Establishing a therapeutically beneficial new chemical entity (NCE) can be broadly classified into research (discovery) and development phases. Drug development is generally divided into nonclinical (animal) and clinical (human) testing stages, with regulatory approval of an investigational new drug application separating the two. In contrast, drug discovery does not require regulatory oversight and follows the general path of: target validation —→ assay development —→ high-throughput screening —→ hit to lead —→ lead optimization (LO). Promising drug candidates are selected and transitioned from late LO into nonclinical development, although the specific processes for doing so vary between organizations.

Along with potency and safety, a third key factor in determining an NCE’s viability as a potential clinical drug candidate is its ADME (absorption, distribution, metabolism, and excretion) properties. In the not-so-recent past, poor ADME properties were responsible for more clinical drug trial failures than efficacy or safety.1 Typically ADME properties are initially evaluated in early LO with in vitro experiments, while more specific and resource-intensive in vivo exposure data is collected during later stage drug candidate optimization and selection. Both in vitro and in vivo data are used to assess for liabilities such as poor bioavailability, high clearance, potential for drug-drug interactions, etc. A more thorough mechanistic understanding of how ADME properties can be optimized has evolved in the past decade or so, which has enabled pharmaceutical scientists to select more robust NCEs. This enhanced understanding also enables earlier stage in silico modeling and simulation.

Human exposure predictions during the nonclinical development phase have also improved. Historically, compartmental pharmacokinetic (PK) methods have used allometric scaling of preclinical animal data to predict exposures in humans. In compartmental PK, the body is arbitrarily represented by either one or several theoretical compartments without specifying anatomy and physiology. These analyses, however, require in vivo data, which is often unavailable during early compound optimization. Furthermore, allometric scaling in simple compartmental models has no mechanistic basis, which can limit the predictive ability.

Physiologically based pharmacokinetic (PBPK) methods are an alternative approach to address these challenges. Although the concept of PBPK modeling is not new, its use within the pharmaceutical industry has been limited until recently following regulatory acceptance by the Food and Drug Administration. The PBPK approach utilizes anatomical and physiological parameters for either in silico/in vivo extrapolation (ISIVE) or in vitro/in vivo extrapolation (IVIVE), which predicts full PK in animal species or humans. These predictions require only a compound’s in silico or in vitro ADME properties. A significant disadvantage of PBPK has been the complexity, requiring hundreds of differential equations and biopharmaceutical parameters. This shortcoming, however, has been rectified in whole-body PBPK models incorporated within several commercially available software products such as PK-Sim (Bayer Technology Services), Simcyp (Certara), and GastroPlus (Simulations Plus, Inc.).

Traditionally PBPK modeling and simulation is used in the late nonclinical to early clinical development space to support dose selection for first-in-human studies, potential drug-drug interaction effects, and possible exposure differences resulting from a change in formulation. This powerful methodology can, however, also be employed in the early to late discovery stage as discussed further in a few select examples below.

**CASE STUDY EXAMPLES**

**In Silico/In Vivo Extrapolations for Estimating Oral Bioavailability**

A database of 62 drugs was constructed using literature data that included molecular structures, dose, and oral bioavailability (F%). The major reported clearance pathways for these compounds were all CYP mediated.2 The total reported F% range across all
A tool for sparing resources in the discovery environment.

By Siladitya Ray Chaudhuri, Ph.D., Johnson & Johnson; Michael B. Bolger, Ph.D., Simulations Plus Inc.; Michael Lawless, Ph.D., Simulations Plus Inc.; Anand Balakrishnan, Ph.D., Bristol-Myers Squibb; and John Morrison, Ph.D., Bristol-Myers Squibb

Compounds varied from 3 percent (fluphenazine) to 99 percent (diazepam, galantamine, glimepiride, indomethacin, and tamsulosin) with an overall average of 60 percent. These compounds were therefore representative of early drug candidates.

Bioavailability for each of the 62 drugs was predicted by integrating quantitative structure activity relationship modeling data and PBPK modeling data. GastroPlus utilizes an advanced compartmental absorption and transit model to estimate mechanistic absorption (MA) and PBPK. This simulates the process by which orally dosed drugs dissolve in the stomach, transit into the intestine, are absorbed through the gut wall, and circulate systemically. The MA/PBPK model also incorporated estimates of intestinal absorption as well as both intestinal and liver first pass extraction.

