APPLICATION OF A GENERIC PHYSIOLOGICALLY BASED PHARMACOKINETIC MODEL TO THE ESTIMATION OF XENOBIOTIC LEVELS IN HUMAN PLASMA

Cyprotex Discovery Ltd., Macclesfield, Cheshire, United Kingdom

Received March 24, 2005; accepted October 11, 2005

ABSTRACT:
Estimation of xenobiotic kinetics in humans frequently relies upon extrapolation from experimental data generated in animals. In an accompanying paper, we have presented a unique, generic, physiologically based pharmacokinetic model and described its application to the prediction of rat plasma pharmacokinetics from in vitro data alone. Here we demonstrate the application of the same model, parameterized for human physiology, to the estimation of plasma pharmacokinetics in humans and report a comparative evaluation against some recently published predictive methods that involve scaling from in vivo animal data. The model was parameterized through an optimization process, using a training set of in vivo data taken from the literature, and validated using a separate test set of published in vivo data. On average, the vertical divergence of the predicted plasma concentrations from the observed data, on a semilog concentration-time plot, was 0.47 log unit. For the training set, more than 80% of the predicted values of a standardized measure of the area under the concentration-time curve were within 3-fold of the observed values; over 70% of the test set predictions were within the same margin. Furthermore, in terms of predicting human clearance for the test set, the model was found to match or exceed the performance of three published interspecies scaling methods, all of which showed a distinct bias toward overprediction. We conclude that the generic physiologically based pharmacokinetic model, as a means of integrating readily determined in vitro and/or in silico data, is potentially a powerful, cost-effective tool for predicting human xenobiotic kinetics in drug discovery and risk assessment.

Physiologically based pharmacokinetic (PBPK) models are mathematical descriptions of the flow of blood throughout the body, developed for the simulation of xenobiotic absorption, distribution, and elimination. Such models have been used by scientists from a number of different disciplines who are interested in the simulation and prediction of exposure (Grass and Sinko, 2002; Leahy, 2003).

The application of a generic form of a PBPK model to the prediction of xenobiotic plasma levels in rat following an intravenous dose has been reported in an accompanying publication (Brightman et al., 2006). Here we describe the work that we have done to parameterize the same PBPK model for humans and to assess the reliability of the model in estimating plasma levels of xenobiotics, where these values are known from experimentation. In addition, we draw comparisons with alternative methods for predicting human pharmacokinetic properties that involve extrapolation from experimental data generated in animals.

Just as there are numerous published compound-specific PBPK models for the rat that use data derived from in vivo studies (Sugita et al., 1982; Igarı et al., 1983; Tsuji et al., 1983; Bernareggi and Rowland, 1991; Kawai et al., 1994; Blakey et al., 1997), there are many examples of comparable PBPK models for humans that rely upon scaling from in vivo animal data to simulate the human pharmacokinetics of a particular compound, and frequently incorporate observed clearance data from human subjects (Igari et al., 1983; Sawada et al., 1985; Tsuji et al., 1985; Bernareggi and Rowland, 1991; Kawai et al., 1994, 1998). The PBPK model for humans presented herein seems to represent the only truly generic model to be published to date, since it has been parameterized for human physiology, independently of any specific compound, and the manner in which in vivo distribution and elimination kinetics are predicted is the same for all xenobiotics. Furthermore, the only compound-dependent inputs that the model requires are readily determined in vitro, or even in silico.

In this paper, we have concentrated on predicting the in vivo pharmacokinetics of compounds for which plasma levels have been determined following an intravenous dose. Work that we have done to extend the model to predict both human and rat plasma levels following an oral dose will be reported separately.

Materials and Methods
Model Inputs. A generic PBPK model, which enables the prediction of the pharmacokinetic behavior of any given compound dosed intravenously in a specified human population, without recourse to data derived through in vivo studies, is presented herein. The compound-dependent inputs required by the model are the same as those listed previously (Brightman et al., 2006).

