Computational Methods to Address Challenges in Chemical Risk Assessment

Bio-Seminar in the Department of Electrical & Computer Engineering at Texas A&M
31 March 2017

Weihsueh A. Chiu, PhD
Professor, Veterinary Integrative Biosciences
College of Veterinary Medicine and Biomedical Sciences

Collaborators

- TAMU-CVM
  - Ivan Rusyn
  - David Threadgill
  - Postdoctoral associates
    - Nan-hung Hsieh
    - Chimeddulam Dalaijamts
    - Fabian Grimm
- TAMU-GERG
  - Tony Knap
  - Terry Wade
- TAMU-EN
  - Stratos Pistikopoulos
- TAMU-HSC
  - Tommy McDonald
- Pacific Northwest National Laboratory
  - Erin Baker
  - Justin Teeguarden
- Colorado State University
  - Brad Reisfeld
  - Sudipto Ghosh
- L'Institut national de l'environnement industriel et des risques (INERIS, France)
  - Frederic Bois

Outline

- Overview of chemical risk assessment
- Examples of key challenges and role of computational methods
  - Risk from complex and varied exposures
  - Addressing population variability
  - Quantifying risk and uncertainty
- Risk assessment as translational science

Scientific Components of Risk Assessment

Source-to-Outcome Continuum

Source/stressor formation

Fate & Transport

Environmental concentrations

Exposure

External doses

Toxicokinetics

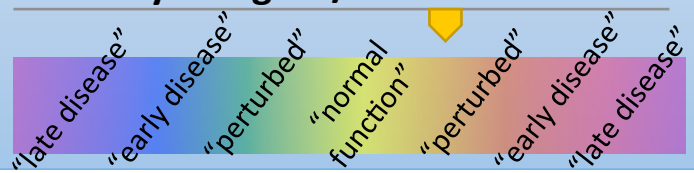
Internal concentrations

Toxicodynamics

Biological response measurements

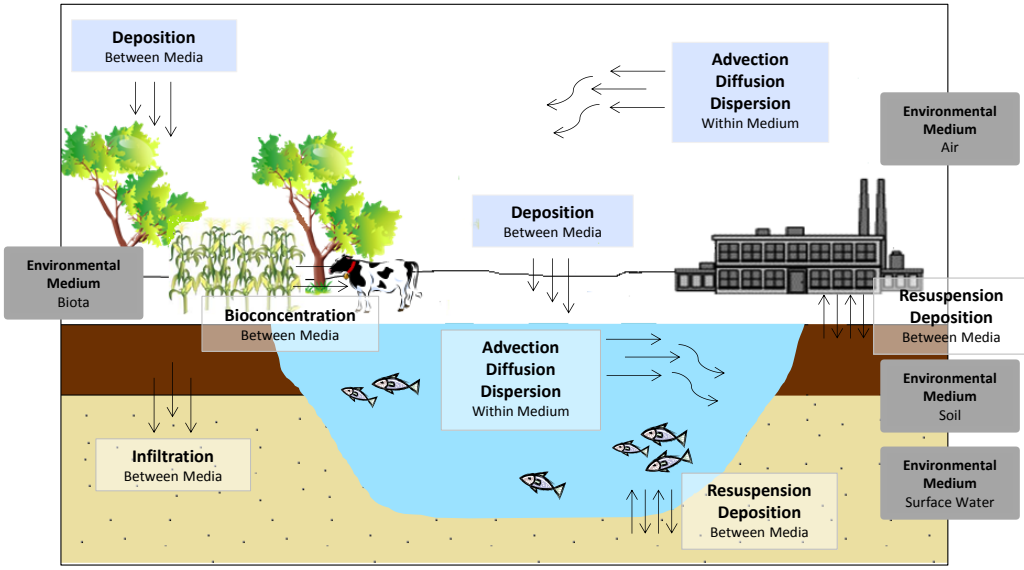
Systems dynamics

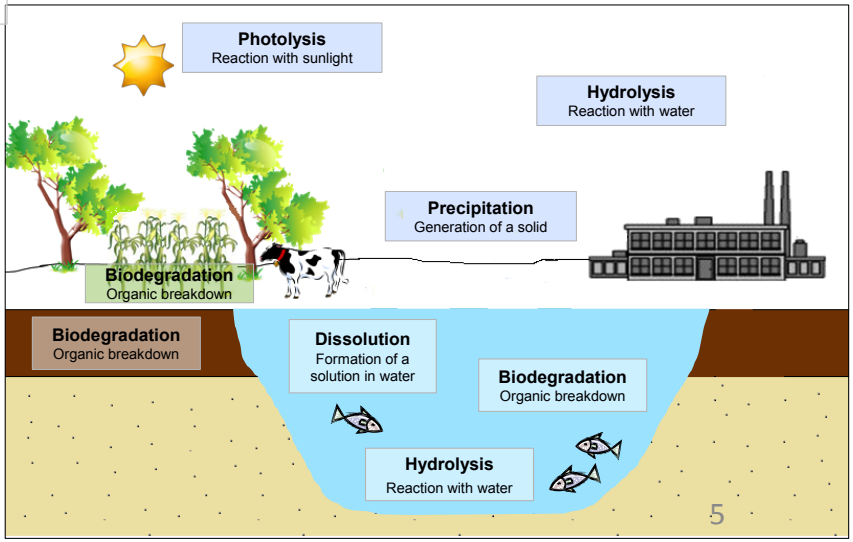
Physiological/health status



Exposure Assessment

Transport and transformation of chemicals in the environment





Exposure modeling

Storm surge from Hurricane

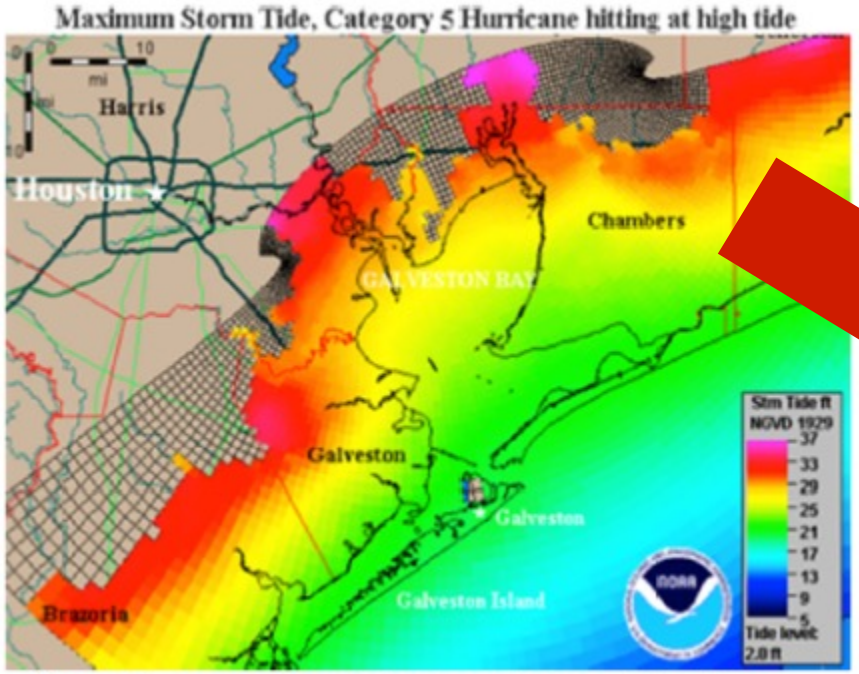
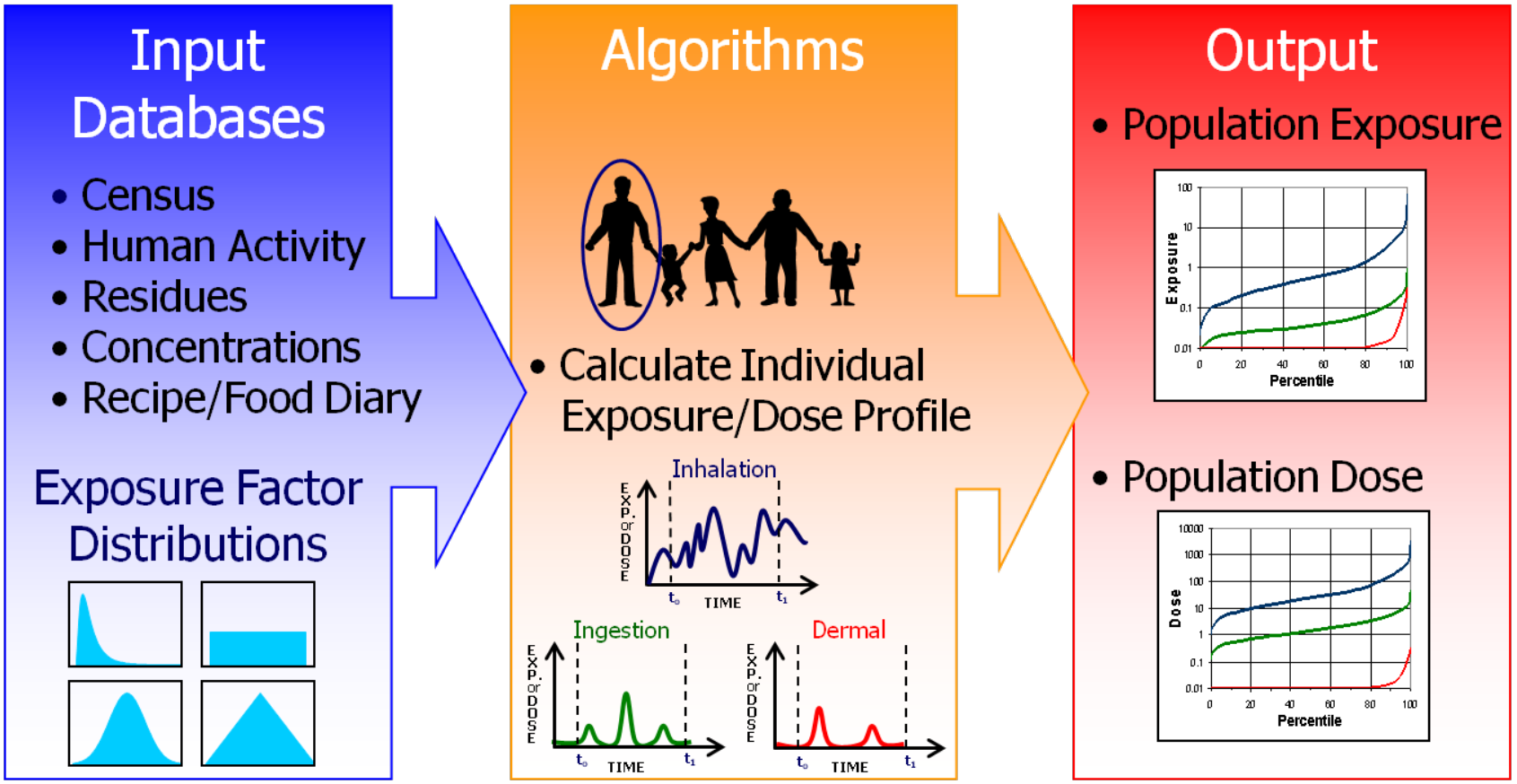


Figure 8. Coastal inundation in a Category 5 hurricane.

Sediment deposition

Figure 4. Sediments in residential New Orleans post-Katrina (Photo: Geoff Plumlee).

Estimating Human Exposure in the Population

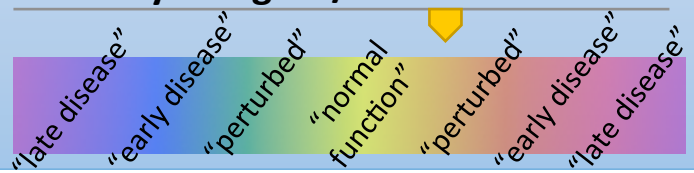


Source: SAP SHEDS Overview, 7/14/2010

Scientific Components of Risk Assessment

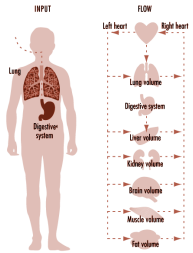
Source-to-Outcome Continuum

- Source/stressor formation
- Fate & Transport
- Environmental concentrations
- Exposure
- External doses
- Toxicokinetics
- Internal concentrations
- Toxicodynamics
- Biological response measurements
- Systems dynamics
- Physiological/health status



Exposure Assessment

Pharmacokinetics/Toxicokinetics



Toxicokinetics = "Fate and transport within the body"

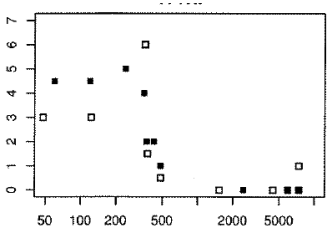
- Exposure alone is not sufficient to elicit toxicity

- Interaction between an exogenous agent and a biological target
  - What is the agent or toxic moiety?
  - How does it get to the biological target?
  - How much of the agent gets there?
  - How long does it stay there?

