Computational Models for the Prediction of Pharmacokinetic Properties at Pharmacopeia, Inc.

Bill Egan & Kenneth M. Merz, Jr.
Center for Informatics & Drug Discovery
Pharmacopeia, Inc.

March 26, 2000
**ADME Parameters**

**Drugs**

**Excretion**

**Metabolism & Toxicology**

**Distribution**

**Absorption**

**Solubility**

**Serum Levels**

**Decision Points**

- Soluble pH 6-8
- >30% absorbed
- <95% PB
- Serum Levels
- half-life, CNS
- rate, toxicity
- rate
Requirements for Successful Predictive Modeling

1. Data must cover the entire region of chemical space related to the property of interest
   ⇒ need larger data set
   ⇒ range of property of interest must be covered

2. Data must be good quality
   ⇒ identify/remove outliers
   ⇒ measurement precision must be good enough

3. Mathematical models
   ⇒ must be appropriate for the amount, type, and quality of the data
Intestinal Absorption

Passive Diffusion
- Transcellular (membrane permeation)
  - lipid/aqueous interface “Fluid/Mosaic Model”
- Paracellular (tight junction passage)
  - limited to < ~200 MW
  - Effects from size, hydrophilicity
    (H-bonding), lipophilicity, and charge

Active Transport
- Influx - multiple transporter mechanisms (7+)
- Efflux - p-Glycoprotein (MDR)
Caco-2 Assay Variability

Random Survey of Intra-laboratory Variability:
RSD = 5.6%, ~<10%, 10.3%, 12.7%, 28.3%, 95.5%

Inter-laboratory variability:

Absorption Model Data

**Literature Data**
- 199 compounds ≥90% human gut absorption
- 35 compounds <30% human gut absorption
- 40+ yrs of data with high inter-laboratory variation, differing assays, and non-linear relationships preclude the use of mid-range data

**Pharmacopeia Compounds**
- 197 compounds with Caco-2 cell permeability
- 63 compounds w/ $P_{app} > 100$ nm/s (~ 90% absorption)
- 78 compounds w/ $P_{app} < 34$ nm/s (~ <30% absorption)
Robust Outlier Detection

Compounds >90% absorbed
Compounds <30% absorbed

Factor 2

Standard 95% confidence region

Robust 99.9% confidence region

Factor 1
Results of Outlier Detection

- compounds >90% absorbed
- actively transported compounds >90% absorbed
- compounds <30% absorbed
Clinically-Used Drugs

*Physician’s Desk Reference*

- 440 orally delivered compounds (tablets, capsules, liquid suspensions)

- 77% are predicted to have ≥90% absorption, (95% confidence) (83.1% excluding actively transported compounds)

- 84.8% are predicted to have ≥90% absorption, (99% confidence) (91.4% excluding actively transported compounds)
PDR Orally Available Drugs

- PDR actively transported compounds
- PDR compounds < 30% absorbed
- PDR compounds 30-90% absorbed
- Remaining orally available PDR compounds
Drug-like Compounds

Comprehensive Medicinal Chemistry Database

• 7,577 compounds
• 5,836 drug-like compounds by listed class

• 75.1% drug-like are predicted to have ≥90% absorption (95% confidence)

• 83.7% drug-like are predicted to have ≥90% absorption (99% confidence)
Model Predictions vs Caco-2 Results

Absorption Model Score

$P_{\text{app}}$ (nm/s)
Two Recent Optimization Series

Controlling Validation

\[ P_{\text{app}} \text{ (nm/s)} \]

Absorption Model Score
### Observed Probability of Caco-2 Permeability vs. Absorption Model Score

<table>
<thead>
<tr>
<th>$T^2$</th>
<th>$&lt;34$</th>
<th>34-100</th>
<th>&gt;100</th>
<th>“Good”</th>
</tr>
</thead>
<tbody>
<tr>
<td>$&lt;6.12$</td>
<td>18.4%</td>
<td>25.5%</td>
<td>56.1%</td>
<td>81.6%</td>
</tr>
<tr>
<td>6.12-9.5</td>
<td>35.9%</td>
<td>30.3%</td>
<td>33.8%</td>
<td>64.1%</td>
</tr>
<tr>
<td>$&gt;9.5$</td>
<td>65.0%</td>
<td>26.1%</td>
<td>8.9%</td>
<td>35.0%</td>
</tr>
</tbody>
</table>

