DDT Vol. 6, No. 12 (Suppl.), 2001

Effect of diverse datasets on the predictive capability of ADME models in drug discovery

George M. Grass and Patrick J. Sinko

Accurate methods of predicting drug absorption in humans early in the drug discovery process are becoming increasingly important in the drive to evaluate chemical candidates for their product potential faster and cheaper than traditional experimentally based methods. The prediction of drug absorption in humans is used here as an example to describe the unique requirements and advantages of appropriately designed datasets in developing predictive models to simulate *in vivo* response from *in vitro* inputs. This approach is being successfully implemented in other ADME areas to develop a series of models that will ultimately predict bioavailability using chemical structure inputs.

***George M. Grass**
Lion Biosciences
9880 Campus Point Drive
San Diego, CA 92121 USA
*fax: +1 708 810 9345
e-mail: grass@sierra.net
Patrick J. Sinko
Department of Pharmaceutics, College of Pharmacy, Rutgers University
160 Frelinghuysen Road
Piscataway, NJ 08854
USA

▼ During the 1990s, drug discovery entered an age of industrialization. The creation and investigation of limited numbers of compounds that were produced manually and evaluated in low-throughput assays gave way to large-scale synthesis and HTS technologies. Although advancements in combinatorial chemistry and HTS have been very effective at producing 'hits', that is, identifying structures that demonstrate activity against a particular target of interest, they have not yet resulted in an overall increase in effectiveness or efficiency in bringing new drug candidates to market.

Part of the reason for this inability to progress early lead-candidates efficiently is that they do not necessarily demonstrate acceptable properties that will ultimately make them valuable and successful product candidates. Characteristics such as potency, selectivity, potential for oral absorption, rate of metabolism, and potential to cause toxic effects, are all properties that require evaluation of, and might require modification to, the original lead structure found in primary screening (i.e. lead optimization). Newer computational technologies aim to accelerate the process by decreasing the actual number of experiments required to make the crucial decision of following a product candidate through clinical trials.

The current drug discovery paradigm

Only a small fraction of New Chemical Entities (NCEs) will eventually advance to clinical trials, and approximately 80% of those will fail to become products[1,2]. The traditional drug discovery–development paradigm has focussed on the selection of 'winners' – NCEs with a high chance of success in the clinic. However, the selection and elimination of potential failures is also important in today's resource-constrained environment. Statistics show that approximately 75% of the total cost of drug development – estimated to be greater than US$500 million per compound in the 1980s – was attributed to development failures[3]. During the same period, the percentage of successful development candidates compared to the total number of candidates entering development was low and reported to be only 7.4%, 10% and 15.4% for Switzerland, the US and the UK, respectively[4–7]. From the mid-1960s to the mid-1980s, the average development times for NCEs increased four-fold with a consequent reduction in the effective patent life[5]. More recently, Lehman Brothers estimated the total cost of successful drug product development to be US$608 million per compound, suggesting that the process has not become more efficient[8].

The fundamental reasons for the failure of NCEs have also been examined. Prentis observed that during the 20-year period of 1968–1988, the majority of NCEs were dropped from development because of pharmacokinetic difficulties and a lack of efficacy in humans[5]. However, the failure

1359-6446/01/$ – see front matter PII: S1471-1931(01)00150-1

rate resulting from pharmacokinetic problems could actually be greater than that which was reported because poor pharmacokinetic properties, such as lack of absorption, rapid metabolism or elimination, and unfavorable distribution properties might be clinically manifested as a lack of efficacy. Although the quantitative values of each category might be different in current drug discovery efforts, poor pharmacokinetic properties are still recognized as significant contributors to compound failure[9–11].

The entire discovery lead-optimization process is a multi-dimensional problem that aims to find active, potent and selective compounds that can be orally absorbed, have ideal metabolism, distribution and elimination characteristics, and are non-toxic in humans. Currently, these multiple properties are routinely evaluated in a sequential fashion, the concept being that a very large number of compounds can be filtered down to a smaller number of potential lead candidates. This sequential solution of a multi-dimensional problem is inefficient at best, frequently difficult and time consuming, and often unsuccessful.

Box 1. Absorption versus bioavailability: definition

In a thorough review of the subject of absorption, Sietsema noted that the terms absorption and bioavailability are often incorrectly and interchangeably used[a]. Sietsema defined absorption as, 'The drug passing from the lumen of the gastrointestinal (GI) tract into the tissue of the GI tract. Once in the tissue, the drug is considered absorbed'.

Sinko[b] has proposed an updated definition for absorption because more-recent scientific advances have demonstrated that drug elimination from intestinal tissues, other than for drug absorbed into blood or lymph, can occur by two mechanisms: metabolism[c] and secretion (i.e. efflux into lumen)[d]. Therefore, the passing of drug into tissues is commonly referred to as drug 'uptake'. A more appropriate definition of oral drug absorption is, 'The amount of drug that passes through the intestinal tissues and enters into the portal vein'. In order for a drug to be bioavailable, it must pass through the liver and the lungs before it reaches general circulation[a]. Therefore, the difference between absorption and bioavailability is the amount of drug eliminated by means of secretion or metabolism on the first pass through the liver and lungs.