- **Toxicokinetics** is the study of the movement of chemicals in and out of the body ("what the body does to the chemical")
  - Absorption
  - Distribution
  - Metabolism
  - Excretion

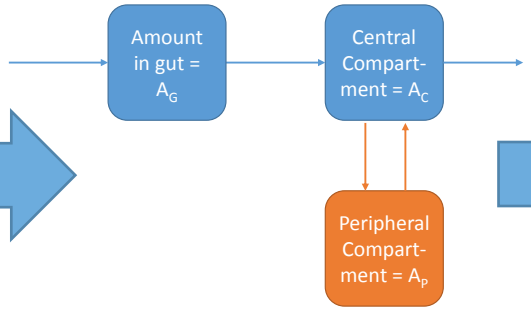
For pharmaceuticals – mostly use simple empirical models

Chemical-specific data



in vivo

+



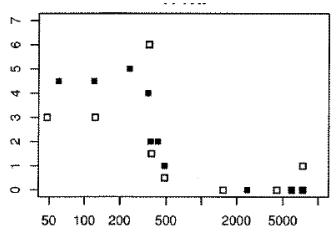
Empirical models

=

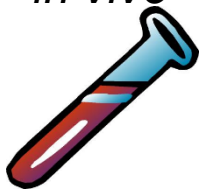
Predictions about similar scenarios

More complex models trade off simplicity for predictive power

Chemical-specific data

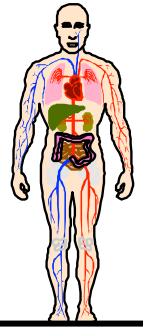


in vivo

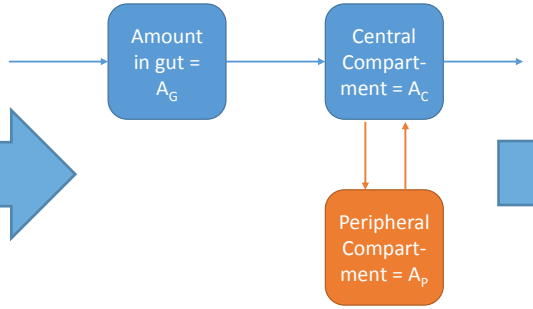


in vitro

Physiological Data



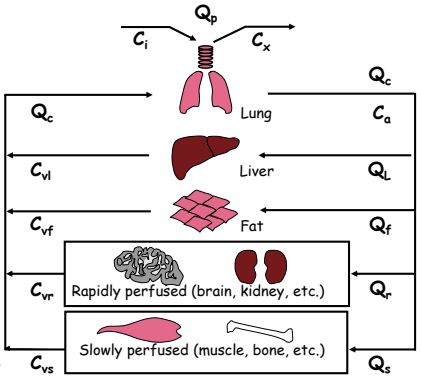
+



Empirical models

(simple & quick)

+



PBPK models

(complicated & time-consuming)

=

Predictions about similar scenarios

=

Predictions about scenarios with different:

- Exposure routes, durations, levels, patterns
- Species
- Individuals

Scientific Components of Risk Assessment

Source-to-Outcome Continuum

Source/stressor formation

Fate & Transport

Environmental concentrations

Exposure

External doses

Toxicokinetics

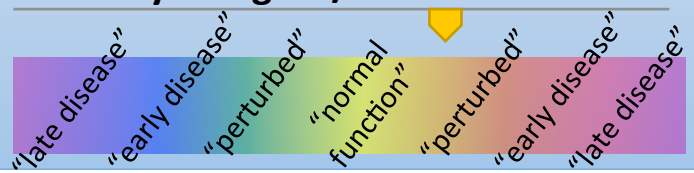
Internal concentrations

Toxicodynamics

Biological response measurements

Systems dynamics

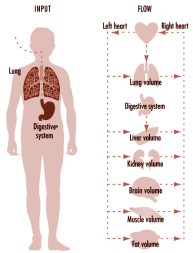
Physiological/health status

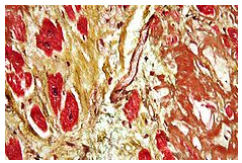


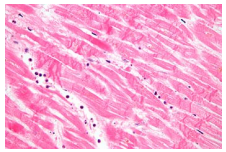
Exposure Assessment

Hazard Identification and Dose-Response Assessment

Pharmacokinetics/Toxicokinetics







Hazard Identification

- Determination of whether a particular chemical is or is not ***causally linked*** to particular health effects
  - Increased ***incidence***
  - Increased ***severity***

What adverse effects have been observed or are anticipated?

- Human data
- Laboratory animal data
- In vitro data
- Physical/chemical/molecular property data

For each adverse effect, what is the evidence that the agent can cause it in humans?

- Availability of data
  (absence of evidence ≠ evidence of absence)
- Consistency within and across the different types of data.
- Biological plausibility / mechanistic basis

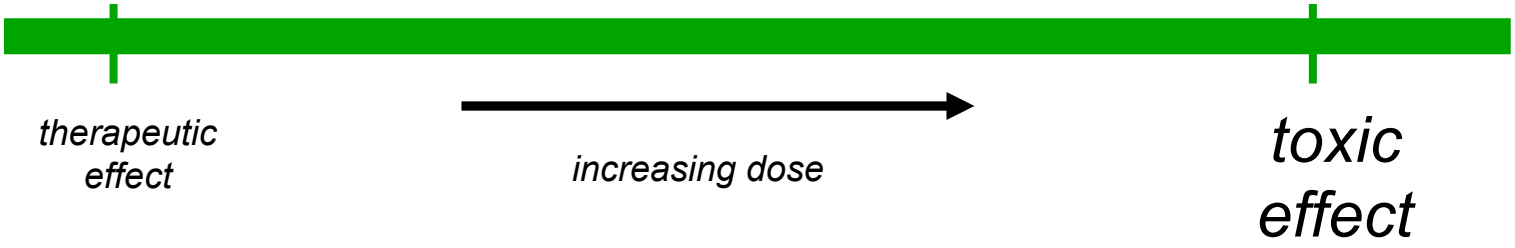
Recent emphasis has been on applying systematic review methods to evaluate evidence of causality (not discussed further today)

Dose-Response – Many still ascribe to the principles of Peracelsus…

Peracelsus
(Phillippus Aureolus Theophrastus Bombastus von Hohenheim)

1493-1541

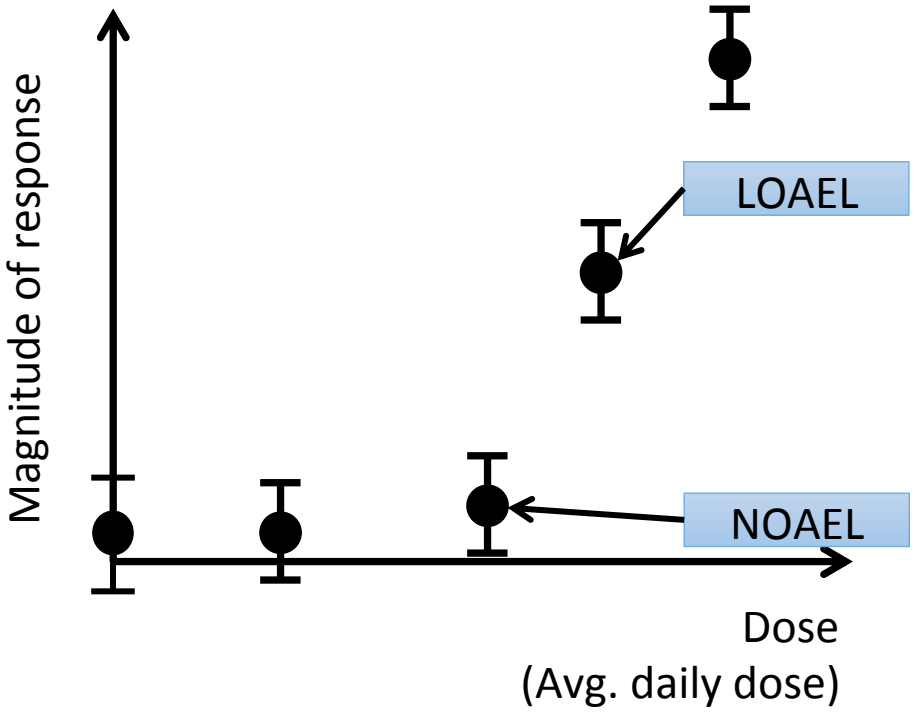
Known as the 'father of toxicology'. The saying "Dosis facit venenum" (The dose makes the poison) is attributed to him. His actual quote translates "All things are poisons, for there is nothing without poisonous qualities...it is only the dose which makes a thing poison."



Slide courtesy of D. Threadgill

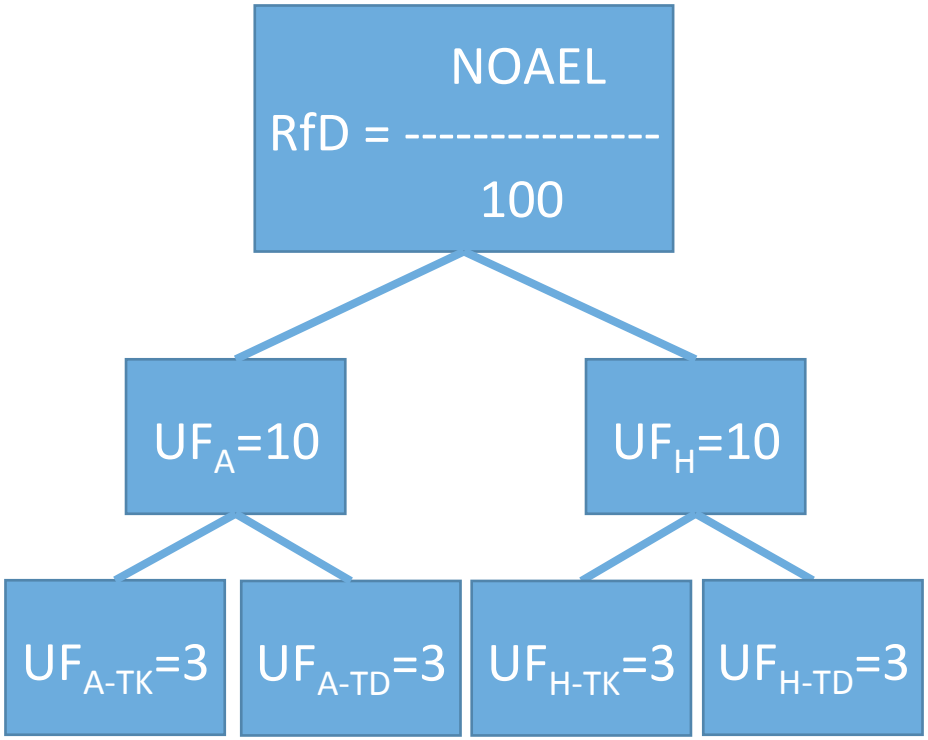
Traditional interpretation: Existence of a "threshold" below which there are no effects

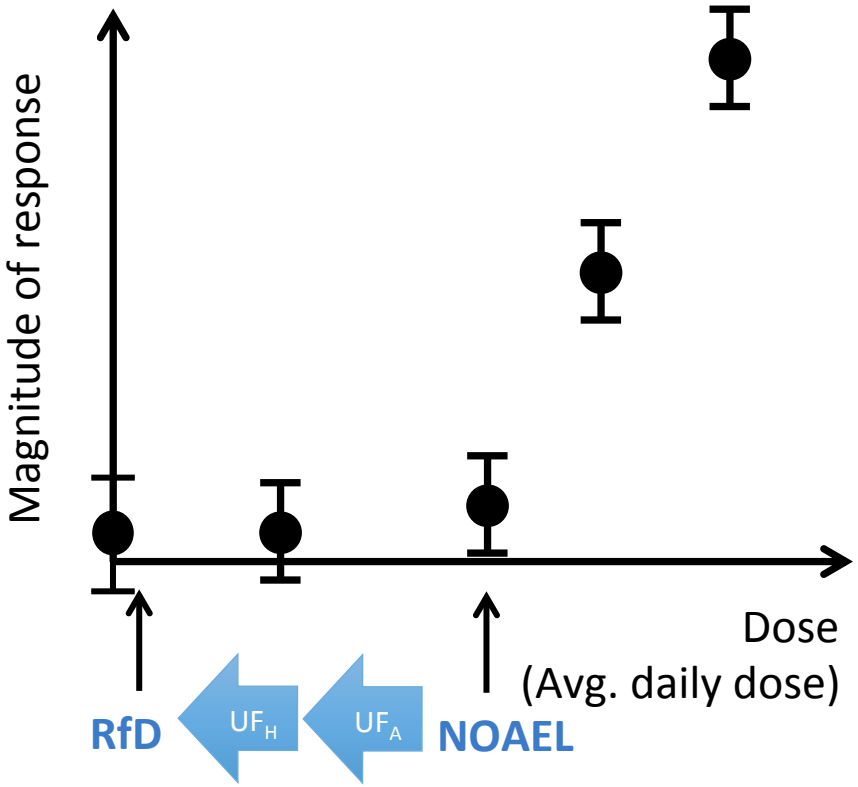
- **NOAEL:** Greatest concentration or amount of a substance, found by experiment or observation, that causes no adverse alteration …of the target organism distinguishable from those observed in normal (control) organisms of the same species and strain under the same defined conditions of exposure.*
- Commonly viewed **(incorrectly)** as an experimental dose threshold.