Model Description. The PBPK model is based upon that published by Bernareggi and Rowland (1991), as shown in their Fig. 1, but with substantial

ABBREVIATIONS: PBPK, physiologically based pharmacokinetic; ADME, absorption, distribution, metabolism and elimination; AUC, area under the concentration-time curve; AUCt1-tlast-DN, dose-normalized AUC from the first to the last recorded time points; Clint, hepatic intrinsic metabolic clearance; fup, fraction unbound in plasma; fuu, fraction unbound in the interstitial fluid; IQ, interquartile; PK, pharmacokinetic(s); wMLFE, weighted mean log -fold error.
modification of the tissue distribution and elimination components, and comprises a series of compartments representing 14 major organs and tissues in the body, interconnected by further compartments representing arterial and venous blood pools, according to the principles developed by Bischoff and others (Bischoff, 1975).

The additional features of the model, including adaptations to facilitate modeling of “diffusion-limited” distribution of an intravenously administered compound into the various tissues and organs, as well as the various processes involved in renal excretion, are described in detail in the accompanying paper discussing its application to the prediction of rat in vivo pharmacokinetics (Brightman et al., 2006).

**Model Parameters.** The physiological parameters used in the model were obtained from the literature and are given in the Appendix (which is available online as supplemental data); these were scaled according to the actual body weights of the subjects used in the clinical studies being simulated. Tissue and organ volumes were largely derived from a single comprehensive compilation of physiological data for use in pharmacokinetic models (Brown et al., 1997), supplemented by data from other sources for skin (Mapleson, 1963) and testes (Spector, 1956) volumes, and represent the extravascular (combined interstitial fluid and intracellular space subcompartments) volumes only. Blood flow rates were taken almost exclusively from Bernareggi and Rowland (1991), with the exception of the blood flow to the testes (Williams and Leggett, 1989). The glomerular filtration rate and urine flow rate were from Davies and Morris (1993) and Tang-Liu et al. (1983), respectively, whereas the renal tubular lumen volume was obtained from a textbook of physiology (Pitts, 1974). A hematocrit of 0.441 (Altman and Dittmer, 1971) was assumed.

Parameterization of the distribution and elimination components of the generic PBPK model for humans required the development of a number of correlation models. Two such models, for the prediction of parameters corresponding to the effective in vivo lipophilicity and plasma protein binding, were derived through a process of optimization of the performance of the PBPK model, as described in the companion paper (Brightman et al., 2006). A comprehensive training set of in vivo data was used for this purpose and is described in greater detail below.

The following parameters were derived as detailed previously (Brightman et al., 2006): permeability-surface area products for organ and tissue distribution; intracellular space/interstitial fluid (unbound) partition coefficients; blood/plasma concentration ratio ($R$); parameters governing renal excretion; hepatic microsomal intrinsic clearance (CLint), and fractions unbound in plasma ($fu_p$) and interstitial fluid ($fu_t$).

Stochastic simulations were performed to incorporate known variability in the subject body weights and imprecision in the values of the physicochemical parameters, as described previously (Brightman et al., 2006).

**Training Dataset.** The set of in vivo data used in training the model comprised 180 instances of data (where an instance corresponds to a single plasma concentration-time profile) for 69 different compounds and was derived from numerous published clinical studies of intravenous dosing in human subjects. The compounds in this diverse training set were drawn from many therapeutic areas. The required model inputs for the training set compounds were obtained as described previously (Brightman et al., 2006). As before, no attempt was made to eliminate compounds from the training set on the basis of any features of the in vivo pharmacokinetic behavior.

**Test Dataset.** To objectively evaluate the performance of the model, an independent set of in vivo test data was constructed. This consisted of 39 instances of plasma concentration-time data for 18 compounds dosed intravenously in human subjects. These data were derived from the literature.