References

a Sietsema, W.K. (1989) The absolute oral bioavailability of selected drugs. *Int. J. Clin. Pharmacol. Ther. Toxicol.* 27, 179–211

b Sinko, P.J., ed. (1999) *Screening acceptable pharmacokinetic qualities in early drug discovery*, Current Drugs

c Sinko, P.J. *et al.* (1993) Mass balance approaches for estimating the intestinal absorption and metabolism of peptides and analogues: theoretical development and applications. *Pharm. Res.* 10, 271–275

d Makhey, V.D. *et al.* (1998) Characterization of the regional intestinal kinetics of drug efflux in rat and human intestine and in Caco-2 cells. *Pharm. Res.* 15, 1160–1167

While the pharmaceutical industry has been engaged in 'industrialization' of the early stages of drug discovery, such as the production of chemical libraries by combinatorial techniques and screening large numbers of compounds utilizing high-throughput approaches, other industries have engaged in the application of information technologies. With few exceptions, the technologies of the 'information age' have yet to be fully implemented and yet to have a significant impact on the pharmaceutical industry, especially the drug discovery process. The sequencing of the human genome has initiated what should be an unprecedented rate of new-target identification for human disease therapeutics to a pharmaceutical industry that is, for the most part, largely unprepared to capitalize fully on the opportunity to convert this much larger number of targets to drug products efficiently. This is evidenced by the fact that most drug discovery is conducted by mass experimentation rather than through computation.

The pharmaceutical industry is one of the few remaining major industries that develops new products primarily through a build-and-test method. The technological innovations of the past decade, specifically combinatorial chemistry and HTS, have been evolutionary rather then revolutionary to this process, that is, enabling companies to build more and test faster. Other major industries have moved away from the physical evaluation of large numbers of prototypes. The aerospace industry, for example, routinely uses computer simulation to both design and test new aircraft, enabling the investigation of many possible designs without physically producing them for testing. Drug discovery, however, remains focussed on searching ever-larger numbers of chemical candidates though sequential processes of physical testing, which encompass years of evaluation along with extremely high failure rates.

Although the processes involved in drug discovery are potentially more complex and less understood than those of other industries, solving these complex modeling challenges will lead to a significant change in the drug discovery process. One of the first areas of focus has been the prediction of pharmacokinetic parameters because bioavailability has become a rate-limiting assessment in the current approach to lead optimization.

Prediction of drug absorption

Drugs delivered orally for intestinal absorption represent the vast majority of dosage forms administered to human patients every year. Drug absorption is a complex process that is dependent upon numerous biochemical, physiological and physicochemical factors (Box 1). *In vivo* animal models have not been

DDT Vol. 6, No. 12 (Suppl.), 2001

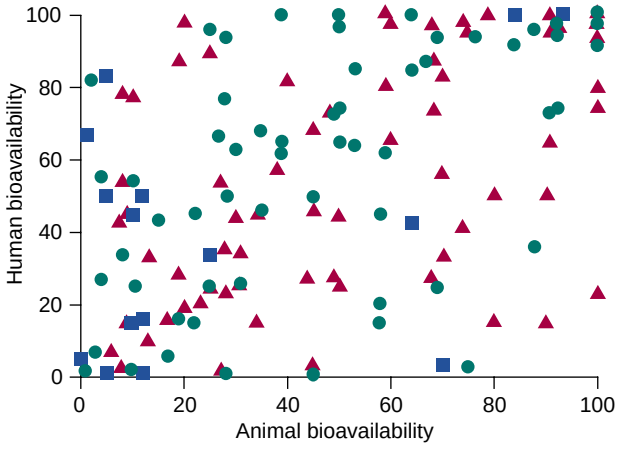

Figure 1. Absolute bioavailability of various drugs in dogs (red triangles), primates (blue squares) and rodents (green circles) versus the absolute bioavailability reported in humans. Three observations can be made: (i) there is no apparent relationship between animal bioavailability and human bioavailability; (ii) the number of false negatives is high; and (iii) the number of false positives is high. High drug bioavailability in animals does not necessarily translate into high human bioavailability. Athough the results show poor comparisons, gaining an understanding of the pharmacokinetics in animals is valuable in validating the pharmacokinetic behavior for toxicology studies that must be performed for regulatory agencies. These results were plotted from data in Ref. 12.

reliable predictors of bioavailability behavior in humans. Sietsema compared the absolute bioavailability of various drugs in dogs, primates and rodents with the absolute bioavailability reported in humans (Fig. 1)[12]. There are a number of observations that can be made from these data.

First, there is no apparent relationship between animal bioavailability and human bioavailability for any of the species examined – there is no one species that can be relied upon to predict pharmacokinetics in man accurately. The most likely reason for the lack of correlation is differences in physiology, metabolism and plasma protein binding. A review of the similarities and differences between species with respect to certain ADME properties has been published[13]. Animal studies attempt to correlate simultaneously all of the potentially independent parameters of an animal species and relate them all simultaneously to the same parameters in humans. For an individual compound, some parameters might correlate and some might not. Unfortunately, the group of non-correlating parameters that determine which drugs are not correctly predicted, are also not easily predicted.