*WHO definition

Implementation: "Safe Human Dose" Established by Use of "Uncertainty" or "Safety" Factors

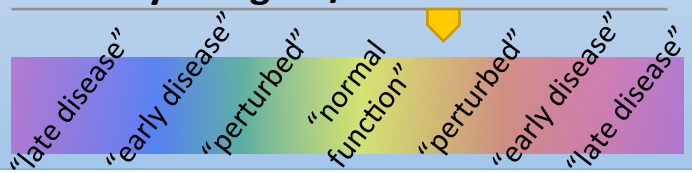




Scientific Components of Risk Assessment

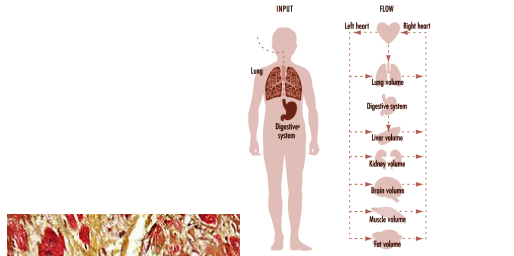
Source-to-Outcome Continuum

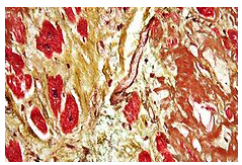
- Source/stressor formation
- Fate & Transport
- Environmental concentrations
- Exposure
- External doses
- Toxicokinetics
- Internal concentrations
- Toxicodynamics
- Biological response measurements
- Systems dynamics
- Physiological/health status

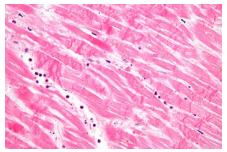


Exposure Assessment

Hazard Identification and Dose-Response Assessment

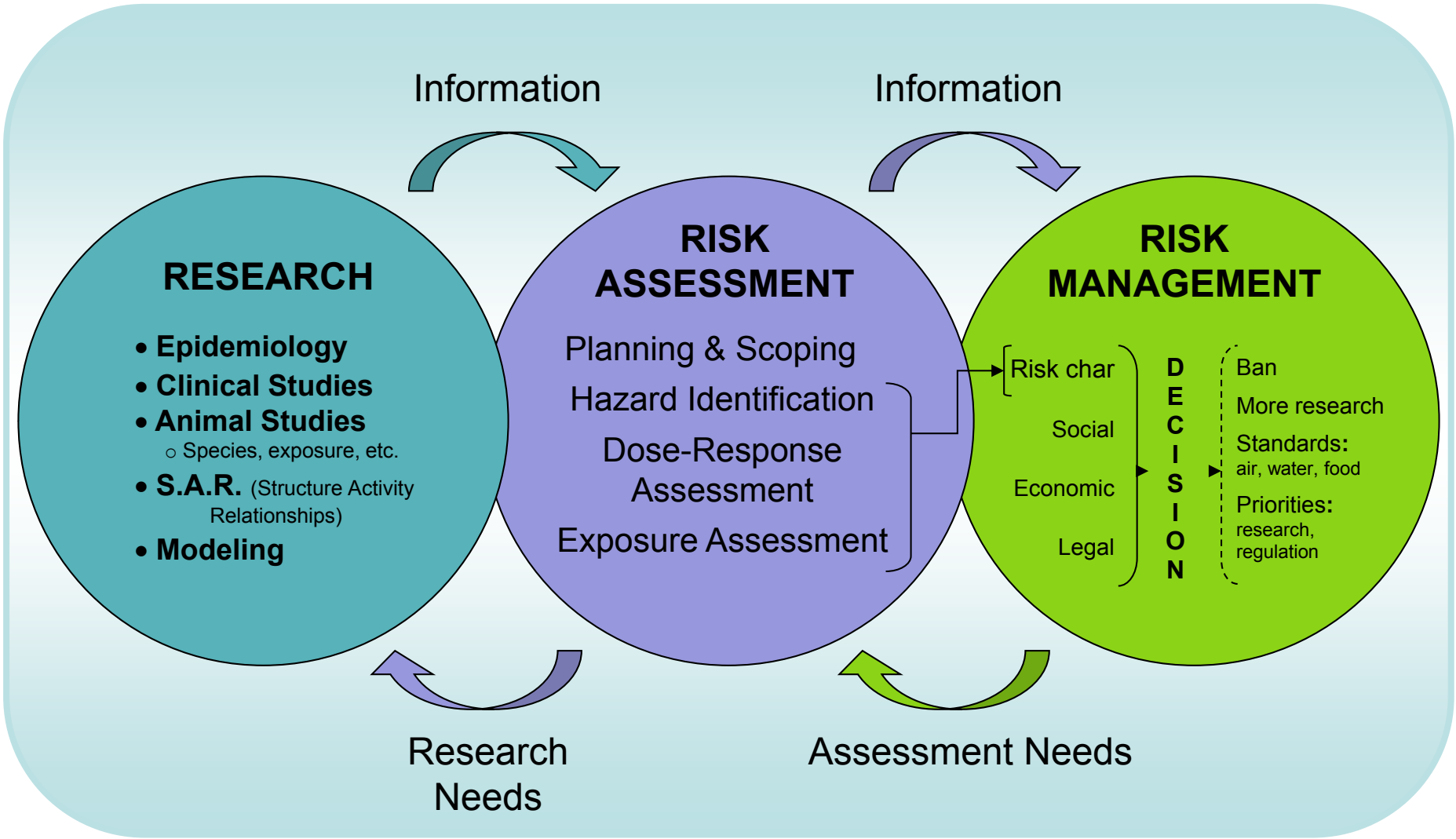






Risk Characterization

Risk Assessment in the Context of Research & Decision-Making



Multi- and trans-disciplinary nature of risk assessment

- Requires data and models from multiple scientific disciplines.

- Requires methods and approaches for integrating diverse information to draw scientific conclusions about risk.

- Requires consideration of not only scientific, but also social, economic, and legal factors in order to inform decisions about managing risk.

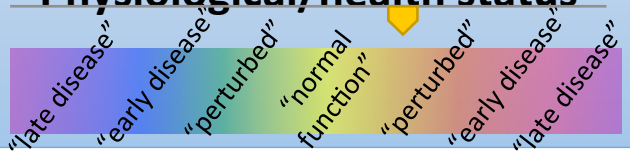
Examples of challenges and computational methods

- Complex and varied exposures with incomplete data on chemical risks
- Incomplete understanding of population variability in susceptibility to chemical risks
- Inadequate quantification of chemicals risk and its uncertainties

Example Challenge: Exposure assessment for environmental mixtures

Source-to-Outcome Continuum

- Source/media concentrations
- Exposure
- External doses
- Toxicokinetics
- Internal concentrations
- Toxicodynamics
- Biological response measurements
- Systems dynamics
- Physiological/health status



Storm surge from Hurricane

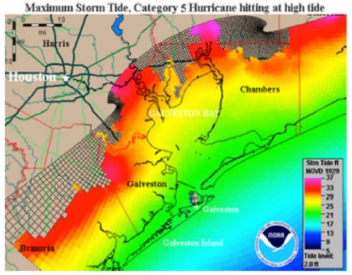


Figure 8. Coastal inundation in a Category 5 hurricane.

Sediment deposition

Figure 4. Sediments in residential New Orleans post-Katrina (Photo: Geoff Plumlee).

Usual Approach is to perform "targeted" chemical analyses:

"Known known" contaminants

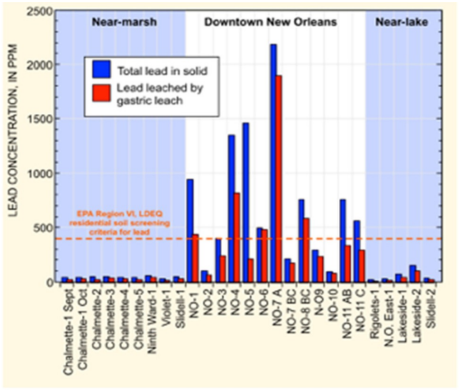


Figure 3. Lead in New Orleans soils after Katrina (1).

"Known unknown" contaminants

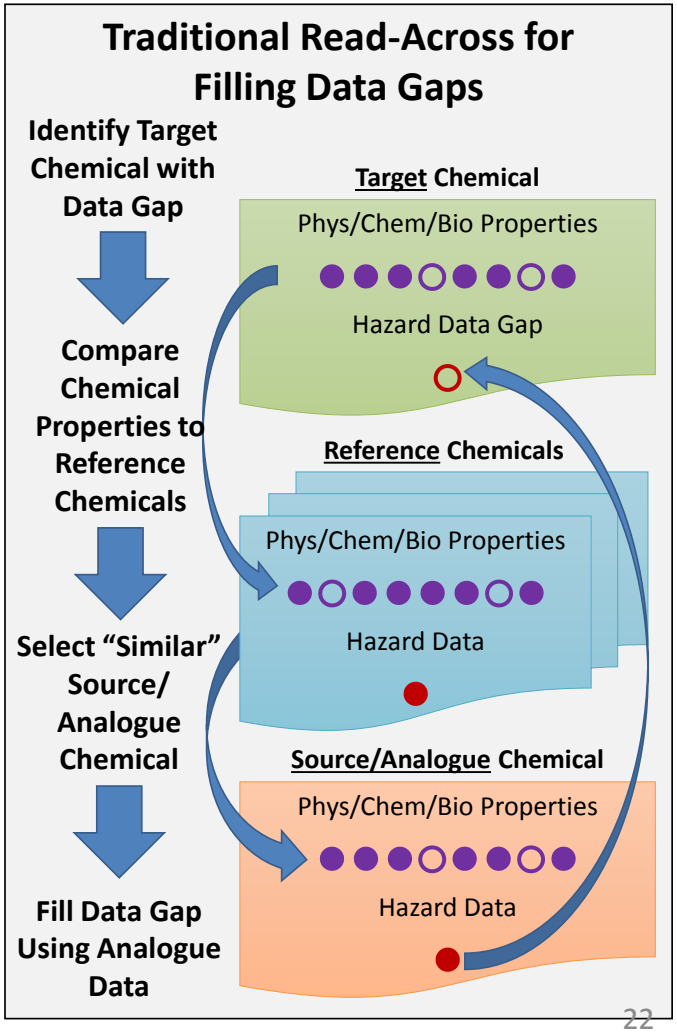
Table 2: Number of compounds per class of Contaminant and their detection limits at GERG Laboratories

Analyte Class Contaminant	Individual Component	Detection Limit
PAH	53	1 ng/g
PCB	209	1 ng/g
Dioxin and Furans	17	0.5 to 2.5 pg/g
Organochlorine Pesticides	20	1 ng/g
Organophosphate Pesticides	4	2 ng/g
Plasticizers	7	1 ug/g
Phenols	3	1 ug/g
Chlorophenols	8	1 ug/g
Nitrophenols	4	1 ug/g
Chlorinated Hydrocarbons	20	1 ug/g
Lead	1	0.4 ug/g
Mercury	1	0.01 ug/g
Cadmium	1	0.1 ug/g
Chromium	1	5 ug/g
Cobalt	1	10 ug/g
Copper	1	0.4 ug/g
Arsenic	1	0.1 ug/g
Zinc	1	0.4 ug/g
Selenium	1	0.10 ug/g
Nickel	1	0.10 ug/g

- How do you prioritize "known unknowns" given limited time and resources?
- What about "unknown unknown" contaminants?

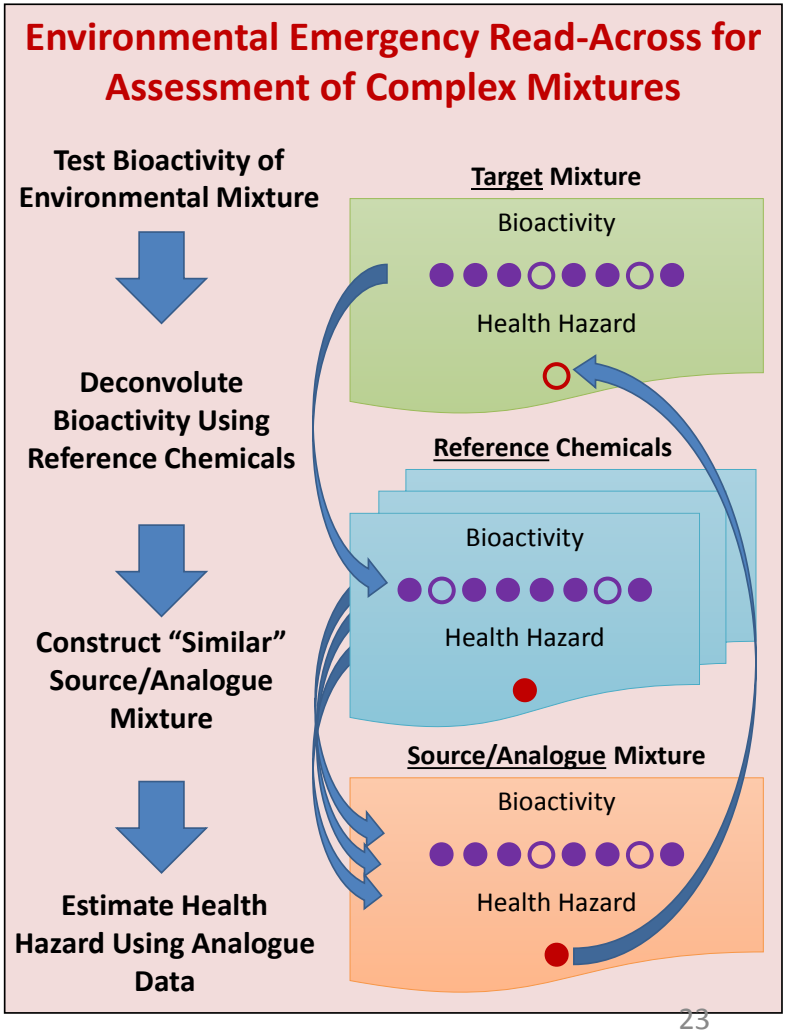
Proposed solution based on the principle of "read-across."

