The Role of Holistic Assay Platforms and Computational Methods to Prioritize Pharmaceutical Candidates

Russ Naven
Principal Scientist, Compound Safety Prediction,
Pfizer Inc, Eastern Point Road, Groton CT 06340
russell.naven@pfizer.com
• Safety assessment in drug discovery
• In vivo toxicity data collation
• Development of Compound Safety Evaluator
  - Physicochemical properties
  - In vitro assays
  - MPO approach
• External evaluation
• Improvements
• Summary
Safety Assessment in Drug Development

Hit identification ➔ Lead identification ➔ Lead optimisation ➔ Candidate drug ➔ Clinical trials ➔ Market

Target Safety ➔ In silico/in vitro Toxicity screening ➔ Ames Mutagenicity ➔ in vivo tox studies

Confidential and Internal use only – Wyeth is now a part of Pfizer. The merger of local Wyeth and Pfizer entities may be pending in various jurisdictions and integration is subject to completion of various local legal and regulatory obligations. Accordingly, changes will not become effective outside of the U.S. and Puerto Rico until the relevant in-country go-live date. Until that time, Pfizer and Wyeth must continue to operate as two separate companies. All content in this message is subject to works council and/or union consultations and other legal requirement.
Safety Assessment in Drug Development

- Hit identification
- Lead identification
- Lead optimisation
- Candidate drug
- Clinical trials
- Market

- Target Safety
- In silico/in vitro Toxicity screening
- Ames Mutagenicity
- in vivo tox studies

Safety of Impurities
Safety of Degradants
Safety of Extractables
Leachables
Sequential Investment in Tox Screening

**Tier 0**

- In silico support of compound design **“on paper”**
  - Models automatically triggered when compound scheduled for synthesis
  - Includes physchem property space, target models, structural alerts
  - Specific in silico models for basic tox endpoints (ion channels, cytotox)

**Tier 1**

- Rapid chemotype screening for broad mechanisms of toxicity
  - Used to steer chemistry into more productive chemical space
  - Requires high throughput assays - cytotox, mitochondrial dysfunction, 2ndary pharmacology assays
  - Probabilistic assessment of low dose toxicity (Compound Safety Evaluator)

**Tier 2**

- Deeper dive on toxicity mechanisms.
  - Reserved for favored chemical series
  - High content mechanistic screens, full pathway coverage (e.g. lipid metabolism, steroidogenesis etc), reporter panels (e.g. stress response, ER stress, mechanisms of mitochondrial dysfunction etc)

**Tier 3**

- Project-specific SAR mechanistic work
  - Resource intensive – identification/mitigation of toxicophore
Safety Assessment in Drug Development

- Hit identification
- Lead identification
- Lead optimisation
- Candidate drug
- Clinical trials
- Market

Target Safety
In silico/in vitro assessment
in vivo tox studies
Knowledge and Data Collation

Objective: Centralize and standardize in-house & external compound-specific toxicity data and annotations, identify data gaps & backfill data (on-going effort)

Acute In vivo Toxicology Studies list: ~1100 compounds
- Detailed annotations of findings at dose group level w/ TK data in most cases

Attrited Compound List: ~800 compounds
- Internally developed candidates, high level annotations on safety issues

Compound Tox List: ~1600 compounds (includes ~650 with organ tox)
- Marketed, withdrawn & prototypic compounds with organ toxicity classes
- Covers hepatotox, cardiotox, nephrotox

Chemogenomics list: ~3000 compounds
- Compounds ‘specific’ for a particular target, no annotations on in vivo outcome

Toxicity prediction progress is at the mercy of annotation quality
Need to put all 1100+ compounds on a common scale

- Issues: complicated by species, strain, dose/exposure, duration
- Severity of clinical signs, serum chemistry, histopathology etc.