Second, the number of false negatives and false positives is high. Numerous drug candidates would be either abandoned or their development delayed owing to low animal bioavailability even though the drugs were reported to have high human bioavailability. Furthermore, high drug bioavailability in animals does not necessarily translate into high human bioavailability, resulting in the potential to commit to the development of compounds that would ultimately fail to become products.

These observations suggest that it is prudent to exercise caution when interpreting animal bioavailability results for predicting human bioavailability. Better predictions of whether an NCE will fail or succeed in the clinic are more likely to be gauged by determining the factors that control oral bioavailability.

The problem of data

Ideally, predictive models for drug discovery would be based entirely on information derived from chemical structure. The prediction of clinically relevant properties (potency, selectivity, oral absorption, metabolism, distribution, excretion and toxicology) would be accomplished by providing a description of a chemical structure to a properly trained computer program. The attractiveness of this type of model is, of course, the ability to screen large numbers of chemical structures *in silico*, without the need to conduct experiments. In fact, this approach would enable virtual screening of chemicals without physically creating the molecules. Although this remains the ultimate goal of modeling efforts in drug discovery, there is a substantial hurdle to the construction of such a model: the acquisition of the necessary amounts of reliable and consistent data.

The requirement for high-throughput ADME screens has been reviewed by Sinko[14] and more recently discussed by Spalding *et al.* (Ref. 15). However, the information generated from these screens will be of little value if it cannot be related to the outcome *in vivo*. It is the *in vitro* to *in vivo* relationship that is the focus of this review, using the prediction of human drug absorption as an example. However, the principles that are described here for drug adsorption are equally applicable to and have been utilized in models of other important pharmacokinetic variables[16]. Even in cases where models to predict *in vitro* parameters from chemical structure have been developed[17,18] or the parameters are measured[19–24], these data cannot be utilized effectively until their influence upon the performance of drugs in humans is known and predictable.

The iDEA™ dataset for modeling absorption

A number of models have been proposed that attempt to relate *in vitro* parameters to human *in vivo* absorption[25–29]. Historically, these models have been compromised by several common factors that generally reflect the nature of the data used to develop or 'train' the models. These include:

- Relatively small datasets
- Datasets that are limited in scope; usually restricted to marketed compounds
- Datasets that are restricted to compounds with known and robust analytical procedures, or that are commercially available as radiolabeled material.

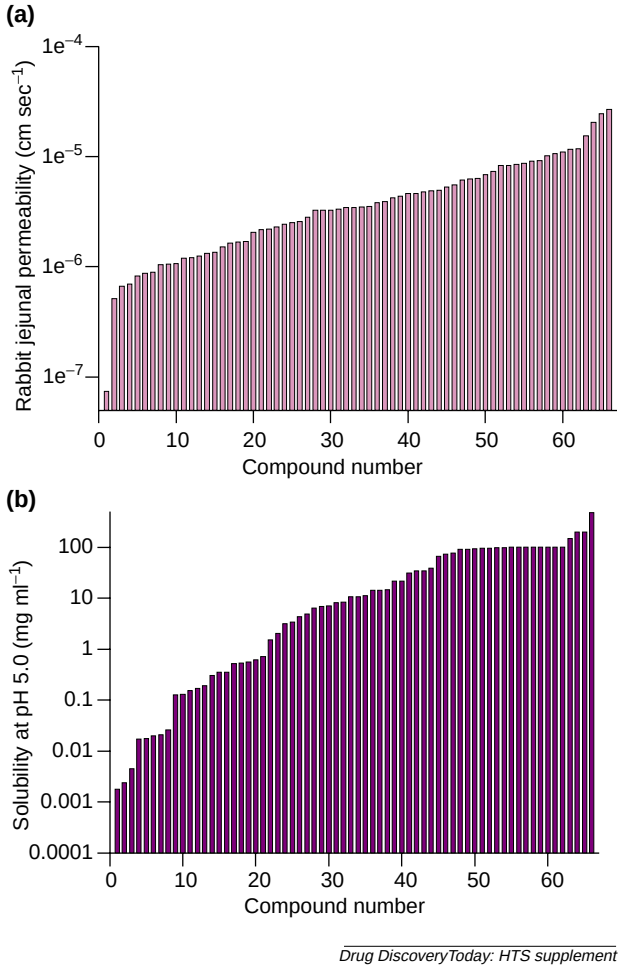

Figure 2. Ranges and distribution of (a) jejunal permeability (pink) and (b) solubility (purple) values of a unique set of compounds. The compounds were used to test published models and to develop a simulation model for absorption. The dataset consists of information for compounds that are not metabolized in humans and was obtained from a variety of published and unpublished sources and includes compounds that are marketed products and those that failed to become products.