- Hypothesis that data gaps can be filled by "analogy"
- Requires:
  - Data and method to determine "similarity"
  - A "reference" set from which to find "analogue"
- Traditionally based on
  - Qualitative similarity in chemical structure & properties
  - Single reference chemical representative of a "group"



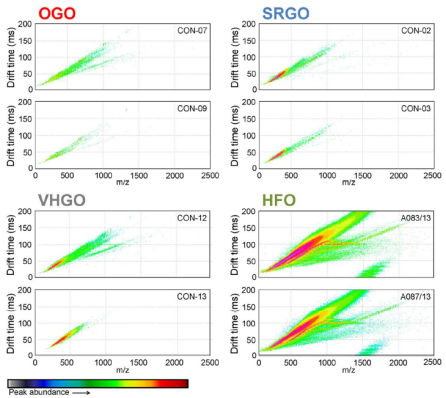
Proposed solution based on the principle of “read-across.”

- Extend the single chemical approach to environmental mixtures
- Use high-throughput chemical and biological profiling to define “similarity”
- Similarity based on mixture of reference chemicals

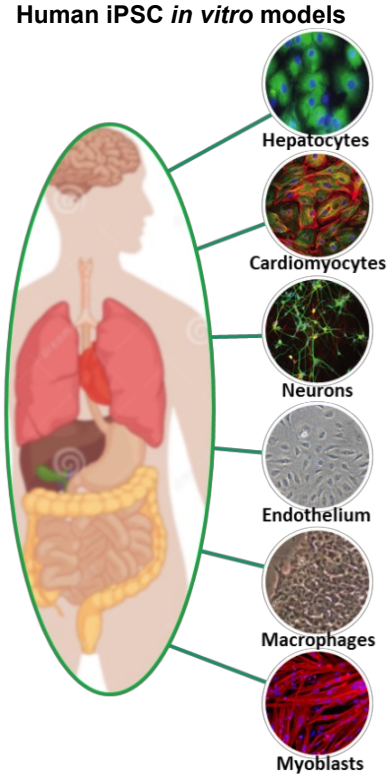


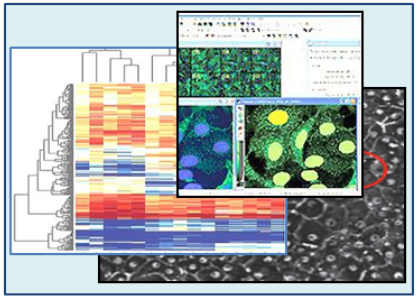
Data for defining similarity

High dimensional untargeted chemical profiling using Ion Mobility Spectroscopy/Mass Spectrometry

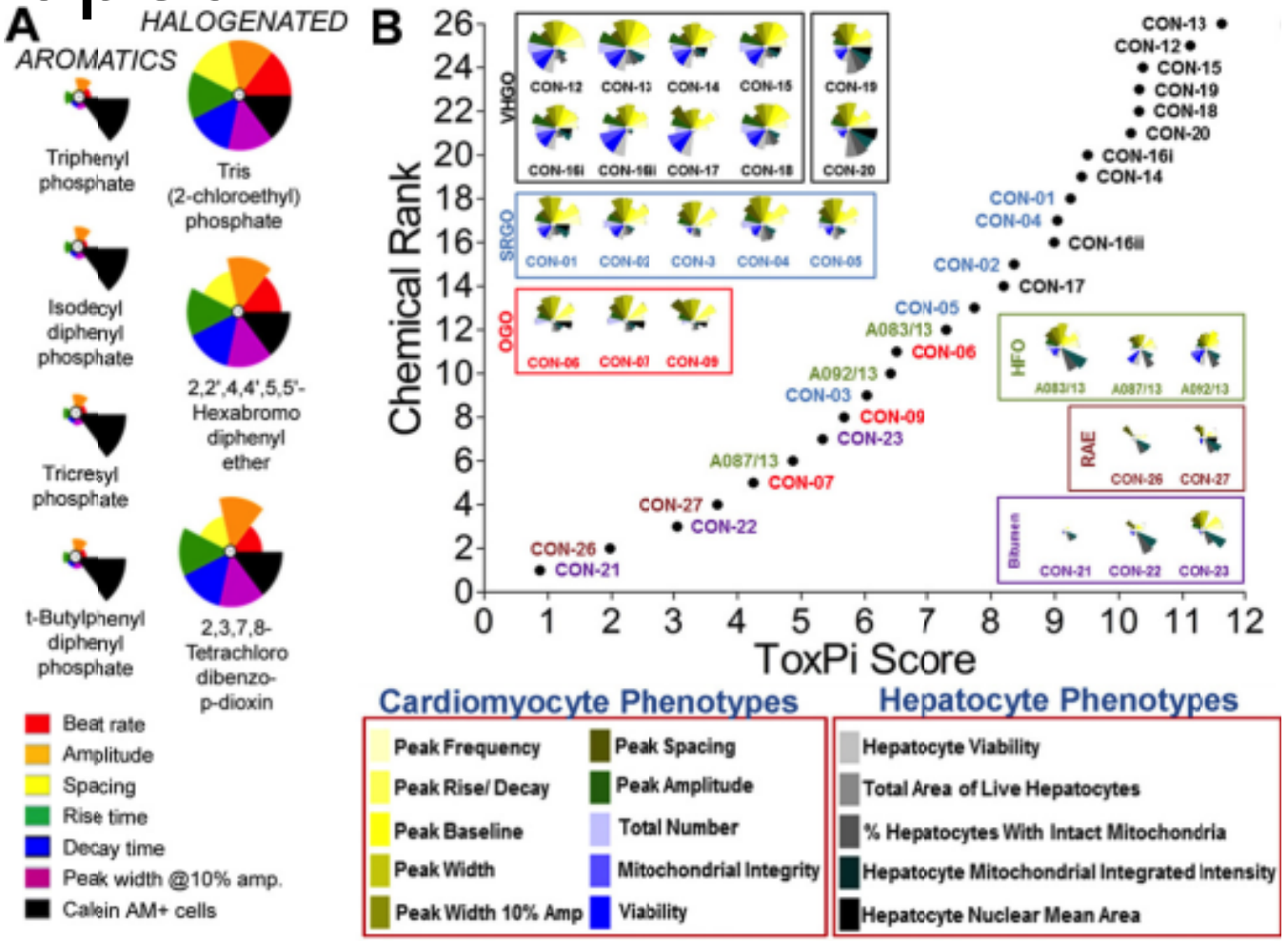


High dimensional biological profiling using induced-pluripotent stem cell-derived human tissues



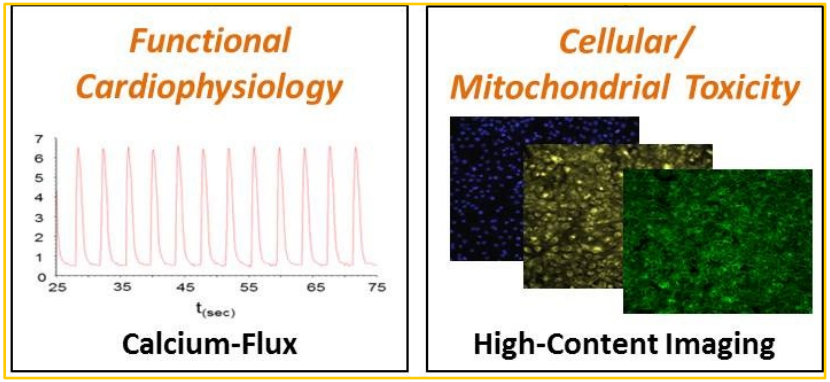


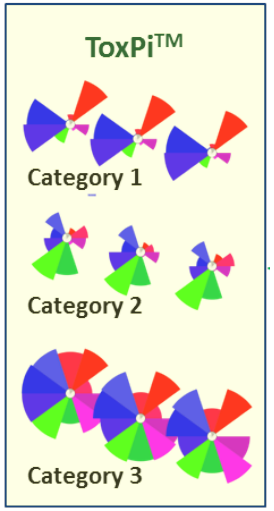
Proof of principle that individual and complex substances can be grouped

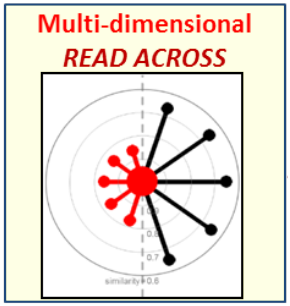


Computational demands

- Multiple types of high-dimensional data processing
  - Multi-dimensional chemical data
  - Imaging data
  - Time-series data
  - Genomic (gene expression) data
- Multivariate data integration to define "similarity"
- Deconvolution to construct "mixture analogues" using reference chemicals
- For quantifying risk, classification is not enough – need a numerical prediction.