First pass extraction in the gut and liver was simulated using cytochrome P450 (CYP) metabolism models for 5 CYP isoforms (1A2, 2C9, 2C19, 2D6, and 3A4) within the ADME Predictor software (Simulations Plus, Inc.). The initial step was to predict whether a drug molecule was a substrate for any of the 5 CYP isoforms, based upon substrate/nonsubstrate classification criteria. Each of the 62 molecules was predicted to be a significant CYP substrate, which was consistent with their major reported clearance pathway. Furthermore, the CYP isoform with highest predicted intrinsic clearance (CLint) was the same isoform as the major reported clearance pathway for the majority of these molecules (68 percent). Michaelis-Menten kinetic parameters were calculated for each compound and included in the MA/PBPK simulations. Overall, 69 percent of the compounds in this sample set had predicted F% values within twofold of their reported in vivo F% as shown in Figure 1.

In Silico/In Vivo and In Vitro/In Vivo Extrapolations for Predicting Plasma Concentration Profiles Following Oral Dosing

Risperidone is an antipsychotic drug molecule with physicochemical and biopharmaceutical properties consistent with many discovery stage compounds. It is designated as a Biopharmaceutics Classification System Ib weak base compound with low aqueous solubility.

Figure 1. Predicted versus Observed Oral Bioavailability (F%) for 62 Drug Molecules Simulated with an In Silico Mechanistic Absorption (MA) and Physiologically Based Pharmacokinetic (PBPK) Model

Notes:  i) Bioavailability (F%) values were estimated using in silico rCYP Km and Vmax values for gut and liver extraction as well as in silico physicochemical parameters with GastroPlus version 9.0 (Simulations Plus, Inc., Lancaster, CA 93534 USA). ii) The model assumed a 35-year-old American male physiology.
Table 1: In Silico Estimates and Experimental In Vitro Values for Risperidone Physicochemical and Biopharmaceutical Properties

<table>
<thead>
<tr>
<th>PROPERTY</th>
<th>ESTIMATE (in silico)</th>
<th>OBSERVED (in vitro)</th>
<th>NOTES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log P (octanol/water)</td>
<td>3.23</td>
<td>3.04</td>
<td>reference 9</td>
</tr>
<tr>
<td>Aqueous solubility (pH 10.0)</td>
<td>0.08 mg/mL</td>
<td>0.04 mg/mL</td>
<td>Marcus Brewster private</td>
</tr>
<tr>
<td>pKa (weak base)</td>
<td>7.8 and 3.0</td>
<td>8.2 and 3.1</td>
<td>reference 9</td>
</tr>
<tr>
<td>Jejunal Permeability (human)</td>
<td>3.35 x 10⁻⁴ cm/s</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>MDCK Cell Permeability</td>
<td>-</td>
<td>19.8 x 10⁻⁴ cm/s</td>
<td>reference 8</td>
</tr>
<tr>
<td>Blood/Plasma Concentration Ratio</td>
<td>0.69</td>
<td>0.67</td>
<td>reference 9</td>
</tr>
<tr>
<td>Fraction Unbound in Plasma</td>
<td>0.08</td>
<td>0.12</td>
<td>reference 10</td>
</tr>
<tr>
<td>Unbound CLint</td>
<td>167 µL/min/mg Prot.</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CYP2D6 S+Km</td>
<td>2.69 µM</td>
<td>N/A</td>
<td>-</td>
</tr>
<tr>
<td>CYP3A4 S+Km</td>
<td>13.3 µM</td>
<td>N/A</td>
<td>-</td>
</tr>
<tr>
<td>CYP2D6 Vmax</td>
<td>not reported</td>
<td>7.5 pmol/min/pmol</td>
<td>reference 13</td>
</tr>
<tr>
<td>CYP3A4/5 Vmax</td>
<td>not reported</td>
<td>0.6 pmol/min/pmol</td>
<td>reference 13</td>
</tr>
</tbody>
</table>

Notes: i) Estimates were obtained using ADMET Predictor (AP ver. 7.2) (Simulations Plus, Inc.)

Physiologically Based Pharmacokinetic Modeling and Simulation for Drug Candidate Optimization and Selection

An MA/PBPK model was used to generate an in silico prediction of the plasma profile following oral administration of a two milligram immediate release (IR) tablet as shown in Figure 2. This purely in silico model provides a reasonable approximation of the in vivo data. The simulated area under the curve was within 10 percent of the observed in vivo value, and both the Cmax and Tmax simulations were within twofold of the observed in vivo values.

However, the addition of in vitro data further improves the modeling accuracy. Risperidone was reported by Fang et al to be a substrate for metabolism by CYP2D6 and CYP3A4, which was accurately predicted by ADMET Predictor. Although experimental values of Km for these enzymes were not reported, in silico estimates from the ADMET Predictor were used along with the reported Vmax values (Table 1).

solubility and high permeability. Table 1 lists in silico estimates and in vitro experimental values for the relevant properties.  

