Quantitative Evidence Integration to Facilitate the Use of Tox21 Data for Rapid Risk Screening

Erin E. Yost, PhD
Ingrid L. Druwe, PhD
Oak Ridge Institute for Science and Education (ORISE)
Postdoctoral Fellows

U.S. EPA National Center for Environmental Assessment
Research Triangle Park, NC

The views expressed are those of the presenters and do not necessarily represent the views of the US EPA. Mention of trade names or products does not constitute endorsement or recommendation for use.
Too many chemicals, not enough risk assessments

Pathways to exposure from contamination

Image from California Environmental Health Investigations Branch, adapted from CDC
• Tox21: *In vitro* concentration-response data has been generated for thousands of chemicals.

• These molecular data can be anchored to human health outcomes using the Adverse Outcome Pathway (AOP) framework:
• Use Tox21 data to inform the risk assessment of data-poor chemicals?

• In vitro potency -> risk screening values -> chemical prioritization

![Diagram](Molecular Cellular Tissue Organ Individual Population)

Tox21 data

Human health outcomes

Adapted from Noffisat Oki
Addressing uncertainty in chemical potencies
How can this HTS data be quantitatively integrated to inform the probabilistic identification of hazard and estimation of risk specific concentration/doses?

1. Approaches for data integration:
   a) Meta-regression across orthogonal assays (Erin Yost)
   b) Bayesian Methods (Ingrid Druwe)

2. Case study: Benzo[k]fluoranthene
Meta-regression across orthogonal assays

Dimerization assays

Ligand binding assays

DNA binding assays

Transactivation assays

Nucleus

Cell

Transcription
Identifying orthogonal assays for ERα activation

- Identified two orthogonal assays (from PubChem database) that had been used to screen large chemical inventories:
  - Invitrogen ERα-UAS-bla GripTite™ assay (AID 743078)
  - BG1-Luc-4E2 assay (AID 743079)
How reproducible are chemical potencies within each assay?

• Looked at active chemicals in each assay
  • ERα-UAS-bla GripTite™: 3145 chemicals
  • BG1-Luc-4E2: 3441 chemicals

• Calculated the mean potency (EC50) for each chemical across replicates

• Determined how much each replicate differed from the mean (residuals)
  • Residual close to zero -> assay is reproducible
How reproducible are chemical potencies within each assay?

- **Result:** Residuals close to zero -> Chemical potencies are generally reproducible within each assay
How reproducible are chemical potencies *across orthogonal assays*?

- Composited data for chemicals that were active in both assays
  - 2028 chemicals total
- Calculated the mean potency (EC50) for each chemical
- Calculated residuals
  - Residual close to zero -> assay is reproducible
How reproducible are chemical potencies across orthogonal assays?

Estimated mean residual: 0.0011 μM
(95% CI: -0.025 to 0.027 μM)

• Result: Residuals close to zero -> Chemical potencies are generally reproducible across these assays
Integrating data across multiple orthogonal assays

Activity of Bisphenol-A in ER-alpha-UAS-bla-GripTite Assay

Response Ratio (%)

Activity of Bisphenol-A in BG1-Luc-4E2 Assay

Response (%)
Integrating data across multiple orthogonal assays

Activity of Bisphenol-A in BG1-Luc-4E2 Assay

Activity of Bisphenol-A in ER-alpha-UAS-bla-GripTite Assay

Median Concentration-Response for Bisphenol-A

POD = 0.034 μM
Threshold
Integrating data across multiple orthogonal assays

Activity of Bisphenol-A in BG1-Luc-4E2 Assay

Activity of Bisphenol-A in ER-alpha-UAS-bla-GripTite Assay

Max. Activity

1:1000 risk-specific conc. = 0.031 μM

1:1000 risk factor
Conclusions:
Meta-regression across orthogonal assays

• Simple and transparent method for rapid data integration

• Orthogonal assays provide weight of evidence for chemical perturbation of key events, and possibly the associated active concentrations

• **Ultimate goal:** Use bootstrap natural spline-based metaregression to rapidly develop risk screening values for data-poor chemicals
How can this HTS data be quantitatively integrated to inform the probabilistic identification of hazard and estimation of risk specific concentration/doses?

1. Approaches for data integration:
   a) Meta-regression across orthogonal assays (Erin Yost)
   b) Bayesian Methods (Ingrid Druwe)

2. Case study: Benzo[k]fluoranthene
What is Bayesian Analysis?

Bayesian Analysis is a statistical method which aims to estimate parameters of an underlying distribution based on the observed distribution.
Example: A woman is being screened for breast cancer, as is recommended for women over 40. She does not have a family history of breast cancer and she was uncertain about her status in regards to breast cancer, and wanted to learn more about it.