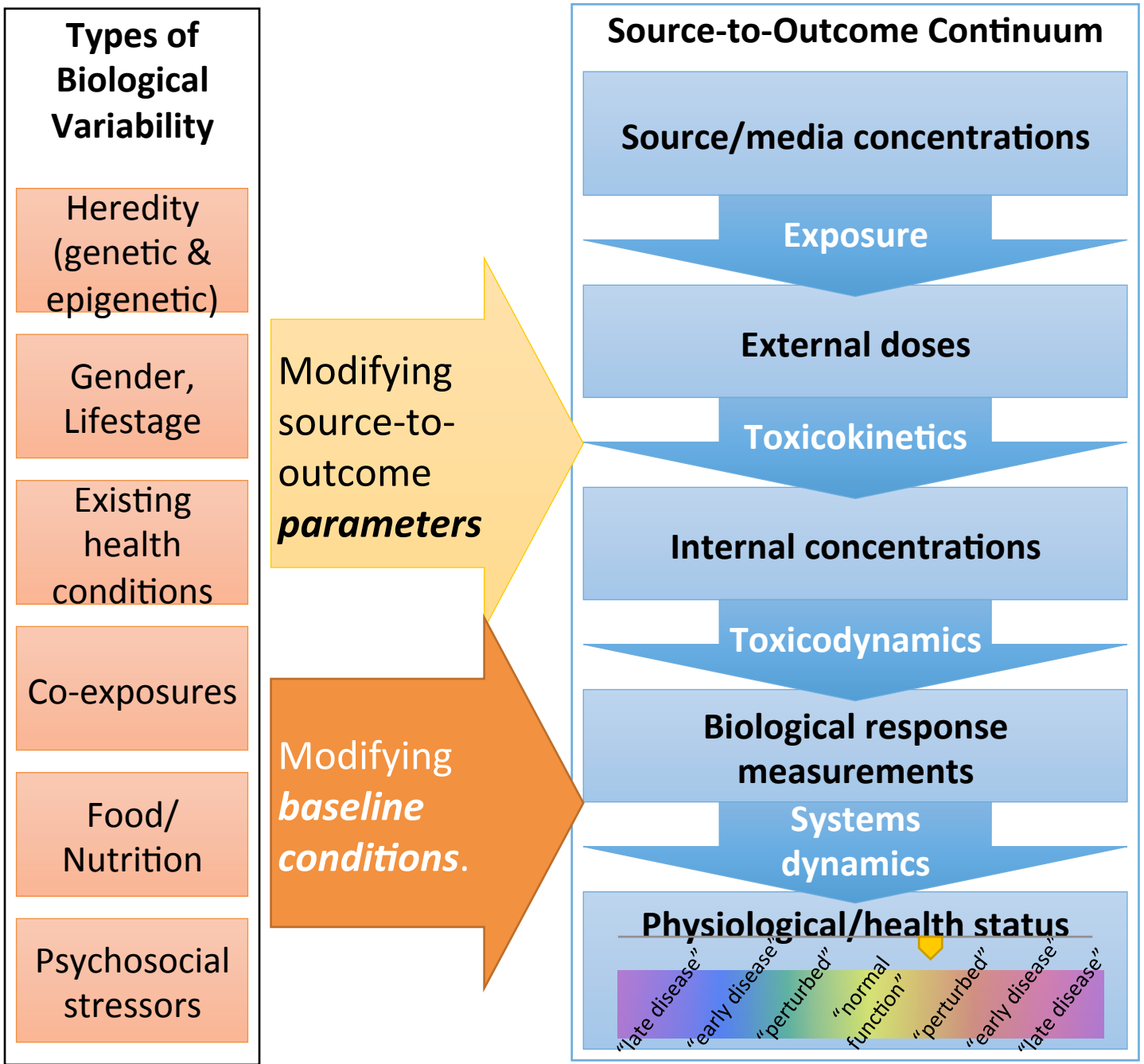






$$y=\sum k=1\uparrow n\square a\downarrow k\ x\downarrow k$$

Example Challenge: Characterizing human variability

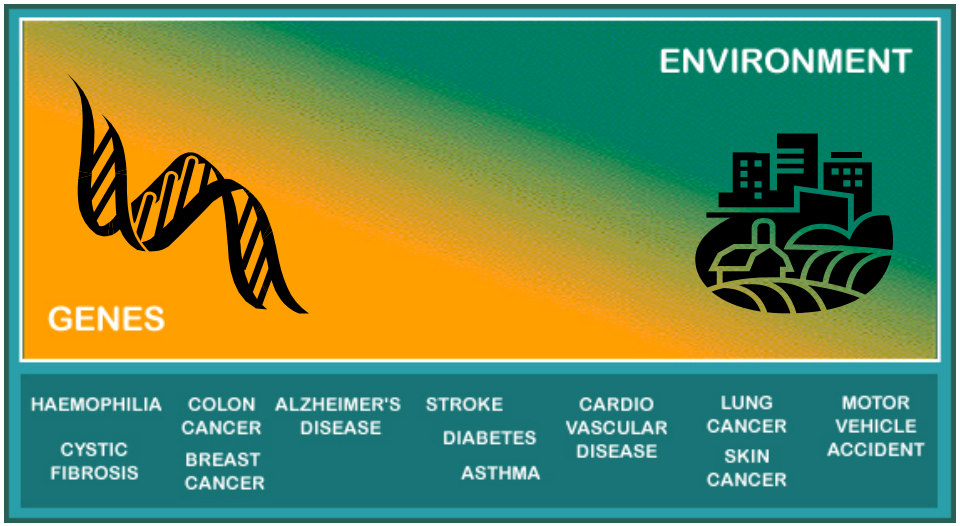


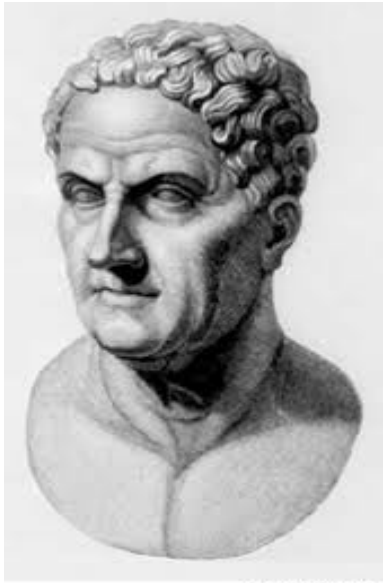
Example Challenge: Characterizing human variability

Claudius Galenus (Galen of Pergamum)

129-217 AD

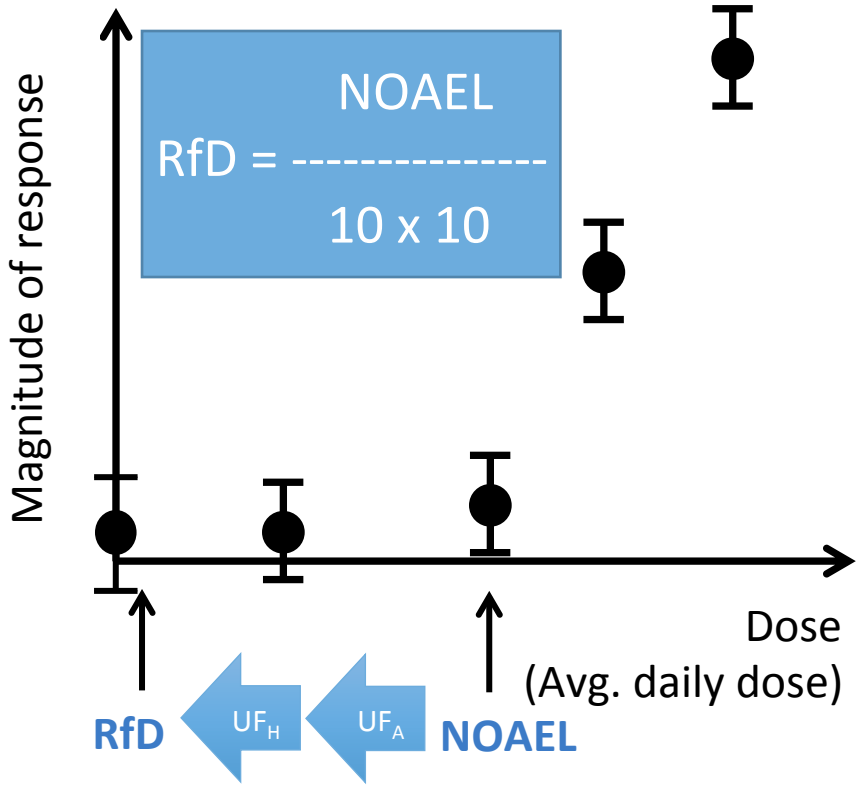
"But remember throughout that no external cause is efficient without a predisposition of the body itself. Otherwise, external causes which affect one would affect all."




Library of Congress

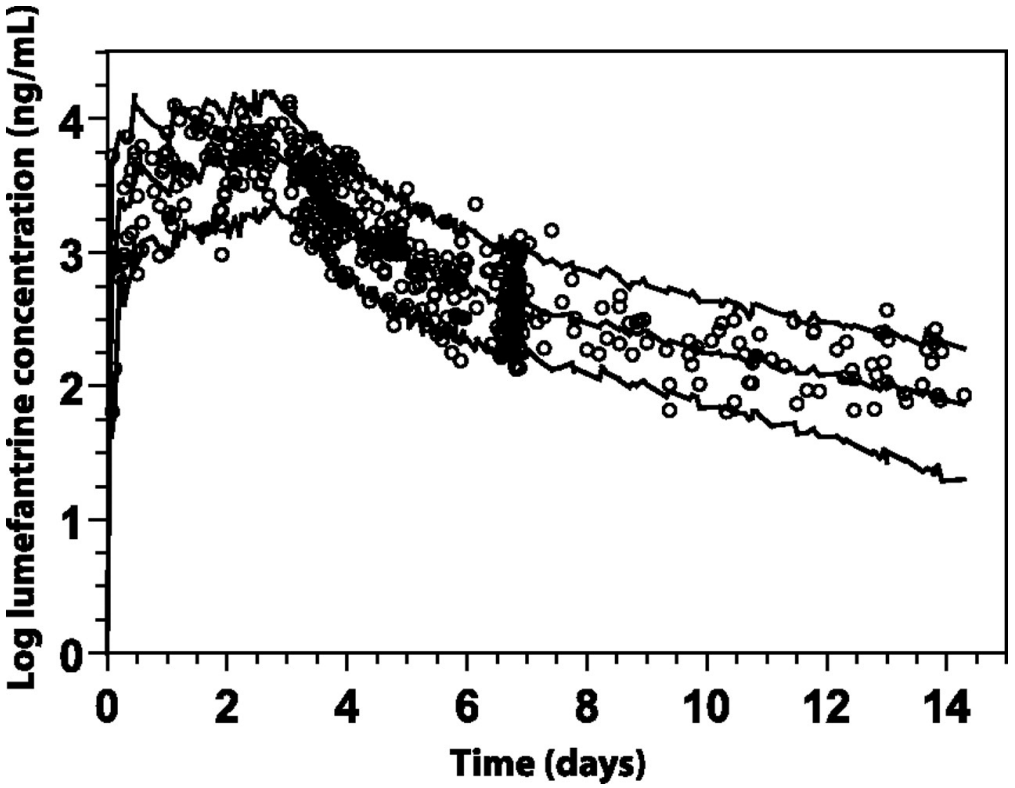
Slide courtesy of D. Threadgill

Can we do better than dividing by a factor of 10?



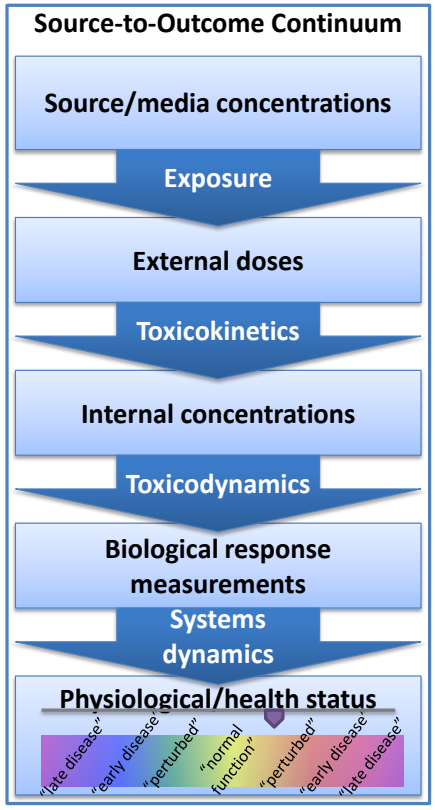
For pharmaceuticals (and some environmental chemicals), generally have direct, human empirical data

- Long history of methodological development (population PK-PD).
- Both frequentist and Bayesian statistical approaches.
- What can you do in the absence of empirical data?



Joel Tarning et al. Antimicrob. Agents Chemother. 2009;53:3837-3846

Limitations to characterizing variability for environmental chemicals



Epidemiologic & clinical studies

- Available for relatively few chemicals (~100).
- **Limited power to examine population variability/susceptibility.**
- **Generalizing from occupational/patient cohorts to the population.**

Animal bioassays

- Available for relatively few chemicals (<1000).
- Uncertain interspecies differences.
- **Homogeneous (genetics, diet, etc.) experimental animals.**

TK models

- Available for relatively few chemicals (~100 PBPK; <1000 total).
- **Few examples analyzing population variability or uncertainty.**

In vitro assays, Toxicity pathways

- Available for more chemicals (~10,000).
- Uncertain relationship to health risk.
- **Genetically homogeneous in vitro systems.**

Adverse Outcome Pathways

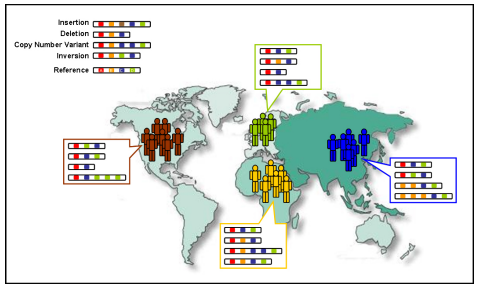
- Available for only a few endpoints (~10?).
- Qualitative, not quantitative.
- Most are artificially linear constructs.
- **Variability/susceptibility not included.**

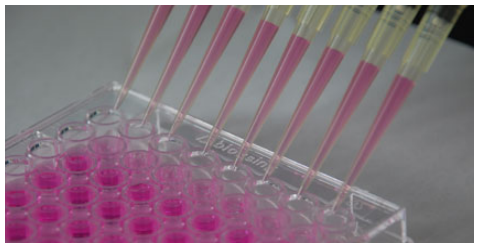
Toxicity values and risk characterization

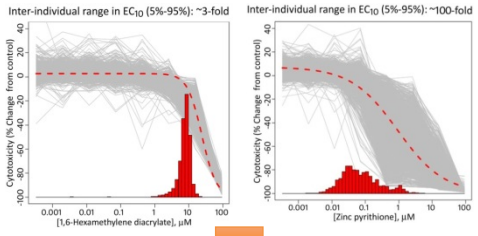
- Available for relatively few chemicals (<1000).
- **Do not adequately address uncertainty, variability, susceptibility (10-fold factor).**
- In most cases, do not explicitly estimate risk.

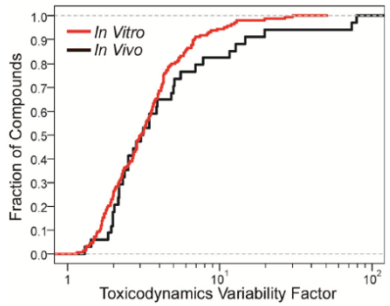
Possible approaches without direct empirical data

