280(1):177-189, ISSN 0041-008X, http://dx.doi.org/10.1016/j.taap.2014.07.009.**
- **Tropsha, A., Gramatica, P. and Gombar, Vijay K. 2003. The Importance of Being Earnest: Validation is the Absolute Essential for Successful Application and Interpretation of QSPR Models. QSAR Comb. Sci., 22: 69–77. doi: 10.1002/qsar.200390007**
- **Venkatapathy R, Wang CY, Bruce RM, Moudgal C. 2009. Development of quantitative structure-activity relationship (QSAR) models to predict the carcinogenic potency of chemicals I. Alternative toxicity measures as an estimator of carcinogenic potency. Toxicol Appl Pharmacol. 2009 Jan 15;234(2):209-21. doi: 10.1016/j.taap.2008.09.028. Epub 2008 Oct 15.**
- **Wignall JA, Shapiro AJ, Wright FA, Woodruff TJ, Chiu WA, Guyton KZ, Rusyn I. 2014. Standardizing benchmark dose calculations to improve science-based decisions in human health assessments. Environ Health Perspect 122:499–505; http://dx.doi.org/10.1289/ehp.1307539**