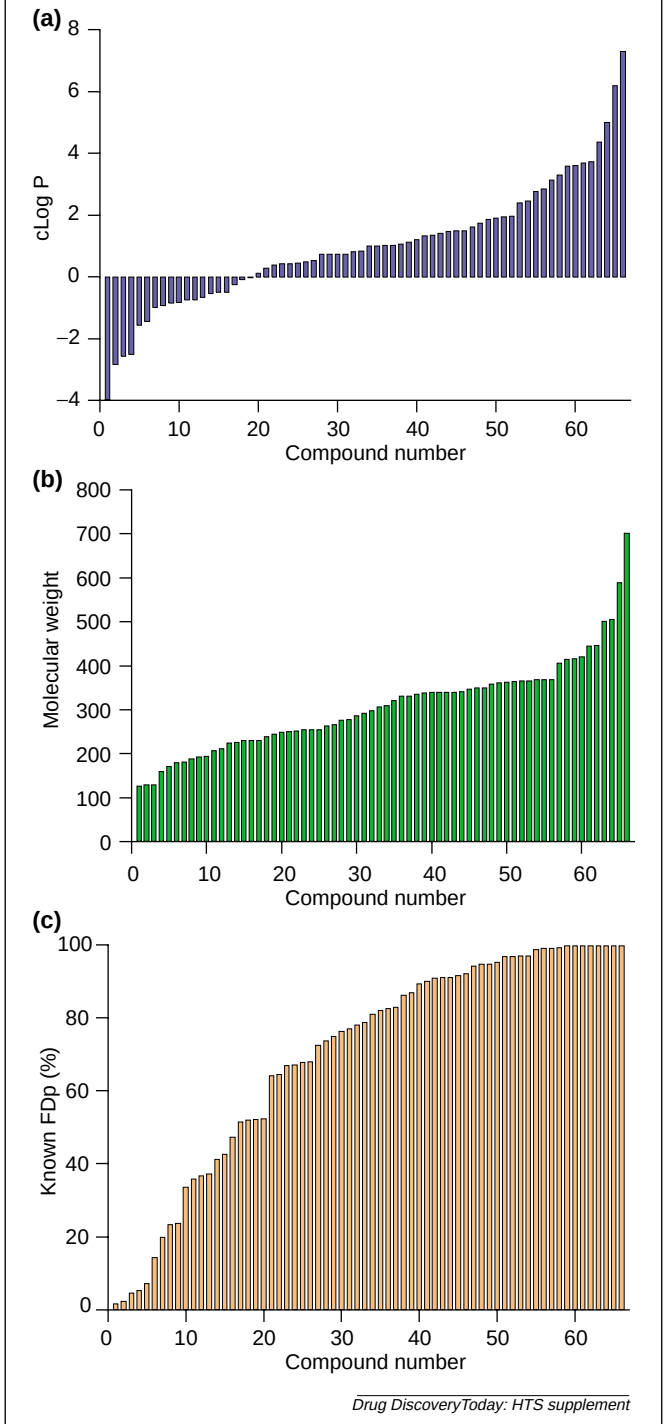

Figure 3. Ranges and distribution values of (a) cLog P (blue), (b) molecular weight (green), and (c) the fraction of drug absorbed to the portal vein in humans (FDp) (orange) of the dataset from Fig. 2.

To address these problems, a unique dataset was assembled through a consortium effort and termed the iDEA™ (*in silico* Determination for the Estimation of ADME) database[30]. Participating pharmaceutical companies supplied preclinical and clinical data for the development of a simulation model that would predict *in vivo* human absorption from *in vitro* data. Included in this dataset were compounds that had become marketed products and, more importantly, representative compounds that had failed in the development effort. The dataset contained compounds from a large number of different intended therapeutic applications and covered a broad diversity. All drugs were either not metabolized in man or had low hepatic clearance and were therefore not subject to significant first pass metabolism. Pharmacokinetic data available for each of the compounds was: (i) plasma concentration versus time curves, following both oral and intravenous administration to healthy human subjects; and (ii) data from human mass balance studies. An