The test set compounds were varied in terms of physicochemical properties and represented diverse therapeutic areas. The requisite model inputs for these compounds were derived as described previously (Brightman et al., 2006).
Calculation of the Plasma Concentration Weighted Mean Log -Fold Error (wMLFE). For each pair of in vivo and simulated plasma concentration-time profiles, the log -fold prediction error was determined at each simulated time point for which there were corresponding in vivo data, and the mean of these errors over all time points was calculated, to give an overall mean prediction error for each instance of simulated data. The wMLFE represents the weighted mean of these individual means. The weights used in the calculation arise from there being multiple instances and/or sources of in vivo data for several compounds, and hence the contribution of each individual log -fold prediction error to the overall mean is weighted accordingly; i.e., so that each compound contributes equally, whatever the number of instances of in vivo data for that compound.

Results

Model Validation. For any given set of input data, output from the PBPK model is in the form of a predicted plasma concentration-time profile. When stochastic simulations are performed for a single set of input data, each iteration generates a predicted profile, and hence the total output consists of a population of profiles that reflect the inherent uncertainty in the input data. Examples of typical simulation results, plotted on the same axes as the corresponding in vivo data, are given in Fig. 1 for selected training set compounds and in Fig. 2 for a similar selection of test set compounds.

The simulated profiles in Fig. 1, A and B, and Fig. 2, A and B, illustrate accurate estimation of plasma concentrations over time, for selected training set and test set compounds, respectively. There is little variation within the population of profiles generated for either of the training set compounds, dexamethasone and verapamil (Fig. 1, A and B), and the fit of the model output to the single set of observed data is precise and almost exact in both cases. Conversely, Fig. 2, A and B, demonstrates accurate, but less precise simulation of the in vivo plasma data for the test compounds, biperiden and acecainide; although the median predicted profiles for these compounds depart very slightly from the observed profiles, the ranges of predicted profiles encompass the in vivo data. The variation within the model output for biperiden, for example, was generated from both known variability in the subject weights and a combination of variability and uncertainty arising from multiple estimates of in vivo $f_{un}$ and $CL_{int}$.

Some other simulation results are shown in Fig. 1, C and D, and Fig. 2, C and D. Somewhat inaccurate estimation of the in vivo tissue distribution of the training set compound phenytoin, resulting in a tendency to underpredict plasma levels, is indicated by Fig. 1C. However, there is clearly a degree of variability in the observed data for this compound, which is reflected in the model output. In contrast, the elimination half-life of alprazolam has been underestimated (Fig. 1D). The median predicted profiles for the test set compounds dofetilide and budesonide deviate from the observed profiles (Fig. 2, C and D), but the in vivo data are captured by the ranges predicted; hence, the model results can still be considered acceptable.

For the remainder of the results presented here, the median of the population of predicted profiles generated from each set of input data was used as an individual estimate of the plasma concentration-time course. To assess the overall performance of the model in terms of successfully predicting in vivo plasma levels, the plasma concentration wMLFE was determined for both the training and test sets; this

![Fig. 2. Predicted and observed plasma concentration-time profiles for selected test set compounds: A, biperiden; B, acecainide; C, dofetilide; and D, budesonide. The in vivo data (filled symbols) are taken from the literature (Ryrfeldt et al., 1982; Grimaldi et al., 1986; Coyle et al., 1991; Smith et al., 1992). The simulated data are the median values (solid line) of a population of predicted profiles generated from 100 stochastic simulations; also indicated are the 10th (dashed line) and 90th (dotted line) percentiles of the population.](image-url)
normalized AUC from the first to the last recorded time points to observed data in terms of a standardized parameter, the dose-vary in the estimates they yield, the simulation results were compared time to the first time point and from the last time point to infinity can curve (AUC), allowing direct comparisons to be made with analogous these can be used for estimation of standard pharmacokinetic (PK) predicted in vivo plasma concentration-time profiles that are generated, this set has a mean prediction error of more than 5-fold. In fact, very similar proportions of the two sets (59% of the training set and 61% of the test set) are on average within a factor of 2 above or below the observed data, and the same percentage of each set than for the training set, in that plasma concentration predictions shown in Fig. 3. Interestingly, the results seem rather better for the test set than for the training set, in that plasma concentration predictions are on average within a factor of 2 above or below the observed data points for a far greater percentage of the test set compounds. However, given the relatively small size of the test set, this discrepancy is unlikely to be significant. In fact, very similar proportions of the two sets (59% of the training set and 61% of the test set) are on average within 3-fold of the observed data, and the same percentage of each set has a mean prediction error of more than 5-fold.