An MA/PBPK model was used to generate an in silico prediction of the plasma profile following oral administration of a two milligram immediate release (IR) tablet as shown in Figure 2. This purely in silico model provides a reasonable approximation of the in vivo data. The simulated area under the curve was within 10 percent of the observed in vivo value, and both the Cmax and Tmax simulations were within twofold of the observed in vivo values.

However, the addition of in vitro data further improves the modeling accuracy. Risperidone was reported by Fang et al to be a substrate for metabolism by CYP2D6 and CYP3A4, which was accurately predicted by ADMET Predictor. Although experimental values of Km for these enzymes were not reported, in silico estimates from the ADMET Predictor were used along with the reported Vmax values (Table 1).
Furthermore, Risperidone’s relatively high log P and high basic pKa (Table 1) were anticipated to render the compound susceptible to lysosomal trapping in both enterocytes and hepatocytes as observed for other central nervous system targeted drugs. This effect has been demonstrated in Caco-2 cells treated with bafilomycin A1 (an inhibitor of vacuolar-type H+-ATPase). Relative to untreated cells, lysosomal pH increases, and the ion-trapping phenomenon for lipophilic weakly basic compounds is diminished. Figures 3a and 3b show the in silico cellular simulations of Caco-2 apparent permeability in the apical to basolateral direction for Risperidone using MembranePlus (Simulations Plus, Inc.), assuming lysosomal pH = 4.0 (untreated) and pH = 6.5 (bafilomycin A1 treated), respectively. Both the rate and extent of Risperidone basolateral transfer (blue curves) increased with the simulated bafilomycin treatment (blue curves in Figure 3b vs Figure 3a) and the lysosomal concentration (pink curve) greatly decreased indicating a strong potential for lysosomal trapping.

Since lysosomal trapping cannot be directly simulated in the MA/PBPK model, enterocyte protein binding (at 92 percent bound) was used instead as a mechanistic surrogate. This had the effect of increasing the dwell time of Risperidone prior to crossing the basolateral membrane and entry into the portal vein. In addition, CYP metabolism as well as intestinal and liver first pass estimates were incorporated into the model for a two milligram IR tablet dose of Risperidone (Figure 2b).

The F% for the IVIVE (Figure 2b) is similar to the ISIVE result (Figure 2a). However, the Cmax and Tmax results are much more accurate—within 10 percent of the observed clinical values. These results illustrate the potential value of in silico property estimates and PBPK simulations for use in early discovery and also demonstrate the value of in vitro experimental data when refining the early PBPK results.

### Modeling and Simulation to Assess the Role of Physicochemical, Biopharmaceutical, and Formulation Properties on Oral Exposures

Historically, evaluation of early discovery drug candidate physicochemical, biopharmaceutical, and formulation properties required in vitro assays and preclinical in vivo testing. Despite the use of high-throughput solubility and permeability assays, the large number of early drug candidates to be evaluated imposes practical limits on the thoroughness of such assessments. Fortunately this data can instead be complemented with PBPK-based tools for a more mechanistic early link to in vivo oral exposures.

In the early discovery space, in vivo resources are limited, and PBPK-based absorption modeling can provide a preliminary risk indication of key physicochemical, physiological, and formulation properties (Table 2). It must be noted that the reliability of such “bare bones” PBPK modeling is not equivalent to that for a clinical asset. However, the ability to triage compounds and refine the in silico models as additional data becomes available is invaluable. The use of PBPK-based simulations in early discovery enables pharmaceutical scientists to derive greater benefit for initially sparse datasets, identify risk factors, and prioritize appropriate follow-up studies.

### CONCLUSIONS

The above examples help demonstrate the accuracy with which in vivo PK data can be simulated using either simple in silico inputs or more precise in vitro data. These simulations also provide an efficient means to test mechanistic hypotheses, for instance, lysosomal trapping, CYP mediated clearance, GI physiology implications, or food effects. Such models offer an opportunity for pharmaceutical scientists to provide critical

**Table 2: Examples of Modeling and Simulation for Physicochemical, Biopharmaceutical, Physiological, and Formulation Properties on Oral Drug Exposures**

<table>
<thead>
<tr>
<th>APPLICATION</th>
<th>NOTES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impact of changes in API properties (e.g., particle size, solubility, etc.)</td>
<td>Reference 16</td>
</tr>
<tr>
<td>Assessing effects of pH-dependent solubility</td>
<td>Reference 17</td>
</tr>
<tr>
<td>Impact of variability in GI physiology</td>
<td>Reference 18</td>
</tr>
<tr>
<td>Understanding effect of food on oral absorption</td>
<td>Reference 19</td>
</tr>
<tr>
<td>Assessment of formulation choices</td>
<td>Reference 20</td>
</tr>
</tbody>
</table>
input during the early stages of compound design and optimization. It is anticipated that such resource-sparing modeling will become more prevalent in the drug discovery environment.

**REFERENCES**