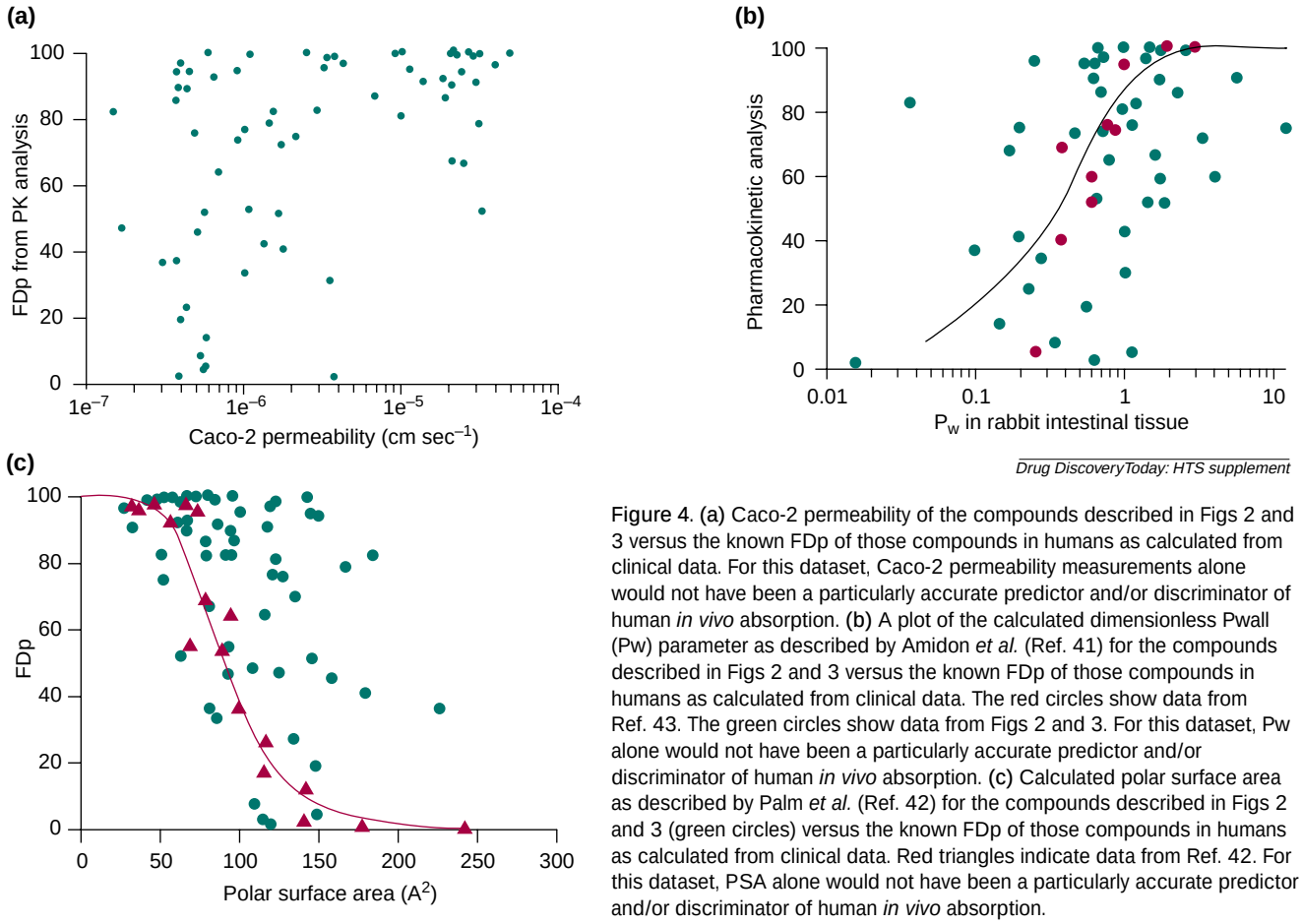

Figure 4. (a) Caco-2 permeability of the compounds described in Figs 2 and 3 versus the known FDp of those compounds in humans as calculated from clinical data. For this dataset, Caco-2 permeability measurements alone would not have been a particularly accurate predictor and/or discriminator of human *in vivo* absorption. (b) A plot of the calculated dimensionless Pwall (Pw) parameter as described by Amidon *et al.* (Ref. 41) for the compounds described in Figs 2 and 3 versus the known FDp of those compounds in humans as calculated from clinical data. The red circles show data from Ref. 43. The green circles show data from Figs 2 and 3. For this dataset, Pw alone would not have been a particularly accurate predictor and/or discriminator of human *in vivo* absorption. (c) Calculated polar surface area as described by Palm *et al.* (Ref. 42) for the compounds described in Figs 2 and 3 (green circles) versus the known FDp of those compounds in humans as calculated from clinical data. Red triangles indicate data from Ref. 42. For this dataset, PSA alone would not have been a particularly accurate predictor and/or discriminator of human *in vivo* absorption.

early version of the model resulting from this work has been previously described[30]. Figs 2 and 3 show the diversity of this dataset for solubility, permeability, clogP, molecular weight and percentage absorption.

Pharmacokinetic analysis

For each compound in the database, best-fit curves were determined for the intravenous (IV) and oral plasma (PO) concentration of each drug compound at each dosage level in order to estimate the fraction of dose absorbed to the portal vein (FDp). A two-compartment disposition model with elimination from the central compartment was used to fit the IV and PO curves simultaneously using weighted non-linear regression. The oral input function was fitted to the equation:

$$FDp\,(t) = \frac{D\,FDp}{2}\left(1 - erf\frac{1-\frac{t}{t_{50}}}{\frac{1}{P_{ec}}\sqrt{\frac{t}{t_{50}}}}\right) \quad (1)$$

where $FDp(t)$ is the fraction dose absorbed to the portal vein at time t; D is the dose; FDp is the fraction dose absorbed to the portal vein at infinity; t_{50} is the time for 50% of the dose to be absorbed; and P_{ec} is a parameter related to the slope of the error function (erf).