<table>
<thead>
<tr>
<th>Model</th>
<th>Prior Probability</th>
<th>Likelihood for M+</th>
<th>Prior X Likelihood</th>
<th>Posterior Probabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast Cancer</td>
<td>0.0045</td>
<td>0.724</td>
<td>0.0033</td>
<td>0.107</td>
</tr>
<tr>
<td>No Breast Cancer</td>
<td>0.9955</td>
<td>0.027</td>
<td>0.0273</td>
<td>0.893</td>
</tr>
</tbody>
</table>

Likelihood:  \[ L = p^s (1 - p)^f \]

Prior distribution:  \[ pr = p^a (1 - p)^b \]
Example (Continued)

<table>
<thead>
<tr>
<th>Model</th>
<th>Prior Probability</th>
<th>Likelihood for S-</th>
<th>Prior X Likelihood</th>
<th>Posterior Probabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast Cancer</td>
<td>0.107</td>
<td>0.11</td>
<td>0.012</td>
<td>0.014</td>
</tr>
<tr>
<td>No Breast Cancer</td>
<td>0.893</td>
<td>0.94</td>
<td>0.839</td>
<td>0.986</td>
</tr>
</tbody>
</table>

Posterior distribution:  

Eq 3: \( p_{S_1} = L \times pr \)  
\[ = p^s (1 - p)^f \times p^a (1 - p)^b \]  
\[ = p^{s+a} (1 - p)^{f+b} \]
Fig 2. Model of Approach

1. Select chemical in PubChem.
2. Use data to build Bayesian model.
3. Integrate data into knowledge base.
4. Generate likelihood of adversity at each dose.
5. Repeat steps 1 & 2 and integrate new data to update model.

• All statistical analyses and modeling were performed in R (v3.1.1).
About the data

Chemical: Bis-Phenol A (Chemical ID: 6623)
Luciferase ligand binding assay using, BG1Luc4E2 cell line (AID: 743079)

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Replications</th>
<th>Concentration Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dataset A</td>
<td>46 reps</td>
<td>(1.23 nM to 94.36 µM)</td>
</tr>
<tr>
<td>Dataset B</td>
<td>3 reps</td>
<td>(1.23 nM to 70.35 µM)</td>
</tr>
</tbody>
</table>

BG1-Luc-4E2 assay
Bootstrap natural spline-based meta-regression plot of data based on Dataset A
Prior distribution: Eq 2: \[ pr = p^a (1 - p)^b = 1 \]
Posterior distribution: Eq 3: \( \rho_{S_1} = L \times \rho_r \)
\[= \rho^s (1 - \rho)^f \times \rho^a (1 - \rho)^b \]
\[= \rho^{s+a} (1 - \rho)^{f+b} \]
Conclusions

• Bayesian Method yields concentration-specific probability of hazard.
• Quantitative data integration yields a visual representation of assay reproducibility
• The Bayes enables quantitative data integration with decreased subjectivity.
Future Directions

• Calculate risk specific concentrations from concentration response data.
• Apply the Bayesian method to data from other chemicals in the PubChem database.
• Incorporate the Bayesian method to our semi-automated workflow for HTS data analysis.
How can this HTS data be quantitatively integrated to inform the probabilistic identification of hazard and estimation of risk specific concentration/doses?

1. Approaches for data integration:
   a) Meta-regression across orthogonal assays (Erin Yost)
   b) Bayesian Methods (Ingrid Druwe)

2. Case study: Benzo[k]fluoranthene
Case Study: Benzo[k]fluoranthene

**Goal:** Apply evidence integration from HTS assays in order to understand the health hazards of benzo[k]fluoranthene (B[k]F).

Lyle D. Burgoon, Ingrid Druwe, Kyle Painter and Erin Yost. **2015.** *Using In Vitro High Throughput Screening Data for Predicting Benzo[k]Fluoranthene Human Health Hazards.* [Currently under review.]
Analysis Plan

Focus on assays that represent **sufficient key events** (events that are sufficient to infer an adverse outcome).

For B[k]F, identified two assays that represent **sufficient key events**:

1. p53 activation
2. HSD17B4 enzyme inhibition
Sufficient Key Event #1: p53 activation (tumorigenesis & cancer)
Concentration-response analysis: B[k]F activity in p53 activation assay

- 5 replicates available total, across largely similar doses
- PubChem BioAssay IDs:
  - AID 651631 (3 replicates)
  - AID 651743 (2 replicates)
Concentration-response analysis: B[k]F activity in p53 activation assay

- 5 replicates available total, across largely similar doses
- PubChem BioAssay IDs:
  - AID 651631 (3 replicates)
  - AID 651743 (2 replicates)
Sufficient Key Event #2: HSD17B4 inhibition (steatosis)
Concentration-response analysis: B[k]F activity in HSD17B4 inhibition assay

- Only one replicate available (PubChem BioAssay AID 893)
Summary: B[k]F toxicity

• p53 transactivation assays indicate that B[k]F likely causes DNA damage
  • Integrated evidence from 5 replicate assays
  • $\text{POD}_{\text{DNA damage}}: 0.751\mu\text{M}$
  • Estimated 1:1,000 risk-specific conc: 0.29μM

• HSD17B4 enzyme inhibition assay indicates that B[k]F may also cause steatosis
  • Only a single replicate
  • Due to lack of repeat evidence, did not calculate POD or risk-specific concentration
Summary: Using Tox21 data to inform rapid risk screening
Acknowledgements

Thank you to:
• Lyle Burgoon, PhD- US Army ERDC
• Kyle Painter, MS- ORISE US EPA
• John Vandenberg, PhD- US EPA
• Reeder Sams, PhD- US EPA
• Ila Cote, PhD- US EPA
• Michelle Angrish, PhD- US EPA