In vitro data

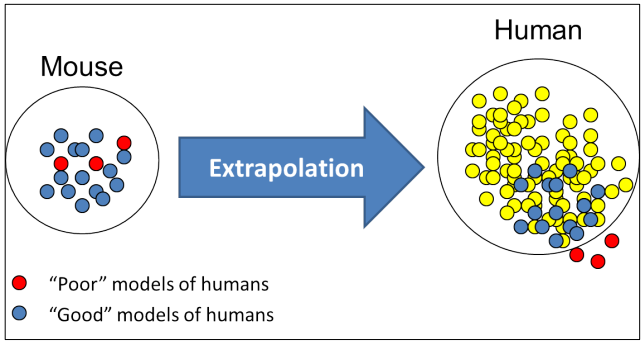


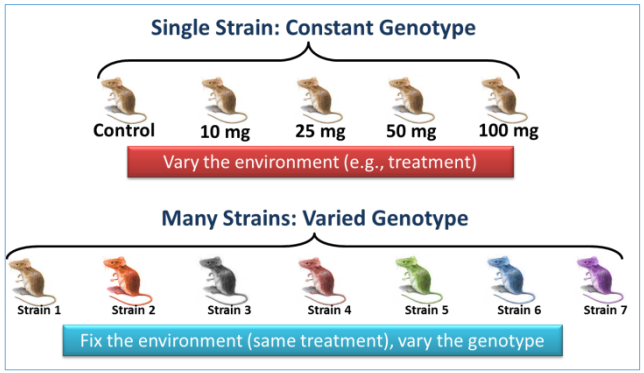


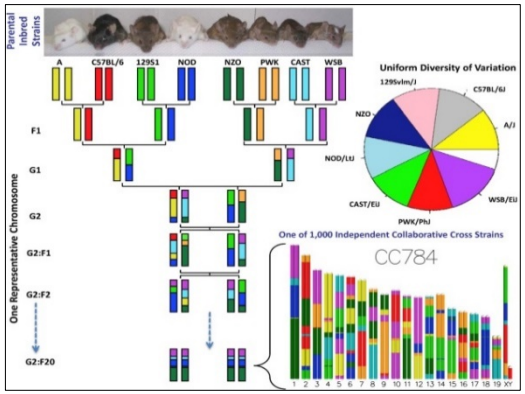




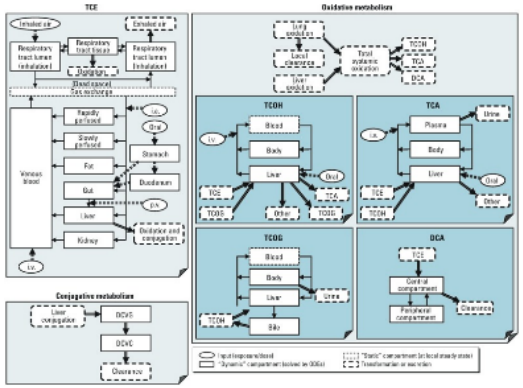
In vivo data

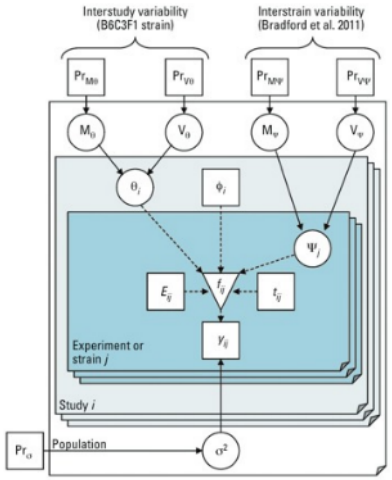






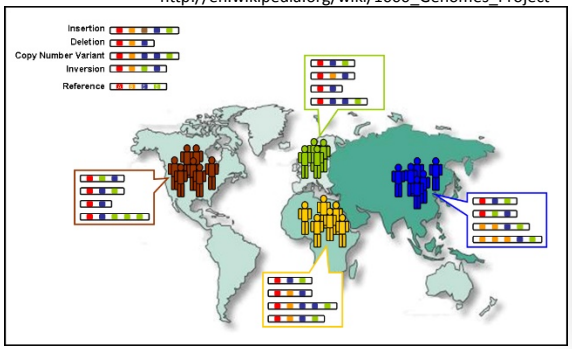
In silico methods

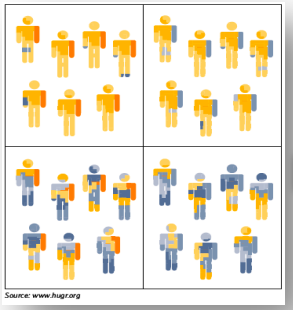




Genetically diverse human population

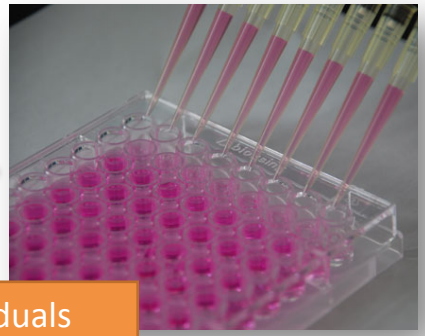
http://en.wikipedia.org/wiki/1000_Genomes_Project





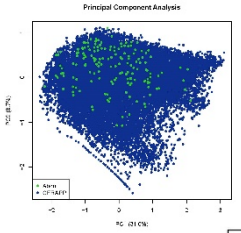
Genetically defined sample of population

High throughput *in vitro* model system



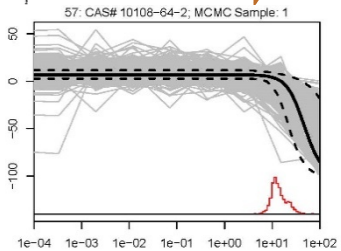
~1000 individuals cytotoxicity screening

Structurally diverse chemical population

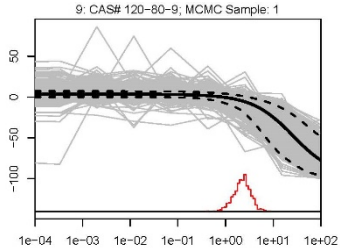


~170 compounds

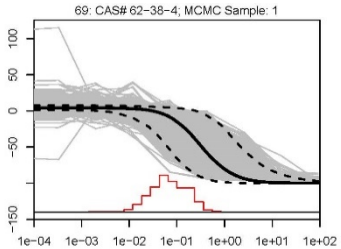
Chemical-Specific TD Variability Factor ($TDVF_{01}$):
The factor estimated to protect up to the most sensitive 1% for human toxicodynamic variability for a chemical

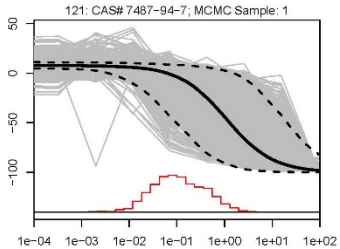


Cadmium Chloride ~2-fold



Catechol ~3-fold



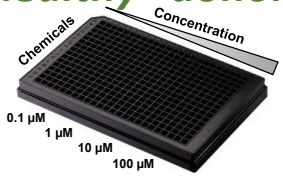


Organic and inorganic mercury compounds >8-fold

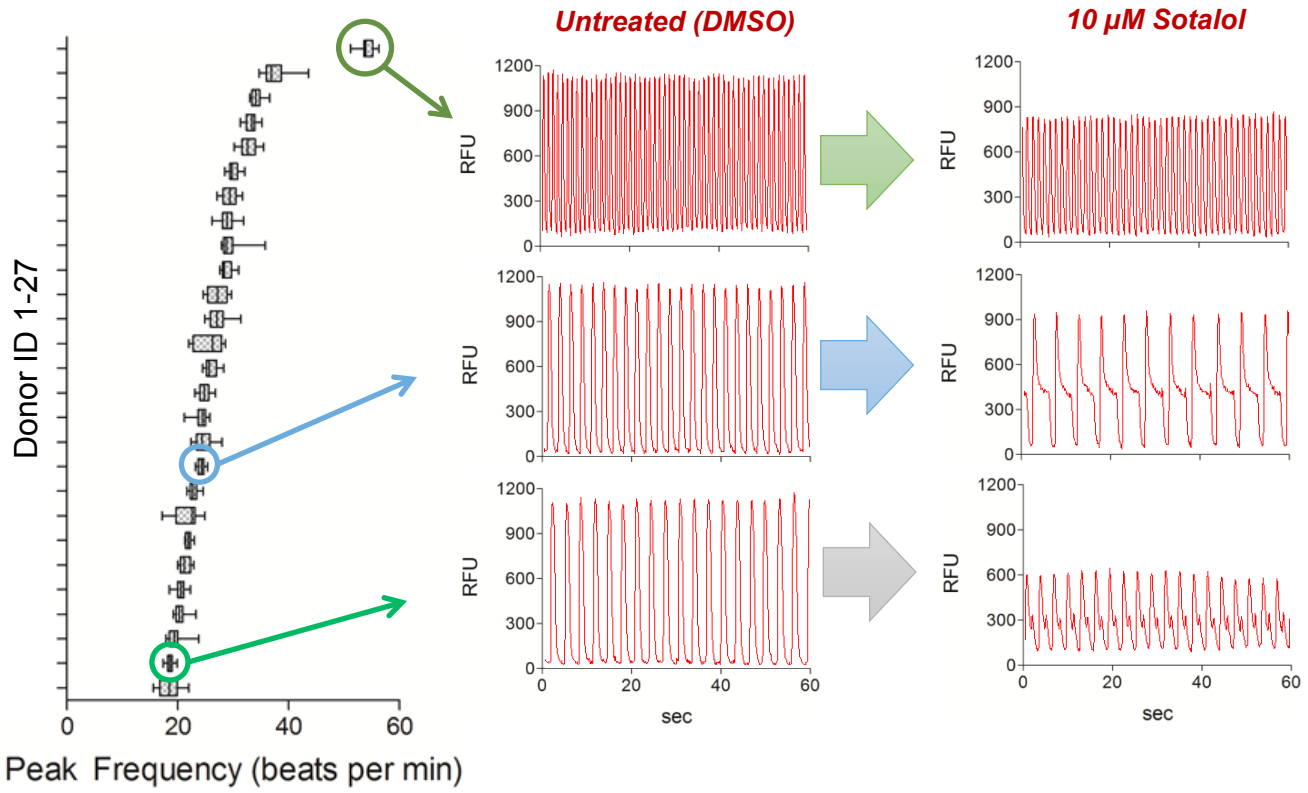
Abdo et al., 2015
Chiu et al., 2017
https://doi.org/10.14573/altex.1608251