Although the primary outputs from the PBPK model are the predicted in vivo plasma concentration-time profiles that are generated, these can be used for estimation of standard pharmacokinetic (PK) parameters of interest, including area under the concentration-time curve (AUC), allowing direct comparisons to be made with analogous in vivo data. Since different methods of extrapolating AUC from zero time to the first time point and from the last time point to infinity can vary in the estimates they yield, the simulation results were compared to observed data in terms of a standardized parameter, the dose-normalized AUC from the first to the last recorded time points (AUC$_{t1-tlast-DN}$).

The capability of the PBPK model to accurately predict AUC$_{t1-tlast-DN}$ in humans has been evaluated in terms of the median values and inter-quartile (IQ) ranges of the predicted/observed ratios, for both the training and test sets, as given in Table 1. These summary data indicate that prediction of AUC$_{t1-tlast-DN}$ is generally successful, although the predictions for the test set are again apparently more accurate than those for the training set. Thus, the median predicted/observed ratio is close to 1.0 for either set of compounds, and half of the test set predictions are within a factor of approximately 1.5 above or below the observed values, but the range is slightly greater for the same proportion of training set predictions.

The frequency distributions of the predicted/observed ratios of AUC$_{t1-tlast-DN}$ for both the training and test sets are more clearly demonstrated by the histograms shown in Fig. 4. The majority (55%) of the predictions for the training set compounds are within 2-fold of the observed values, and more than 80% are within 3-fold. Similarly, over 60% of the test set predictions are within a factor of 2 of the observed values, and greater than 70% are within a factor of 3. The predicted/observed ratios for this small test set are somewhat skewed toward low values.

**Comparison of the PBPK Model with Interspecies Pharmacokinetic Extrapolation Methods.** To further evaluate the potential of the generic PBPK model as a means for estimating the human pharmacokinetics of novel compounds within a drug discovery program, a performance-related comparison was made between the model parameterized for humans and three recently published interspecies scaling methods for the prediction of clearance in humans from rat in vivo data alone, derived for possible use as early pharmacokinetic screening tools (Caldwell et al., 2004; Ward and Smith, 2004). Both of the methods published by Caldwell et al. (2004) use simple allometric scaling, whereas the technique of Ward and Smith (2004) considers clearance as a fixed proportion of liver blood flow.

Predictions of human clearance for the different sets of compounds used in training the PBPK model, by Caldwell et al. (2004) and by Ward and Smith (2004), were evaluated in terms of the average -fold error, as defined by Caldwell et al. (2004), and by the proportion of predictions within 2-, 3-, and 4-fold of the observed values. The statistics presented for the two Caldwell et al. (2004) methods have been reproduced directly from the original source. To calculate comparable statistics for the Ward and Smith (2004) method, it was necessary to regenerate the original results of these authors from rat clearance data presented graphically in the source publication. The majority, but not all (97 of 103), of the data points could be extracted from the source. However, the median prediction -fold error calculated for this subset was the same as that reported in the publication for the full training set, and hence for the purposes of the analysis described here, the -fold error distribution we derived was assumed to be acceptably close to the true distribution.