Rabbit intestinal permeability assay

Compound permeabilities through the duodenum, jejunum, ileum and distal colon were determined in the apical → basolateral direction using vertical Ussing-type, low-volume diffusion chambers and compound permeability was calculated using the equation:

$$P_e = \frac{V}{AC_0}\frac{dC}{dt} \quad (2)$$

where P_e is the effective permeability in cm sec^{-1}; V is the receiver chamber volume in mls; A is the surface area available for transport (cm^2), C_0 is the donor drug concentration; and dC/dt is the slope of the best fit line through the concentration versus time profile in the receiver chamber.

Caco-2 effective permeability assay

Effective permeability (Peff) was measured in the apical → basolateral direction in 20–23-day-old Caco-2 cell cultures and calculated using the following formula:

$$Peff(\mathrm{cm\,sec^{-1}}) = \frac{dX/dt}{AC_0 60} \quad (3)$$

where X is the mass transported; A is the surface area; and C_o is the initial donor drug concentration.

Comparison to other models

This consortium-derived dataset was used to examine the predictive capability of several published models, which included direct correlation of percentage absorption in humans to: (i) Caco-2 permeability[31,32] (Fig. 4); (ii) the rat intestinal perfusion model (Pwall)[33] (Fig. 4); and (iii) calculated polar surface area (PSA)[34,35] (Fig. 4). Other models have recently been developed by both larger pharmaceutical companies[16] and independent technology developers[36]. These include the GastroPlus from Simulations Plus [Lancaster, CA, USA (http://www.simulations-plus.com)], and models from Camitro (Menlo Park, CA, USA (http://www.camitro.com)] and Cyprotex [Manchester, UK (http://www.medeval.com)]. The unpublished algorithms for these models are not accessable nor does there presently appear to be published independent validations of these models. However, the aim of this review is to highlight the importance of compiling an appropriate dataset to train such models to give greater preditive capability, rather than to compare the predictive capability of all existing models.

Each compound in the iDEA™ database was evaluated for the appropriate experimentally determined or calculated parameter for each of the three methods available for comparison. For example, Caco-2 values were generated for the compounds in the dataset and then plotted to determine if a relationship exists as described with the dataset published for that method[37].

When the iDEA™ training set was examined using the previously published methods to predict oral drug absorption, poor correlations resulted (Fig. 4). The reason for the poor performance is probably because the published models were based on relatively limited datasets, often comprising compounds that could be easily obtained and analyzed. When these models are then applied to a diverse dataset, the limitations of their smaller and less diverse training sets results in poor correlation.

The inability of models with relatively small datasets to predict compounds that are not part of the training set has been observed by others. For example, Ren and Lieu have described the difficulty in using Caco-2 as a sole predictor of human absorption for a set of compounds that is more diverse than the originally published training set[38]. This is a manifestation of another common problem with these types of models, which is that the dataset used to develop the model is often the data used to test or evaluate the model, so the models are only internally validated. Although internally validated models can provide a substantial basis for the mechanistic evaluation and understanding of a process, they do not represent the type of validation one would like to have when evaulating NCEs and making the crucial decision to promote a compound to development or to terminate its development.

External validation of model performance

Simulation models that use *in vitro* data to predict oral drug absorption in humans have been previously described, but were limited in predictive capability because there were no comprehensive datasets available to enable adequate training and testing of such models[39]. More comprehensive models such as the iDEA™ simulation system have been recently described[30].

The performance of the iDEA™ simulation system for predicting oral absorption in humans was validated in a blinded test using an external dataset. This validation was conducted in conjunction with F. Hoffmann-La Roche, who were not part of the original consortium of companies that supplied the data. The compounds from Roche were supplied for testing with no information other than their molecular weights (used for LC–MS assay analysis). Chemical structure, clinical data, and therapeutic intent were not made known until after predictions of FDp were completed.

This external set of drugs varied in their physical and chemical properties, e.g. molecular weight (between 200 and 552); solubility (22 ng ml^{-1} to >100 mg ml^{-1}); permeability (rabbit: $0.14–25.0 \times 10^{-6}$ cm sec^{-1}; Caco-2: 0.43 to 32.5×10^{-6} cm sec^{-1}); transport process (including actively effluxed compounds)[40] and pharmacokinetic properties (%FDp between 4% and 100%). *In vitro* permeability values were generated as already described and, together with measured aqueous solubility measurements, were used as inputs to the model to give a prediction of FDp for each compound. These were then compared to the observed values from human clinical data (Fig. 5).