Population Toxicodynamics for Cardiotoxicity using Cardiomyocytes

~100 individual "healthy" donors

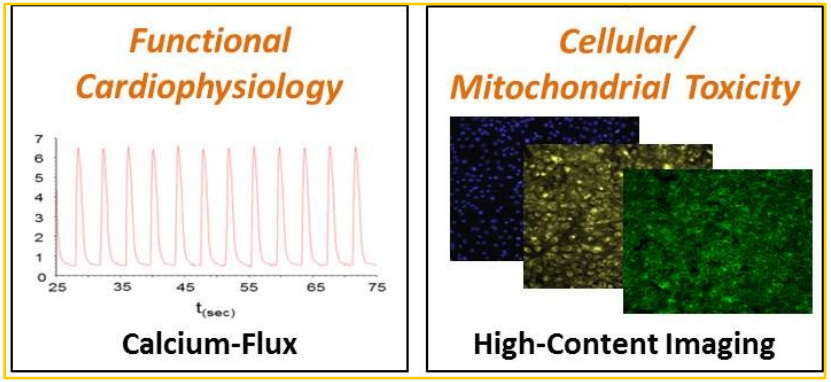


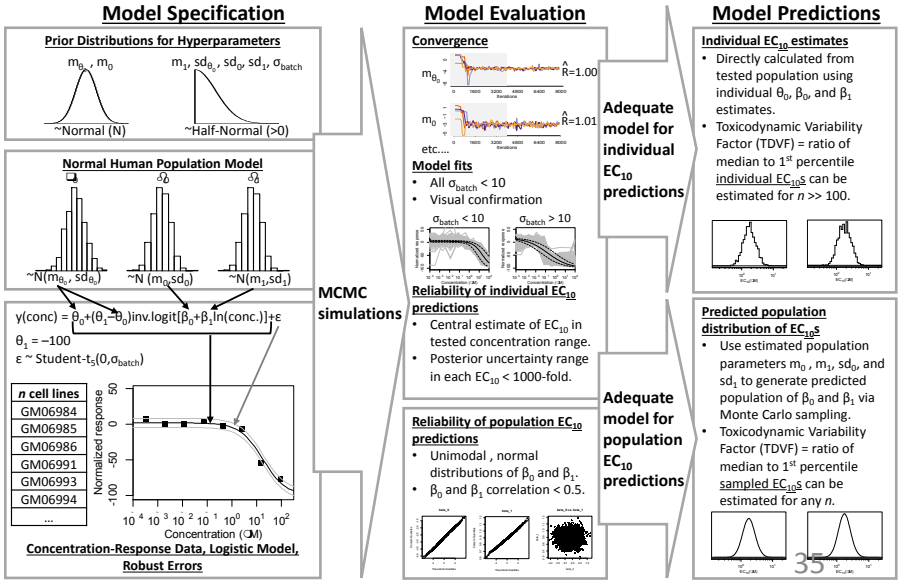
Diverse set of ~140 chemicals



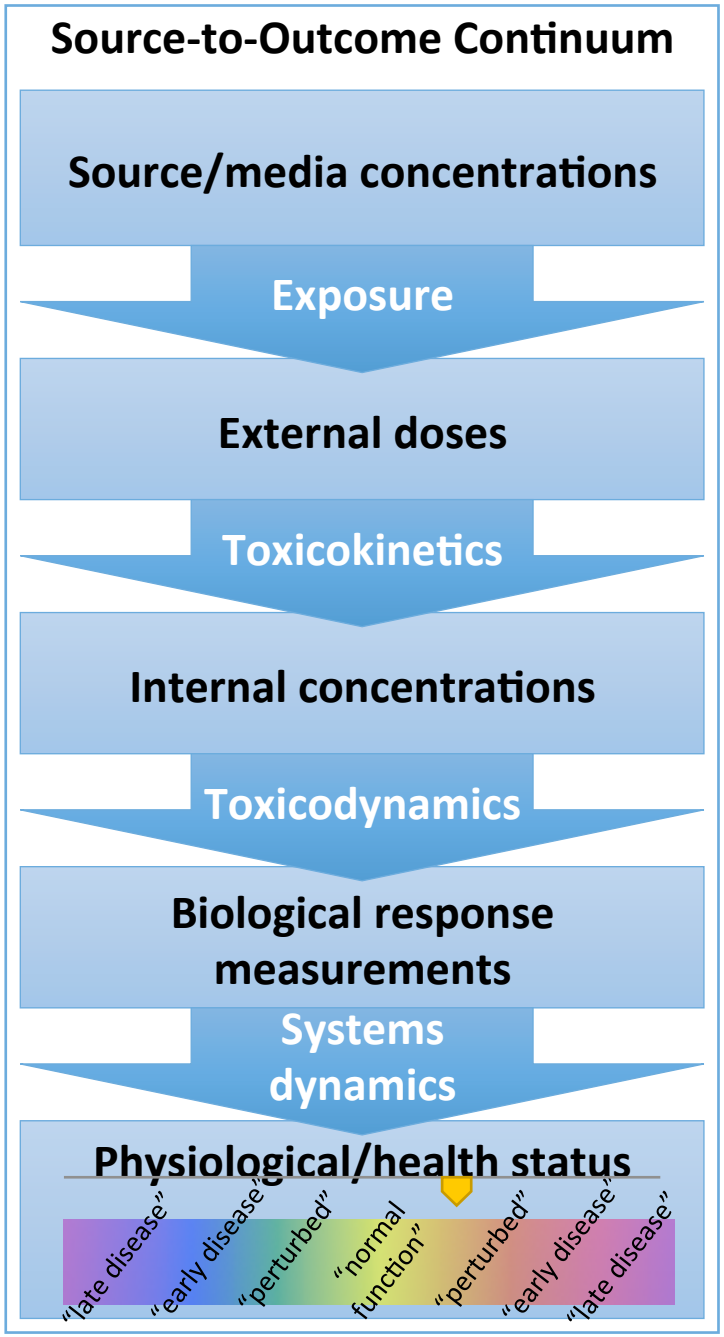
Computational demands

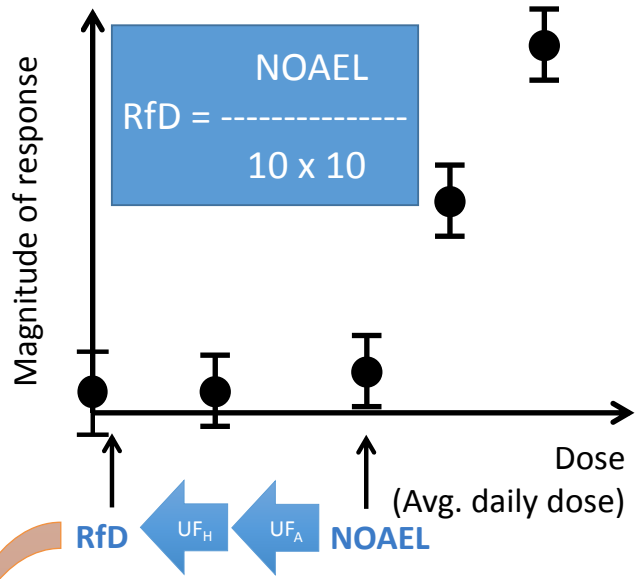
- Pre-processing multiple types of high-dimensional data
  - Imaging data
  - Time-series data
  - Genomic (gene expression) data
- Automated concentration-response modeling
- Distinguishing true population heterogeneity from random errors





Quantifying risk and uncertainty





TK models

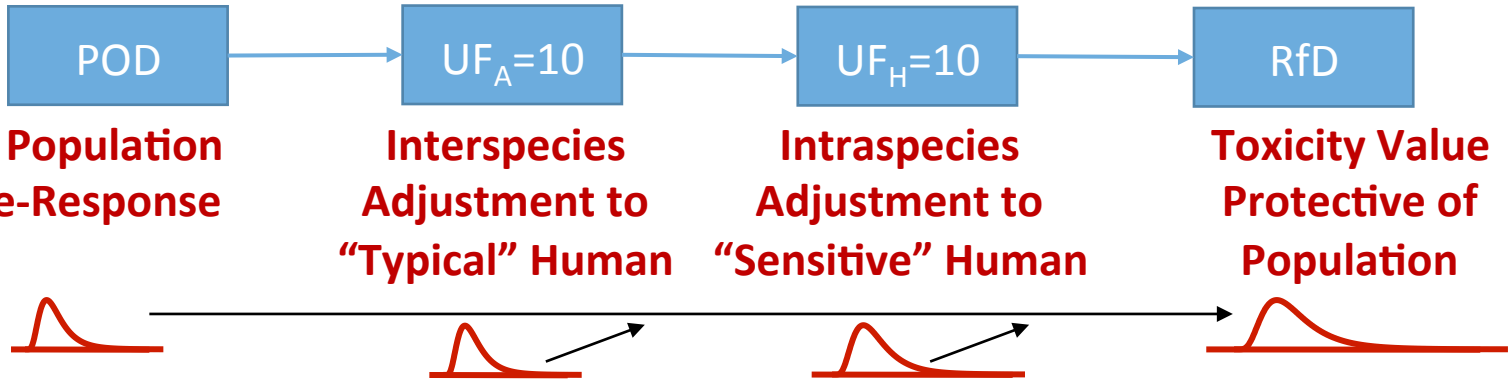
In vitro assays

??

??

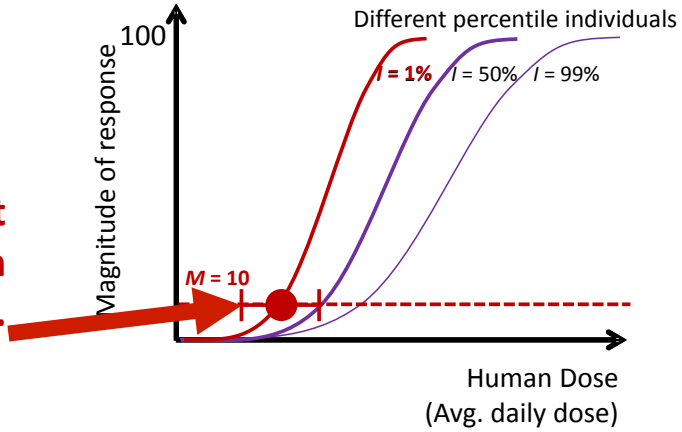
- What is the risk in terms of severity & incidence in the population?
- What are the confidence intervals?

Application of Probabilistic Approaches to Quantify Risk and Uncertainty



- Each "factor of 10" is replaced by a distribution derived from empirical data.
- Dose-response and each adjustment are combined probabilistically to derive a ***confidence interval*** that characterizes ***uncertainty***.
- Result is **Target Human Dose (HD_M^I):** human dose that at which a ***fraction I*** of the population shows an effect of ***magnitude (or severity) M*** or greater (for the critical effect considered).

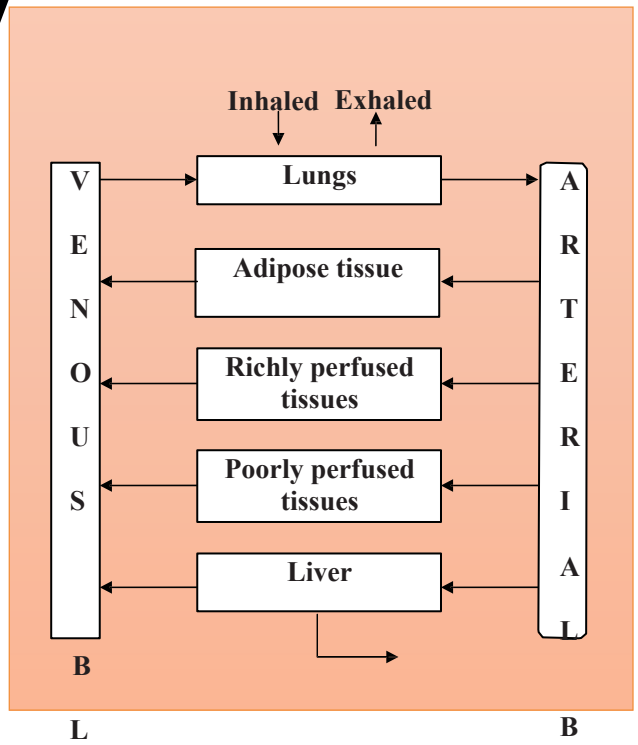
HD_M^I = Dose where at most I=5% of the population experience a M=10% effect.

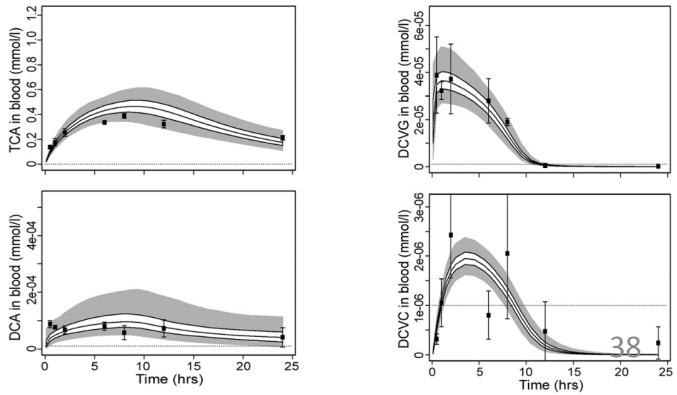


WHO (2014): Guidance on Evaluating and Expressing Uncertainty in Hazard Assessment. Harmonization Project Document 11.
Chiu WA & Slob W (2015): A Unified Probabilistic Framework for Dose-Response Assessment of Human Health Effects. EHP, DOI: 10.1289/ehp.1409385

TK models: approaches to quantify uncertainty

- Physiologically-based pharmacokinetic models are needed in the absence of empirical data
- Many parameters, each with uncertainty and population variability
- Models are not uniquely identifiable from direct observations
- Two approaches have been used:
  1. **Fit** all parameters using Bayesian approach
  2. **Fix** all but a small subset of parameters at nominal values, and fit the remaining using a frequentist approach
- #1 can be computationally prohibitive, whereas #2 can lead to biased results.



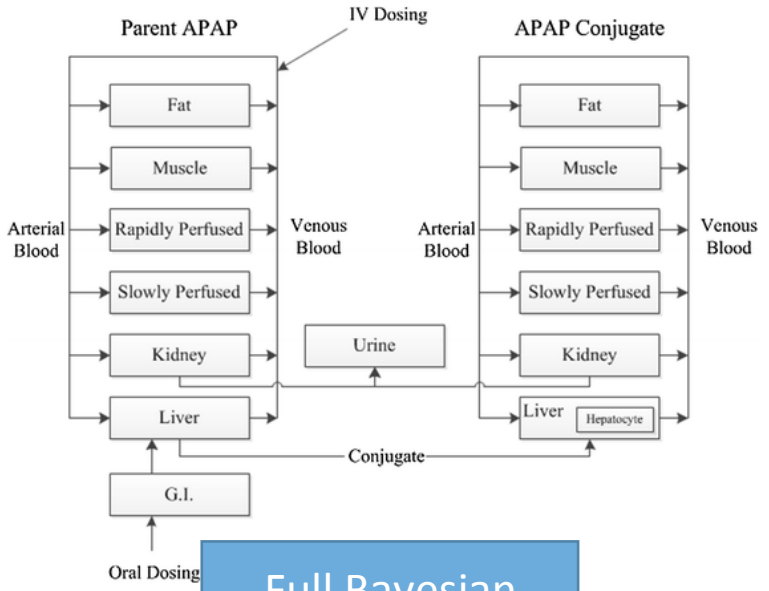


Reducing dimensionality with global sensitivity analyses

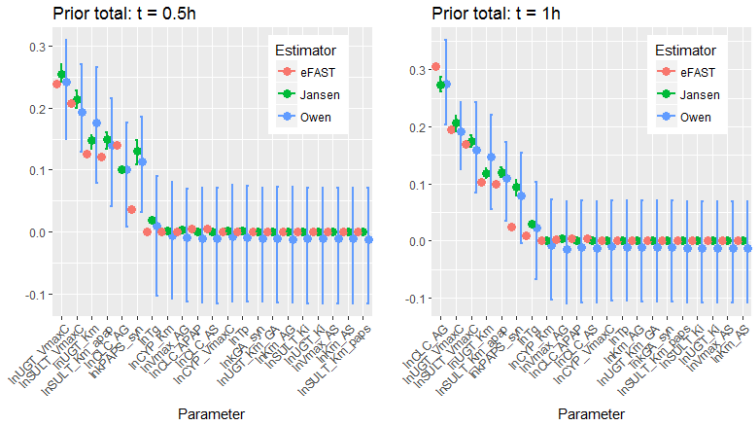
- **Hypothesis:** Can reduce dimensionality of Bayesian analysis by fixing "low sensitivity" parameters at nominal values without introducing significant bias.
- Test hypothesis by comparing reduced model results with those of full Bayesian analysis ("gold standard").