For many of the compounds in the PBPK model training set there were multiple instances of in vivo data, derived from one or more published sources. The statistics shown for the PBPK model have therefore been calculated on the basis of the weighted mean -fold prediction error for each compound and, hence, are directly comparable to those for the published interspecies scaling methods, which

### Table 1

<table>
<thead>
<tr>
<th></th>
<th>Training Set</th>
<th>Test Set</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median</td>
<td>1.14</td>
<td>0.96</td>
</tr>
<tr>
<td>IQ Range</td>
<td>0.65–2.44</td>
<td>0.73–1.55</td>
</tr>
</tbody>
</table>

**Fig. 3.** Frequency distribution of the plasma concentration mean -fold errors for the training set (A) and test set (B) compounds. The training set and test set comprise 69 and 18 different drugs, respectively.

---

**Table 1** Summary of the AUC$_{t1-tlast-DN}$ predicted/observed ratio distributions for the training set and test set compounds.

- **Training Set**
  - Median: 1.14
  - IQ Range: 0.65–2.44

- **Test Set**
  - Median: 0.96
  - IQ Range: 0.73–1.55
were also determined on a per-compound basis. The weighting system used was the same as that used in the calculation of the plasma concentration wMLFE. There was considerable overlap in the compounds comprising the training sets for the four different methods.

The results of this analysis for the individual methods and their respective training sets, as shown in Table 2, demonstrate that the PBPK model apparently compares favorably with the Ward and Smith (2004) method, with the former having a lower average -fold error and predicting a slightly greater proportion of compounds within a factor of 3 of the observed values. However, they also suggest that the current version of the PBPK model might be somewhat less accurate than the methods of Caldwell et al. (2004) in the quantitative prediction of human clearance.

To further investigate the relative capabilities of the four methods, a similar statistical analysis of their predictions of clearance in humans for the PBPK model test set compounds was carried out, and is summarized in Table 3. This test set was completely independent, so that the compounds selected for inclusion were not present in any of the training sets, but also represented those for which suitable rat and human in vivo intravenous dosing data were readily available. Predictions of human clearance made by the Caldwell et al. (2004) and Ward and Smith (2004) methods were calculated from published values of rat in vivo clearance for the test set compounds and evaluated against the corresponding published human clearances. Again, the statistics for the PBPK model were derived from the weighted mean -fold prediction error for each compound.

The predictive accuracy of each method, and particularly of the Ward and Smith (2004) method, is generally poorer for the test set compounds than for those comprising the respective training sets, but still at what might be considered an acceptable level for the early stages of drug discovery (Caldwell et al., 2004; Ward and Smith, 2004). However, the deterioration in overall performance is proportionately less for the PBPK model than for the other methods. Consequently, the expected accuracy of the PBPK model seems to be at least comparable to that of Caldwell et al. (2004) in terms of predicting human clearance for a set of novel compounds. For both methods, the same number of predicted values are within a factor of 2 of the observed values, and an equal or greater proportion of the PBPK model predictions are within a factor of 3 or 4 above or below the observed values, although the average -fold error for the PBPK model is somewhat higher. On the evidence shown, the PBPK model would appear, however, to considerably outperform the method of Ward and Smith (2004) in predicting human clearance for novel compounds. Moreover, the PBPK modeling approach has the obvious advantage over all three interspecies extrapolation methods that no in vivo data are required to predict human pharmacokinetics, with a better or similar degree of accuracy.

Consideration of the median values of the predicted/observed ratios for the test set (Table 3) also reveals an interesting trend in the values predicted by the interspecies scaling methods, which all seem to have a tendency to overpredict human clearance, to varying degrees. This is confirmed by the histograms shown in Fig. 5, illustrating the frequency distribution of the predicted/observed ratios for each method. Although most of the values predicted by either of the Caldwell et al. (2004) methods are within a factor of 2 of the observed values, an almost equal number are between 2 and 10 times the observed values (Fig. 5, A and B). The majority (56%) of the values predicted by the Ward and Smith (2004) method are also between 2 and 10 times greater than the observed values (Fig. 5C). In contrast, the errors in the PBPK model predictions are normally distributed, with the highest single proportion of predictions, and the majority, overall, being within 2-fold of the observed values (Fig. 5D).