**Solution**: Are there any clinical findings at $C_{\text{max}}$ threshold of 10uM?
Initial analyses can only be performed on ~40% of cmpds!
Origins of In vivo Toxicity

Primary pharmacology

Chemical structure

Adverse safety profiles

Off-target/secondary pharmacology

Physicochemical Properties

- Lipophilic basic compounds at risk of:
  - Phospholipidosis
  - Hepatotoxicity
  - QT interval prolongation

Chemical structure: Clozapine

Ariflo

CYP inhibition
Reactive metabolites
Structural Alerts
Physicochemical Properties

- Chance of in vivo toxicity is linked to physicochemical property space:

<table>
<thead>
<tr>
<th>Total drug</th>
<th>TPSA&gt;75</th>
<th>TPSA&lt;75</th>
</tr>
</thead>
<tbody>
<tr>
<td>ClogP&lt;3</td>
<td>0.39 (57)</td>
<td>1.08 (27)</td>
</tr>
<tr>
<td>ClogP&gt;3</td>
<td>0.41 (38)</td>
<td>2.4 (85)</td>
</tr>
</tbody>
</table>

A compound that flags both properties is more than six times as likely to cause findings in a IVT study at $C_{\text{max}}<10\mu\text{M}$ than a compound that does not flag in either of these properties.

What biological endpoints or assays would improve this \textit{in silico} knowledge?
• What in vitro assays can predict in vivo toxicity?

- Structural alerts
- Physico-chemical Properties
- Secondary pharmacology
- Polypharmacology (Cerep profiling)
- Fundamental toxicity assays
- General toxicity assays
Polypharmacology

- Collected in-cerebro knowledge of endpoints with known risks issues
- Selected 15 Cerep Assays - The ‘Promiscuity Panel’
- Covers GPCRs, ion-channels, transporters, PDE
- Provides a lower cost, ‘quick look’ at promiscuity

δ Average inhibition of the 15 targets generally correlates with wider promiscuity

Analysis of ~200 cpds sent to CEREP full panel (Jan-May 2010)
Sorted by Average %I across the 15 ‘Promiscuity Panel’ Targets. Each row is a compound.

The most promiscuous compounds across 15 targets carry on hitting multiple targets in the rest of the Full panel.

CEREP data: Colour-code

- >85%I
- 50-85%I
- 30-50%I
- <30%I

In contrast, the compounds with low average %I in the P-Panel are generally cleaner across the rest of the CERE panel.

~200 cpds sent to CERE panel (Jan-May 2010)
### Polypharmacology hits vs. in vivo tox

<table>
<thead>
<tr>
<th>Cumulative #Hits</th>
<th>Toxic @10uM</th>
<th>Clean @10uM</th>
<th>Relative Odds Ratio (toxic vs. clean)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>42</td>
<td>41</td>
<td></td>
</tr>
<tr>
<td>&gt;0</td>
<td>97</td>
<td>27</td>
<td>3.5</td>
</tr>
<tr>
<td>&gt;1</td>
<td>73</td>
<td>10</td>
<td>6.4</td>
</tr>
<tr>
<td>&gt;2</td>
<td>56</td>
<td>8</td>
<td>5.1</td>
</tr>
<tr>
<td>&gt;3</td>
<td>51</td>
<td>5</td>
<td>7.3</td>
</tr>
<tr>
<td>&gt;4</td>
<td>46</td>
<td>3</td>
<td>10.7</td>
</tr>
<tr>
<td>&gt;5</td>
<td>38</td>
<td>1</td>
<td>25.2</td>
</tr>
<tr>
<td>&gt;6</td>
<td>33</td>
<td>1</td>
<td>20.9</td>
</tr>
<tr>
<td>&gt;7</td>
<td>21</td>
<td>1</td>
<td>11.9</td>
</tr>
<tr>
<td>&gt;8</td>
<td>13</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>&gt;9</td>
<td>10</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>&gt;10</td>
<td>7</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>&gt;11</td>
<td>4</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>&gt;12</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>&gt;13</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>&gt;14</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

A compound hitting just **2 or more** targets in the promiscuity panel is ~**5 times** more likely to cause toxicity @10uM.
• What in vitro assays can predict in vivo toxicity?