- Need **global** rather than **local** sensitivity analysis because of potential nonlinearities across parameter space
- **Sobol indices**: Reduction in output variance if the input parameter were known exactly
  - First order term measures direct effect
  - Interaction term measures effects combined with other parameters
  - Multiple algorithms for calculating indices

Preliminary results



GSA

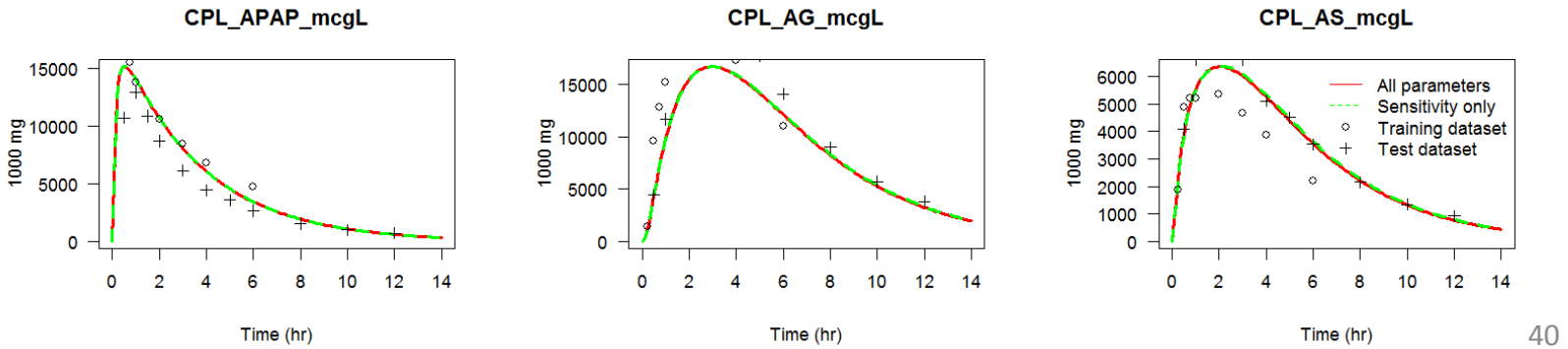


Full Bayesian Analysis

21 parameters, 19 hr simulation time

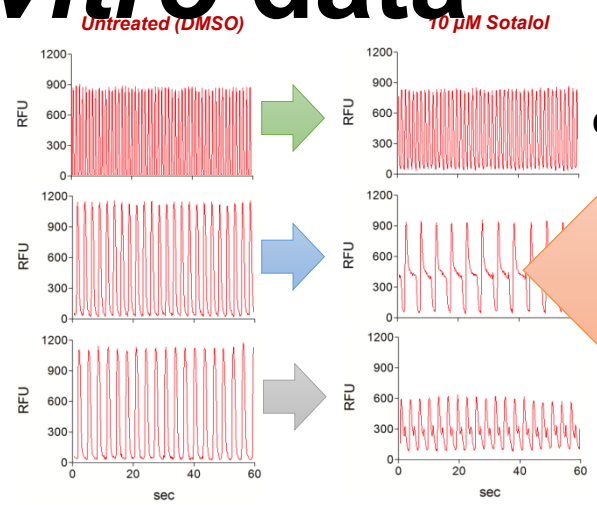
Reduced dimensionality analysis

12 parameters, 10 hr simulation time



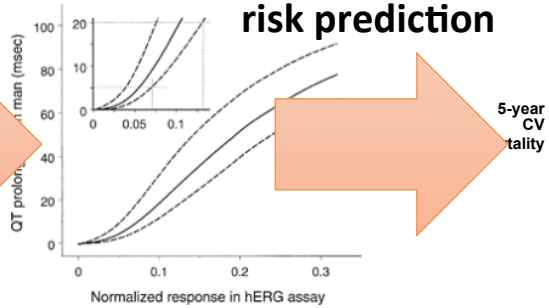
Results nearly indistinguishable

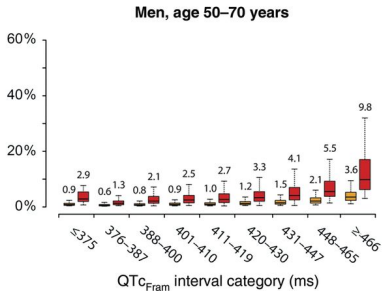
Predicting population risk from *in vitro* data



Calibration to human clinical data

Biomarker-based population risk prediction





Molecular

Chemical blocks ion channel

Chemical

Molecular level

DRUG

Source: hERGAPDbase

Cellular

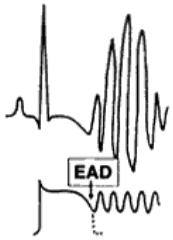
Action potential prolonged

Cellular level

Tissue

QT interval prolonged → Increased likelihood of toursades de pointes

Tissue level (Heart ECG)



Organism

Increased likelihood of myocardial infarction → Increased likelihood of death

Inter-individual variability (TK, TD)

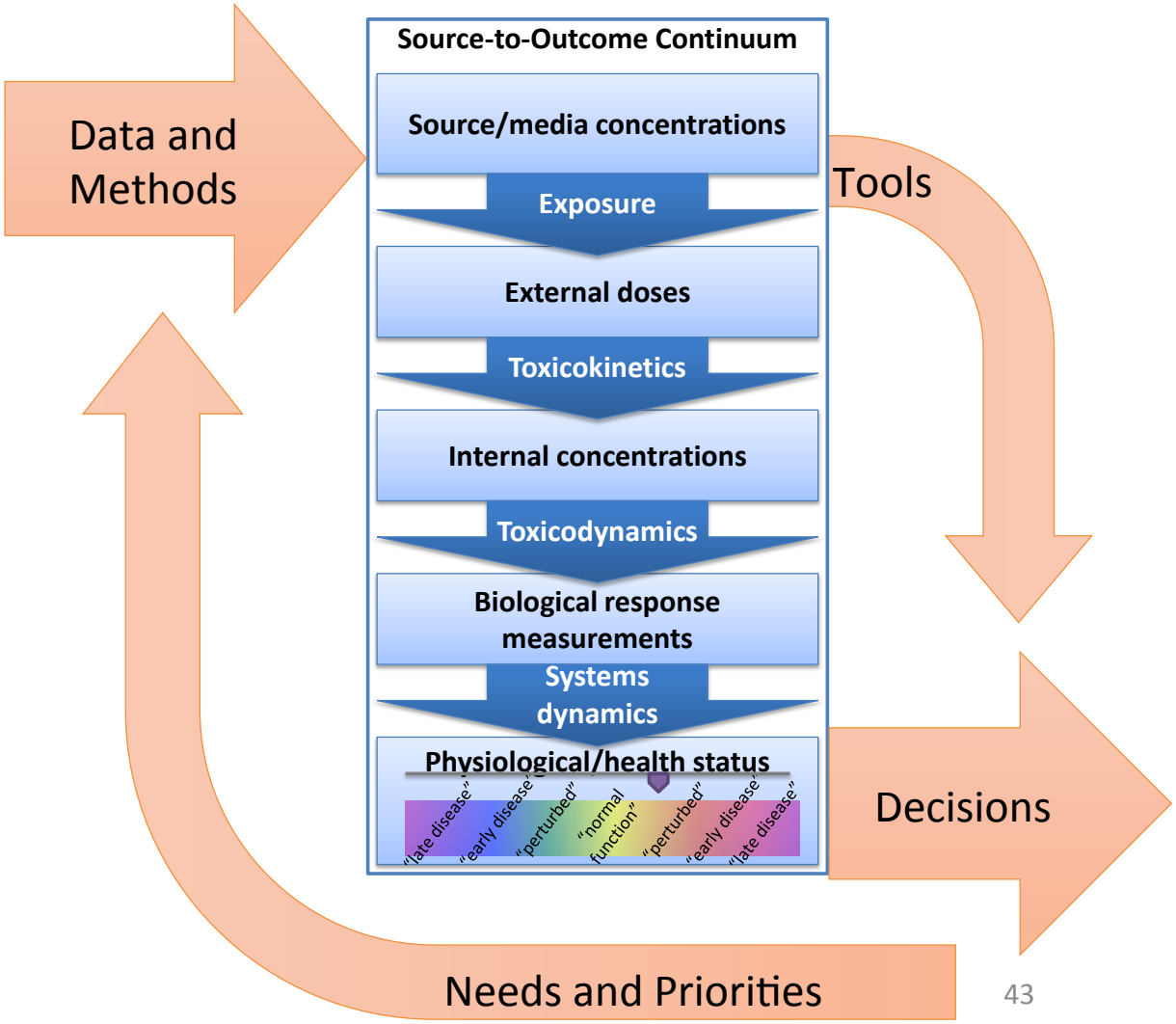
Variability and stochasticity from other stressors/risk factors

Computational demands

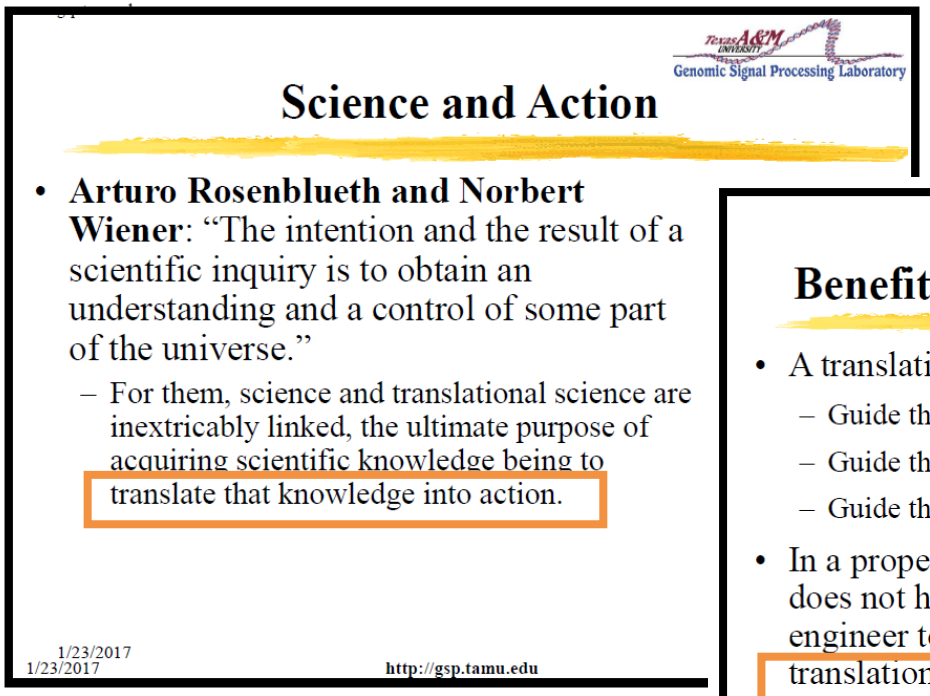
- Monte Carlo simulation
- Bayesian estimation using Markov Chain Monte Carlo
- Stiff ODE solvers
- Global sensitivity analyses

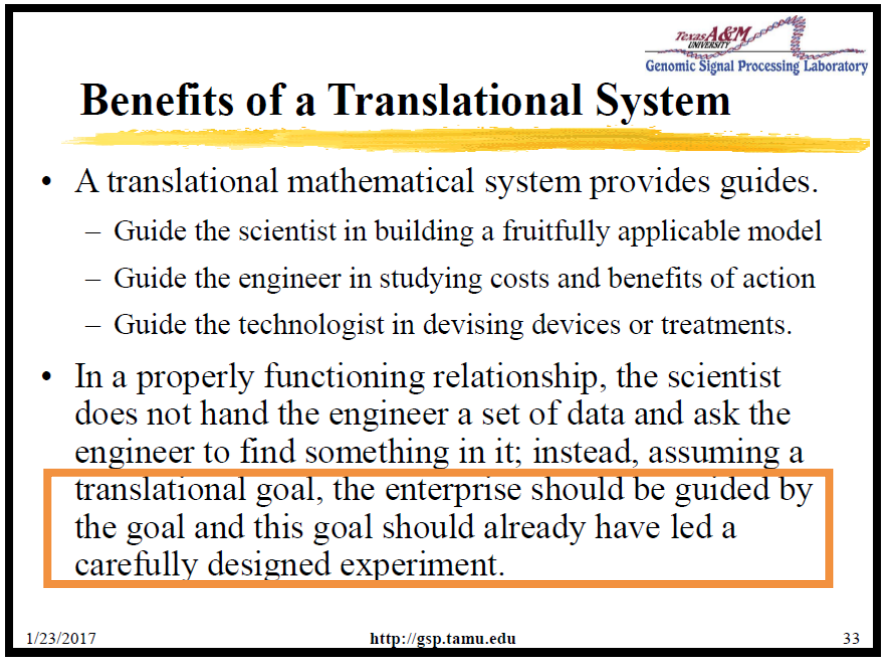
Risk assessment is an inherently ***translational*** science

- Requires integration of data from multiple sources across the source-to-outcome continuum.
- Aimed ultimately at supporting decisions, not testing hypotheses.
- Additional challenges involve moving from a researching methods to developing tools



Echoes Prof. Dougherty's seminar "Modern engineering as a translational science"…





Summary of computational demands of chemical risk assessment

- Multiple types of high-dimensional data processing
  - Multi-dimensional chemical data
  - Imaging data
  - Time-series data
  - Genomic (gene expression) data
- Multivariate data integration to define "similarity"
- Deconvolution to construct "mixture analogues" using reference chemicals
- Automated concentration-response modeling
- Distinguishing true population heterogeneity from random errors
- For quantifying risk, methods for classification are not enough – need a numerical prediction.
- Monte Carlo simulation
- Bayesian estimation using Markov Chain Monte Carlo
- Stiff ODE solvers
- Global sensitivity analyses

Opportunities for students/postdocs: Chemical Risk Assessment suffers from lack of expertise in both <u>developing</u> and <u>applying</u> computational methods.