It is notable that all four methods were found to substantially overpredict clearance in humans of the test set compound digitoxin, which is also a significant outlier for the prediction of AUC₁₋₂₅₀₀₀ by the PBPK model: the predicted/observed ratios for clearance are 14.5, 9.5, 15.0, and 22.1 for the two Caldwell et al. (2004) methods, the method of Ward and Smith (2004), and the PBPK model, respectively. However, the human clearance of quinine, also in the test set, is similarly poorly predicted by the three interspecies scaling methods.


<table>
<thead>
<tr>
<th>Method</th>
<th>Average Fold Error</th>
<th>Fold Error &lt;2 (%)</th>
<th>Fold Error &lt;3 (%)</th>
<th>Fold Error &lt;4 (%)</th>
<th>Median Predicted/Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caldwell et al. (n = 176) Method I</td>
<td>2.16</td>
<td>56</td>
<td>77</td>
<td>88</td>
<td>1.74</td>
</tr>
<tr>
<td>Method II</td>
<td>2.22</td>
<td>52</td>
<td>79</td>
<td>86</td>
<td>1.60</td>
</tr>
<tr>
<td>Ward and Smith (n = 97)</td>
<td>2.64</td>
<td>49</td>
<td>70</td>
<td>79</td>
<td>2.54</td>
</tr>
<tr>
<td>PBPK Model (n = 69)</td>
<td>2.51</td>
<td>46</td>
<td>72</td>
<td>75</td>
<td>1.06</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Method</th>
<th>Average Fold Error</th>
<th>Fold Error &lt;2 (%)</th>
<th>Fold Error &lt;3 (%)</th>
<th>Fold Error &lt;4 (%)</th>
<th>Median Predicted/Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caldwell et al.</td>
<td>2.61</td>
<td>50</td>
<td>56</td>
<td>72</td>
<td>1.74</td>
</tr>
<tr>
<td>Method I</td>
<td>2.53</td>
<td>50</td>
<td>56</td>
<td>67</td>
<td>1.60</td>
</tr>
<tr>
<td>Method II</td>
<td>3.44</td>
<td>28</td>
<td>44</td>
<td>61</td>
<td>2.54</td>
</tr>
<tr>
<td>Ward and Smith</td>
<td>3.44</td>
<td>50</td>
<td>44</td>
<td>46</td>
<td>1.06</td>
</tr>
<tr>
<td>PBPK Model</td>
<td>2.78</td>
<td>75</td>
<td>72</td>
<td>72</td>
<td>1.06</td>
</tr>
</tbody>
</table>

**Fig. 4.** Frequency distribution of the predicted/observed ratios of AUC₁₋₂₅₀₀₀ for the training set (A) and test set (B) compounds. The training set and test set comprise 69 and 18 different drugs, respectively.
predicted/observed ratios of 9.5, 10.5, and 16.6 for the method of Caldwell et al. (2004) and Ward and Smith (2004), respectively), but more successfully predicted by the PBPK model, with a predicted/observed ratio of 4.3.

Finally, the same analysis was applied to another method published by Ward and Smith (2004) for predicting human clearance, requiring scaling from clearance in monkey. The average -fold error for the training set data of these authors was found to be 2.04, with 64%, 78%, and 89% of the predictions being less than 2-, 3-, and 4-fold in error, respectively. The overall level of performance therefore does not greatly exceed that demonstrated by the Caldwell et al. (2004) methods when applied to the corresponding training set. Hence, it would have been interesting to establish whether the same deterioration in predictive capability might be shown by this alternative method of Ward and Smith (2004), when applied to the PBPK model test set, as observed for their analogous method that depends on scaling from rat clearance. Unfortunately, however, appropriate monkey in vivo data were not available for this particular set of compounds. Similarly, Obach et al. (1997) have reported several methods of predicting clearance in humans (as well as volume of distribution and half-life) from preclinical pharmacokinetic data and/or in vitro data, but the compounds used in their analysis were not identified, and hence we were unable to directly compare the performances of the PBPK model and these particular methods.