Computed using probability distributions observed for replicate measurements of 122 PCOP compounds.
## BBB Penetration for 2 PCOP Lead Optimization Projects

<table>
<thead>
<tr>
<th>Compound</th>
<th>Predicted LogBB</th>
<th>Rat LogBB</th>
<th>in vitro BBB Permeability (nm/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>688897</td>
<td>0.461</td>
<td>0.167</td>
<td>0</td>
</tr>
<tr>
<td>397128</td>
<td>-0.095</td>
<td>0.017</td>
<td>0</td>
</tr>
<tr>
<td>388052</td>
<td>-0.020</td>
<td>-0.136</td>
<td>0</td>
</tr>
<tr>
<td>318629</td>
<td>0.191</td>
<td>-0.684</td>
<td>0</td>
</tr>
<tr>
<td>823709</td>
<td>*</td>
<td>-0.717</td>
<td>0</td>
</tr>
<tr>
<td>827681</td>
<td>*</td>
<td>-0.73</td>
<td>0</td>
</tr>
<tr>
<td>416422</td>
<td>*</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>416511</td>
<td>*</td>
<td>0</td>
<td>429</td>
</tr>
<tr>
<td>416712</td>
<td>0.67</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>417321</td>
<td>*</td>
<td>5</td>
<td>57</td>
</tr>
<tr>
<td>425438</td>
<td>*</td>
<td>*</td>
<td>151</td>
</tr>
<tr>
<td>425439</td>
<td>*</td>
<td>*</td>
<td>230</td>
</tr>
<tr>
<td>425759</td>
<td>*</td>
<td>*</td>
<td>166</td>
</tr>
<tr>
<td>425764</td>
<td>borderline *</td>
<td>*</td>
<td>57</td>
</tr>
<tr>
<td>426695</td>
<td>0.71</td>
<td>*</td>
<td>47</td>
</tr>
<tr>
<td>426696</td>
<td>0.43</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>427763</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>430526</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

* = likely very poor
Solubility modeling

- Aqueous solubility is a key property in determining the bioavailability of a drug
- Low solubility can hinder the biological activity of a compound

Experimental Data

- Data for 330 small organic molecules, solubility range $10^{-12}$ to $10^{-2}$
- Less data available for drugs/drug-like compounds
- Wide variety of experimental conditions makes modeling a challenge
Solubility Model Test Data

$r^2 = 0.893$, RMSE = 0.73, 34 compounds
Using LibProp™ to Design Libraries (1)

Factor 1

Before Refinement

Best 19%
Good 30%
Mediocre 25%
Ugly 26%

Factor 2

CIDD
Using LibProp™ to Design Libraries (2)
LibProp™ Absorption Profile of the Pharmacopeia LDS Collection

- 3.3 million compounds, 64 libraries
- 64% should be moderately well-absorbed
- 87% should be good leads
Conclusions & Thoughts

♦ Useful & predictive models can be created
♦ PCOP Labs has implemented the absorption/BBB model in LibProp™ 2.0 for combinatorial library design
♦ Validating the solubility model for inclusion into LibProp™ 2.0
♦ Working on protein binding

♦ ADME problems best solved in a piecewise fashion
♦ A much larger role for in vitro assays for providing consistent data and confirming predictions
# A “Back-of-the-Envelope” View of Drug Development

<table>
<thead>
<tr>
<th>Activity</th>
<th>Abs.</th>
<th>Metab.</th>
<th>Tox.</th>
<th>Excretion</th>
<th>Selectivity</th>
<th>New Drug Success Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Simple estimate of implied step-wise success rate</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>TODAY</strong></td>
</tr>
<tr>
<td>40% × 40% × 40% × 40% × 40%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>= 1%</td>
</tr>
<tr>
<td>46% × 46% × 46% × 46% × 46%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>= 2%</td>
</tr>
<tr>
<td>55% × 55% × 55% × 55% × 55%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>= 5%</td>
</tr>
<tr>
<td>63% × 63% × 63% × 63% × 63%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>= 10%</td>
</tr>
<tr>
<td>72% × 72% × 72% × 72% × 72%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>= 20%</td>
</tr>
</tbody>
</table>

*CIDD*
Acknowledgements

George Lauri
Ailan Cheng
Jonathan Burbaum