- **Secondary Pharmacology**
  - Polypharmacology (Cerep profiling)
  - Fundamental toxicity assays
  - General toxicity assays

- **Physico-chemical Properties**
- **Structural alerts**
• Assay predictivity against in vivo toxicity at 10uM:

**IDEAL SITUATION:**

![Graph showing the distribution of cytotoxicity levels for non-toxic, unknowns, and toxic categories.](image-url)
Assessing Assay Utility

- Assay predictivity against in vivo toxicity at 10uM:

**IDEAL SITUATION:**

- Non-toxic unknowns
- Toxic for other reasons

![Graph showing cytotoxicity (µM) against non-toxic, unknowns, and toxic categories.](image-url)
Assessing Assay Utility

- Assay predictivity against in vivo toxicity at 10uM:

**IDEAL SITUATION:**

- Elimination
- Detoxification
- Metabolism

Cytotoxicity (uM)

- Non-toxic
- Unknowns
- Toxic
Assessing Assay Utility

• **Assay predictivity against in vivo toxicity at 10uM:**

  \[\text{In reality!}\]

  ![Graph showing cytoxicity at 10uM](chart.png)
Assessing Assay Utility

- Assay predictivity against in vivo toxicity at 10uM:

  In reality!

[Diagram showing cytotoxicity (uM) with categories of non-toxic, unknowns, and toxic.]

Select those assays with the highest odds ratios

Minimize mechanistic overlap on combining!
Each variable has its own ability to identify toxic cpds in terms of odds ratio, coverage and false-positive rate.
Each variable has its own ability to identify toxic cpds in terms of odds ratio, coverage and false-positive rate.
Each variable has its own ability to identify toxic cpds in terms of odds ratio, coverage and false-positive rate.
Each variable has its own ability to identify toxic cpds in terms of odds ratio, coverage and false-positive rate.
Objective: To derive a single score prediction of in vivo tox at 10uM to allow comparison of cpds/series across the panel of selected assays and properties.

- **Make use of Multi-Parameter Optimization**
  - the Score is on a 0 to 1 scale with 1 = \( J \) and 0 = \( L \)

- **Approx 35 discrete endpoints are in the combined model, e.g.**
  - Cerep binding assays (%inhib @ 10µM)
  - In vitro cytotoxicity assay (THLE)
  - Genetic Tox assays (BiolumAmes & IVMN)
  - Dofetilide binding and hERG

- **Incorporate knowledge of physical properties**
  - cLogP and TPSA (3/75 guideline)
  - Basic pKa
MPO Scoring Methodology:

\[
\text{CSE Score} = \left( y_1^w, y_2^w, y_3^w, y_4^w, \ldots \right)^{1/(w_1+w_2+w_3+w_4+\ldots)}
\]

For each assay: \( y, X_1 \) and \( X_2 \) and relative weight \( (w) \) were defined e.g.

Cannot disclose all the proprietary assay thresholds, weighting and scoring MPO at this time.
CSE v1 model performance

Findings@<10uM

- True Negatives: 68
- False Negatives: 84
- True Positives: 11
- False Positives: 54

- Sensitivity: 22%
- Specificity: 99%
- Accuracy: 46%

If a study cost ~$100k then 38 studies ~$4M!

Shows significant number of compounds with ‘findings’ had CSE score of <0.75

Clean@<10uM
Identified all Oral Drugs launched since 1990:

- Filtered to MW <600 to remove large biologics etc.
- Must have CEREP data generated in Pfizer database
- Gave 157 launched Drugs for analysis (a snapshot – not comprehensive)

With this data set:
17/18 Drugs with dose >500mgs have CSE Score >0.85
(exception being Gleevec; CSE Score 0.81; Typical oncology dose 400-800mg)

Low dose (<50mg) more forgiving of potential Safety Risks (high potency → high TI)
Caveat – this is only a subset of all launched drugs
Representative CSEv1.0 display

- **Genetic tox. risks**
- **THLE:** Indicators of cell toxicity
- **CEREP Promiscuity panel**
- **Potential CV safety**

**Toxicophore alerts**

**PhysChem properties**

**e.g. Paroxetine**
Delivering CSE to Project Teams

- Project teams are given an annotated report on the behavior of representative chemical series to allow them to consider risk of low dose tox in series selection. Reports are discussed with team in collaboration with DSRD colleagues, to make better informed series selection decisions and to support confidence in safety strategies.