Discussion

In this and the preceding paper (Brightman et al., 2006), we have demonstrated that a generic PBPK model can produce reliable predictions of mammalian PK following intravenous dosing, for a wide range of organic compounds, from a small number of readily determined, compound-dependent inputs. We expect the overall physiological description to be common to all mammalian species, and that the fundamental model could be successfully adapted to any species by changing the physiological parameters.

Because a PBPK model is physiologically realistic, with explicit equations for flows into the major tissues and organs of the body, the compound-dependent parameters required to simulate PK are real properties that can be measured. Consequently, such models can be used to define a minimum set of absorption, distribution, metabolism, and elimination (ADME) properties that must be obtained by in vitro or in silico screening to predict in vivo PK for any compound. In the same way that clinical trials are the point at which all the experimental data generated during drug discovery and preclinical development are integrated to produce a clear picture of the clinical potential of a new drug, so PBPK models can create the framework for integrating ADME, toxicity, and efficacy data throughout discovery and even preclinical development. Our aim is to develop models that are driven by inputs that can be determined within imposed cost and time constraints, and can reliably inform compound selection, by being able to predict human PK with sufficient confidence for a particular phase of the discovery/development process.

The PBPK model that we have described in this paper generates predictions of plasma concentrations and clearance in humans that appear to be sufficiently reliable to inform compound selection during early drug discovery. We have not, so far, been able to validate the human version of the model against drug discovery compounds to the same extent as the rat version. Nonetheless, there is no reason to believe that the relative robustness shown by the latter model in its ability to predict PK for compounds that are distant from the training set is not also a property of the human model. We have, however, demonstrated the equal or even superior ability of the current version of the human model to predict clearance when compared to interspecies extrapolation methods. Further development of the model, for example, through incorporation of additional physiological or biochemical processes currently not simulated, extension of the training set, and/or the use of additional inputs, will serve to further enhance the predictive capability of the model above methods that are dependent on in vivo animal experimentation.

The maximum value that can be extracted from the use of this model within drug discovery largely depends upon the modus operandi of an individual company, drug discovery program, or project. Significant factors that determine the relative merits of applying the model in any particular situation include the number of compounds passing through the successive stages of the discovery process, the
role of early ADME determination, access to low-cost ADME screens, the methods used for lead expansion, and the usage of in silico techniques. Assays are available, for each of the required model inputs, that have sufficiently high throughput to enable human PK prediction using measured values within the time scale dictated by successive synthesis rounds of a typical project in lead optimization. Recent, well documented changes in practice within the pharmaceutical industry have led to ADME data being generated more thoroughly and earlier in drug discovery than was previously the case. Hence, experimental measurement of some or all of the required inputs will be routinely available during lead optimization for many projects. Alternatively, in silico methods can be used to predict one or more inputs, or to reduce cost, increase throughput, and/or reduce the requirement for the physical compound, the availability of which may be restricted during early discovery. Consequently, we anticipate that the use of rapid in vitro/in silico screening of ADME and physico-chemical properties, coupled to the prediction of human PK through PBPK modeling, will enable the optimization of human PK to be a significant determinant of the lead expansion and optimization phases, rendering rat and mouse PK studies increasingly unnecessary.