The demonstrated predictive capability of the model in Fig. 5 is the result of several factors – the diversity of the training set used to build the model and the guided nature of the creation of the model (i.e. the development of the model was not solely data driven). The relationships described by the model have been defined by the known physiology of the system. The model has therefore been optimized through a combination of a known empirical understanding of the system and the influence of the data to optimize the parameters of that system. This differs substantially from typical neural network

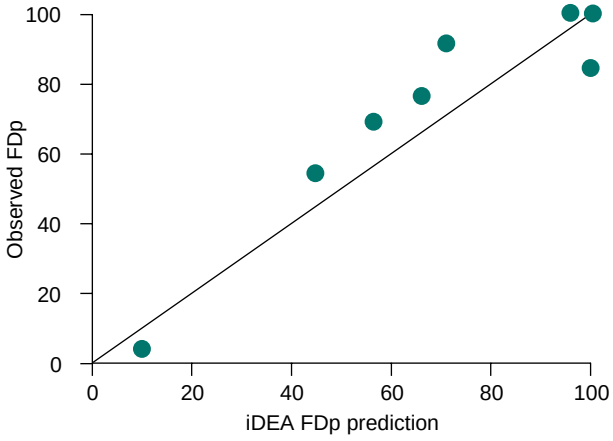

Figure 5. Predicted FDp values of a group of compounds versus the observed FDp as calculated from clinical data. The dataset represents a set of compounds submitted as a blinded external validation set which was used to test a simulation model for oral drug absorption in humans which was trained using the dataset described in Figs 2 and 3. The solid line represents a unity line (1:1 correlation).

solutions where the data drives the relationships within the model, which might have no physiological correlates.

Conclusions

Drug selection is now widely viewed as an important and relatively new, yet largely unsolved, bottleneck in the drug discovery and development process. In order to achieve an efficient selection process, high-quality, rapid, predictive and correlative ADME models are required. However, the requirement for such models to be accurate is perhaps even more important. Although the level of accuracy required can be dependent upon when a model is used in the discovery timeline (with less accuracy required earlier in the process), when projects move towards the decision of whether to proceed to clinical trials, predictive tools need to be accurate in order to be confidently used to support such financially crucial decisions. Systems that can be relied upon to accurately predict performance in humans have not existed and decisions have been made using tools whose capabilities could not be verified until candidates went to clinical trial, leading to the high failure rates historically observed. However, with the sequencing of the human genome, advances in proteomics, the anticipation of the identification of a vastly greater number of potential targets for drug discovery, and the potential of pharmacogenomics to require individualized evaluation of drug kinetics as well as drug effects, accurate predictive tools hold the promise of speeding up the decision-making process and significantly reducing the cost of getting new drugs to the market.

References

1 PhRMA (2000) *Why do drugs cost so much?*

2 Beary, J.F. *et al.* (1993) Basic research and the cost of health care. *Science* 262, 1358

3 Abelson, P.H. (1993) Improvements in health care. *Science* 260, 11

4 Prentis, R.A. and Walker, S.R. (1986) Trends in the development of new medicines by UK-owned pharmaceutical companies (1964–1980). *Br. J. Clin. Pharmacol.* 21, 437–443

5 Prentis, R.A. *et al.* (1988) Pharmaceutical innovation by the seven UK-owned pharmaceutical companies (1964–1985). *Br. J. Clin. Pharmacol.* 25, 387–396

6 Mattison, N. *et al.* (1984) The development of self-originated new drugs by Swiss pharmaceutical firms, 1960–1980. *Regul. Toxicol. Pharmacol.* 4, 157–173

7 Wardell, W.M. and DiRaddo, J. (1980) The measurement of pharmaceutical innovation. *J. Clin. Pharmacol.* 20, 1–9

8 Peakman, T. and Arlington, S. (2001) Putting the code to work: the promise of pharmacogenetics and pharmacogenomics. *Drug Discov. World* 2, 35–40

9 Borchardt, R.T. *et al.* (1996) General principles in the characterization and use of model systems for biopharmaceutical studies. *Pharm. Biotechnol.* 8, 1–11

10 Lesko, L.J. *et al.* (2000) Optimizing the science of drug development: opportunities for better candidate selection and accelerated evaluation in humans. *J. Clin. Pharmacol.* 40, 803–814

11 Peck, C.C. *et al.* (1994) Opportunities for integration of pharmacokinetics, pharmacodynamics, and toxicokinetics in rational drug development. *J. Clin. Pharmacol.* 34, 111–119

12 Sietsema, W.K. (1989) The absolute oral bioavailability of selected drugs. *Int. J. Clin. Pharmacol. Ther. Toxicol.* 27, 179–211

13 Lin, J.H. (1995) Species similarities and differences in pharmacokinetics. *Drug Metab. Dispos.* 23, 1008–1019

14 Sinko, P.J., ed. (1999) *Screening acceptable pharmacokinetic qualities in early drug discovery*, Current Drugs

15 Spalding, D.J.M. *et al.* (2001) Combining high-throughput pharmacokinetic screens at the hits-to-leads stage of drug discovery. *Drug Discov. Today* 5 (Suppl.), S70–S76

16 Poulin, P. and Theil, F.P. (2000) *A priori* prediction of tissue: plasma partition coefficients of drugs to facilitate the use of physiologically-based pharmacokinetic models in drug discovery. *J. Pharm. Sci.* 89, 16–35

17 Camenisch, G. *et al.* (1998) Estimation of permeability by passive diffusion through Caco-2 cell monolayers using the drugs' lipophilicity and molecular weight. *Eur. J. Pharm. Sci.* 6, 317–324