- Approach was rolled out in 2010 to projects in chemotype selection and where necessary compounds going into exploratory in vivo studies are also profiled. Inlicensing opportunities have also been supported.

- > 70 small molecule projects supported in 2010
Improving Sensitivity

δ Version 1.0 sensitivity only at 22%

   • Can we improve this without sacrificing specificity?

   • Score dominated by off-target pharmacology
   • Require more mechanism-based “safety” endpoints
   • Refine variable weightings computationally!

   • Caution: excessive validation may lead to overfitting of training set!
“Simplifying” Promiscuity Assessment

- Reduce 15 assays to one variable:
Optimal Variable Selection

Violin Plots show distribution and density over the 1,000 models
Compound Safety Evaluator v2

Toxic_10uM
Clean@10uM
Findings@10uM

<table>
<thead>
<tr>
<th>Metric</th>
<th>Value 1</th>
<th>Value 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median</td>
<td>0.75</td>
<td>0.49</td>
</tr>
<tr>
<td>Mean</td>
<td>0.72</td>
<td>0.53</td>
</tr>
<tr>
<td>Range</td>
<td>0.86</td>
<td>0.82</td>
</tr>
<tr>
<td>Min</td>
<td>0.14</td>
<td>0.18</td>
</tr>
<tr>
<td>Max</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>StdDev</td>
<td>0.19</td>
<td>0.22</td>
</tr>
<tr>
<td>Q1</td>
<td>0.60</td>
<td>0.36</td>
</tr>
<tr>
<td>Q3</td>
<td>0.87</td>
<td>0.66</td>
</tr>
<tr>
<td>Count</td>
<td>106</td>
<td>235</td>
</tr>
</tbody>
</table>
Next Steps

• How can we incorporate dose and or potency?

- Address “false negative” space
  - Explore novel mechanisms of toxicity
  - Develop new assays and full validation – what does it mean!
- Severity?

![Graph showing dose and risk levels with examples for Lamotrigine and Paroxetine.](image-url)
• **Caveats:**
  - Limited data points for curve fit
  - Imprecise severity scores
  - What about time axis?

• **Benefits:**
  - Enables us to use all compounds
  - Eliminate false positives?
Creating an *in silico* CSE

- In vitro assays now modeled and incorporated into an *in silico* CSE ….

- Evaluated using ~ 80 compounds that were registered before synthesis & testing

### Calculated CSE Score:
- True Positive Rate (sensitivity): 97%
- True Negative Rate (specificity): 83%
- Positive Predictive Value: 80%
- Negative Predictive Value: 97%

[CSE score v2.0 graph]
In our small molecule discovery programs we employ a predictive platform which detects around 60% of the compounds which cause low dose toxicity in preclinical species (with a <10% false-positive rate).

- 2010 - Guided medicinal chemistry in >70 discovery projects at Pfizer.
- Approach initiates safety considerations *early* in projects
- Framework for evaluating the predictivity /utility of new assays
- Building such a tool relies heavily on well characterized training compound sets and excellent engagement across biologists, chemists and computational scientists
- Our current focus for this tool is to address the impact of dose projection, and to model severity of toxicity.
- Value is in steering away from risky chemical space, better survival and resource utilization.
Key Acknowledgements

Compound Safety Prediction
- William Pennie
- Yvonne Will
- Karen Leach
- Mat Pletcher
- Thomas Schroeter
- Minghu Song
- *and all CSP colleagues*

Medicinal Chemistry
- Tony Wood
- Dave Price
- Kevin Dack

Computational Support
- Jackie Klug-McLeod
- Eric Gifford
- Rishi Gupta

...numerous partners in
- Non Clinical Statistics & Research Informatics
- Computational Sciences & Data and Design Analytics
- Drug Safety Research & Development