The logical extension is that all the required inputs could be predicted by in silico methods, enabling virtual human PK screening. The PK of large numbers of virtual compounds can be simulated by the model over a relatively short time scale (simulating a 24-h time course takes a fraction of a second on a 1.3 GHz server with 1 GB of RAM running Red Hat Linux 7.2); calculation of the compound-dependent inputs, rather than execution of the PBPK model, is rate-limiting for some methods of input value prediction. In addition, due to the capability of the model to perform Monte-Carlo simulations, uncertainty in the values of the predicted inputs can be transformed into uncertainty in the predicted PK. This permits assessment of the associated risk when making assumptions based on any combination of predicted inputs and the uncertainty in those inputs. Thus, when using in silico prediction of its inputs, PBPK modeling has a potential role to play during lead identification, and even in prioritizing compounds to be passed through biological activity screens. The balance of how to use available computing resources most effectively, between the number of compounds to simulate and the number of Monte Carlo iterations to perform, depends on the number of compounds (virtual or real) under consideration, the number to be taken forward, the quality of the predicted model inputs, and the acceptable degree of uncertainty in PK when selecting compounds. It is important to note that the uncertainty in predicted PK that results from uncertainty in the inputs is highly dependent on the combination of input values for a compound. Thus, high uncertainty in CLint transforms to relatively low uncertainty in predicted PK for compounds whose unbound CLint (i.e., fup·CLint) is greater than hepatic blood flow, whereas for compounds with low CLint uncertainty in predicted AUC is nearly proportional to uncertainty in CLint.

During such a virtual screening process, care must be taken with the input prediction. Although commercial software is available for predicting many of the inputs to the PBPK model, the ability of such software to generalize with acceptable accuracy to drug discovery compounds is not guaranteed, because in most cases these compounds will lie outside the property space of the training data for the underlying models. Consequently, the suitability of predictive software should be determined by comparison with experimental data for a subset of the compounds to be screened. In those cases where reliability is not sufficient, possible alternative courses of action include building quantitative structure-property relationship models to make corrections to the predicted output for particular chemistry (some commercial software permits this local training) or developing custom-made quantitative structure-property relationship predictions. Fortunately, one of the most difficult properties to predict reliably, CLint, is one of the simplest to determine experimentally. This provides the potential for inexpensively generating a significant experimental clearance database, for use in PK prediction and in the development of QSPR models for predicting clearance of additional compounds.

As compounds progress toward preclinical development, the limitations of the current generic PBPK model, in both predictive reliability and the amount of information concerning the determinants of PK behavior that project teams typically require, are likely to become apparent. The potential for further model development, however, means that such restrictions need only be provisional. In principle, any process that affects PK can be incorporated into the model. The only limitations are the availability of relevant, usable data with which to modify the model and, if the changes require additional inputs, the availability of appropriate data to drive the model for novel compounds. As our understanding of the processes that govern PK increases, along with the availability of cost-effective in vitro techniques for their determination, so generic PBPK models can evolve to incorporate them. Within this scheme, the PBPK model can play an active role, informing in vitro and in silico method development regarding the accuracy and precision required for reliable in vivo prediction. We can, in time, expect to see developments in in vitro and in silico methods that enable successful PBPK-based prediction further along the discovery/development pipeline. Other potential developments of the PBPK model are less dependent on the provision of additional in vitro assays. These include improvements in the prediction of PK differences arising from sex, age, and body weight differences, permitting realistic interindividual variability to be simulated.

As we have already discussed (Brightman et al., 2006), the capacity to reliably predict mammalian PK from in vitro or in silico inputs is also of great potential benefit in assessing the risk posed to populations from exposure to environmental chemicals, without recourse to animal experimentation, and the consequent uncertainties in interspecies extrapolation to humans. Consequently, in the light of the results presented here, we conclude that the generic PBPK model can be a powerful, efficient, and cost-effective tool for xenobiotic PK prediction and reduction of in vivo animal experimentation in industry.

Acknowledgments. We thank Dr. Darwin Cheney for helpful comments on the manuscript.

References


Address correspondence to: Dr. David E. Leahy, Cyprotex Discovery Ltd., 13–15 Beech Lane, Macclesfield, Cheshire, United Kingdom, SK10 2DR. E-mail: d.leahy@cyprotex.com