18 Norinder, U. *et al.* (1997) Theoretical calculation and prediction of Caco-2 cell permeability using MolSurf parametrization and PLS statistics. *Pharm. Res.* 14, 1786–1791

19 Stewart, B.H. *et al.* (1995) Comparison of intestinal permeabilities determined in multiple *in vitro* and *in situ* models: relationship to absorption in humans. *Pharm. Res.* 12, 693–699

20 Stewart, B.H. and Chan, O.H. (1998) Use of immobilized artificial membrane chromatography for drug transport applications. *J. Pharm. Sci.* 87, 1471–1478

21 Dowty, M.E. and Dietsch, C.R. (1997) Improved prediction of *in vivo* peroral absorption from *in vitro* intestinal permeability using an internal standard to control for intra- and inter-rat variability. *Pharm. Res.* 14, 1792–1797

22 Jezyk, N. *et al.* (1992) Permeability characteristics of various intestinal regions of rabbit, dog, and monkey. *Pharm. Res.* 9, 1580–1586

23 Grass, G.M. and Sweetana, S.A. (1989) A correlation of permeabilities for passively transported compounds in monkey and rabbit jejunum. *Pharm. Res.* 6, 857–862

24 Smith, P.L. (1996) Methods for evaluating intestinal permeability an metabolism *in vitro*. *Pharm. Biotechnol.* 8, 13–34

25 Artursson, P. and Karlsson, J. (1991) Correlation between oral drug absorption in humans and apparent drug permeability coefficients in human intestinal epithelial (Caco-2) cells. *Biochem. Biophys. Res. Commun.* 175, 880–885

26 Winiwarter, S. *et al.* (1998) Correlation of human jejunal permeability (*in vivo*) of drugs with experimentally and theoretically derived parameters. A multivariate data analysis approach. *J. Med. Chem.* 41, 4939–4949

27 Gres, M.C. *et al.* (1998) Correlation between oral drug absorption in humans, and apparent drug permeability in TC-7 cells, a human epithelial intestinal cell line: comparison with the parental Caco-2 cell line. *Pharm. Res.* 15, 726–733

28 Yu, L.X. and Amidon, G.L. (1999) A compartmental absorption and transit model for estimating oral drug absorption. *Int. J. Pharm.* 186, 119–125

29 Sinko, P.J. *et al.* (1991) Predicting fraction dose absorbed in humans using a macroscopic mass balance approach. *Pharm. Res.* 8, 979–988

30 Norris, D.A. *et al.* (2000) Development of predictive pharmacokinetic simulation models for drug discovery. *J. Control. Release* 65, 55–62

31 Artursson, P. (1990) Epithelial transport of drugs in cell culture. I: A model for studying the passive diffusion of drugs over intestinal absorptive (Caco-2) cells. *J. Pharm. Sci.* 79, 476–482

32 Artursson, P. and Karlsson, J. (1991) Correlation between oral drug absorption in humans and apparent drug permeability coefficients in human intestinal epithelial (Caco-2) cells. *Biochem. Biophys. Res. Commun.* 175, 880–885

33 Amidon, G.L. *et al.* (1988) Estimating human oral fraction dose absorbed: a correlation using rat intestinal membrane permeability for passive and carrier-mediated compounds. *Pharm. Res.* 5, 651–654

34 Palm, K. *et al.* (1996) Correlation of drug absorption with molecular surface properties. *J. Pharm. Sci.* 85, 32–39

35 Palm, K. *et al.* (1998) Evaluation of dynamic polar molecular surface area as predictor of drug absorption: comparison with other computational and experimental predictors. *J. Med. Chem.* 41, 5382–5392

36 Egan, W.J. *et al.* (2000) Prediction of drug absorption using multivariate statistics. *J. Med. Chem.* 43, 3867–3877

37 Artursson, P. (1990) Epithelial transport of drugs in cell culture. I: A model for studying the passive diffusion of drugs over intestinal absorptive (Caco-2) cells. *J. Pharm. Sci.* 79, 476–482

38 Ren, S. and Lien, E.J. (2000) Caco-2 cell permeability vs human gastrointestinal absorption: QSPR analysis. *Prog. Drug. Res.* 54, 1–23

39 Grass, G.M. (1997) Simulation models to predict oral drug absorption from *in vitro* data. *Adv. Drug Deliv. Rev.* 23, 199–219

40 Makhey, V.D. *et al.* (1998) Characterization of the regional intestinal kinetics of drug efflux in rat and human intestine and in Caco-2 cells. *Pharm. Res.* 15, 1160–1167

41 Amidon, G. *et al.* (1988) Estimating human oral fraction dose absorbed: a correlation using rat intestinal membrane permeability for passive and carrier-mediated compounds. *Pharm. Res.* 5, 651–654

42 Palm, K. *et al.* (1997) Polar molecular surface properties predict the intestinal absorption of drugs in humans. *Pharm. Res.* 14, 568–571