Prediction of intestinal absorption: comparative assessment of gastroplus™ and idea™

Neil Parrott

European Journal of Pharmaceutical Sciences

ELSEVIER

European Journal of Pharmaceutical Sciences 17 (2002) 51–61

EUROPEAN JOURNAL OF PHARMACEUTICAL SCIENCES

www.elsevier.nl/locate/ejps

Prediction of intestinal absorption: comparative assessment of GASTROPLUS™ and IDEA™

Neil Parrott*, Thierry Lavé

F. Hoffmann-La Roche AG, Pharmaceuticals Division, CH-4070 Basel, Switzerland

Received 17 April 2002; received in revised form 9 July 2002; accepted 17 July 2002

Abstract

We have assessed two commercial software tools employing physiologically based models for prediction of intestinal absorption in human. IDEA™ 2.0 and GASTROPLUS™ 3.1.0 were compared both in their ability to predict fraction absorbed for a set of 28 drugs and in terms of the functionality offered. The emphasis was placed on the practical usefulness to pharmaceutical drug discovery. Predictions were assessed for three levels of input data (i) pure in silico input, (ii) thermodynamic solubility and in silico permeability, (iii) thermodynamic solubility and human colon carcinoma cell line (CACO-2) permeability. We found the pure in silico prediction ability of the tools to be comparable with 70% correct classification rate. With measured input data the IDEA™ prediction rate improved to 79% while GASTROPLUS™ stayed at 70%. In terms of functionality GASTROPLUS™ is a powerful system for the trained user. Open access to model parameters, diagnostic tools and the ability to integrate data make it particularly suitable for the later stages of discovery and development. IDEA™ is web based and presents a simple interface suitable for widespread use with minimal training. However the limited functionality and inconvenient handling of multiple compound batches currently restrict the usefulness of version 2.0 for drug discovery.

Keywords: ADME; Oral absorption; Physiologically based pharmacokinetics; Simulation; Modeling

1. Introduction

The ability to be administered by the oral route is a highly desirable property for new pharmaceutical drugs because it is the safest, most convenient and economical method (Goodman et al., 1999). For this reason good oral availability is a required property for drug candidate molecules in a large percentage of pharmaceutical discovery projects. It has been observed that drug development often failed for reasons of poor pharmacokinetics (Prentis et al., 1988). To avoid the high costs associated with such failures the current practice is to consider metabolism and pharmacokinetic properties in parallel with pharmacological tests during the discovery phase. High throughput in vitro technology now allows properties of importance for oral absorption like solubility, permeability, lipophilicity and pK_a to be measured early and for many compounds. Consequently the discovery scientist is presented with large volumes of multivariate data and is faced with considerable data integration and information generation challenges. In particular a need for reliable models of oral absorption exists.

The prediction of in vivo absorption is complex and the number of factors to be considered large. Absorption of drugs from the gastrointestinal tract can be influenced by physicochemical, physiological and formulation factors. The physicochemical factors include pK_a, solubility, stability, lipophilicity, and salt forms. The physiological factors include gastrointestinal pH, gastric emptying, small and large bowel transit times, active transport and efflux, and gut wall metabolism. The formulation factors are related to drug particle size, crystal form and dosage forms such as solution, tablet, capsule or suspension.

Considering the complexity of absorption and the number of processes involved, an integrated approach taking most of the data into account is highly desirable. Physiologically based models provide a rational basis for integration of data and can predict both extent and rate of absorption.

A further step forward is the pure in silico approach where measured in vitro data is replaced with properties predicted from chemical structure alone. Input of in silico predictions into validated predictive models then gives the

*Corresponding author. Tel.: +41-61-688-0813.
E-mail address: neil_john.parrott@roche.com (N. Parrott).

possibility for absorption potential to be assessed before compounds are synthesized. With access to such tools the medicinal chemist is in the position to synthesize only compound collections with a good chance of being well absorbed.

Two physiologically based models to predict oral absorption recently became commercially available. These models are IDEA™ (in vitro determination for the estimation of ADME) From Lion Bioscience (LION Bioscience) and the ACAT (advanced compartmental absorption and transit) model available in GASTROPLUS™ from SIMULATIONS PLUS (Simulations Plus[1]).

In this study we describe a comparative evaluation aimed at assessing the predictiveness of these tools when used with the typical data available to drug discovery scientists. Our evaluation also discusses usability and functionality.

2. Methods

Brief overviews of GASTROPLUS™ and IDEA™ are presented here. For more detailed information we refer the reader to the web sites of the respective companies.

2.1. GASTROPLUS™ 3.1.0

GASTROPLUS™ simulates gastrointestinal absorption and pharmacokinetics for drugs dosed orally or intravenously in humans and animals. The simulation model underlying GASTROPLUS™ is known as the advanced compartmental absorption and transit model (ACAT) (Agoram et al., 2001). The ACAT model of SIMULATIONS PLUS is based upon an original CAT model described by Yu et al. (1996) and is semiphysiological with nine compartments corresponding to different segments of the digestive tract (stomach, seven small intestinal compartments and colon). In addition to human physiology, models for other species (rat, dog, rabbit or cat) are provided. The model accounts for controlled-release profiles, pH dependence of dissolution and permeability, transport of drug material through the gastrointestinal tract and absorption of drug material through the intestinal wall into the portal vein. GASTROPLUS™ provides the ability for customization by allowing absorption constants to be set individually for each intestinal compartment. In addition an optimization module permits fitting of model parameters such as regional permeabilities, physiological variables and formulation factors to observed data. For early drug discovery a generic log *D* model for regional permeability is based upon the premise that as the ionized fraction of drug increases permeability decreases. Therefore permeability in each compartment is scaled according to the pH of that compartment the log *P* and the pK_a values of the drug. The exact function for adjusting compartmental permeabilities was optimized by SIMULATIONS PLUS to explain the observed rate and extent of absorption for a proprietary training set of drugs. The log *D* model and other default settings are recommended when sufficient data needed for construction of drug-specific models are not available (sufficient data might constitute measured regional permeabilities or in vivo concentration versus time data from multiple doses). For pure in silico predictions it is necessary to use a companion product, QMPRPLUS™, that takes an input file of multiple structures and computes properties including solubility, permeability and log *P*. Models based upon 2D or 3D structures are available (in this evaluation we used only 2D). The model for human permeability is based on a combination of in vivo human values and in situ rat wall permeabilities converted to human values based upon a correlation. The model was constructed using partial least square regression on a training set of 47 drugs ($N=47$, $R^2=0.76$, RMSE=0.29). Several solubility models are available. We used the model where the only input is 2D chemical structure (an artificial neural network trained on 1204 examples ($N=1204$, $R^2=0.943$, RMSE=0.47 log units)).

GASTROPLUS™ 3.1.0 and QMPRPLUS™ 2.2 were the versions evaluated. The latest versions GASTROPLUS™ 3.1.1a and QMPRPLUS™ 3.0 were released in February and April 2002 and provide enhanced functionality but without major changes to the models.

2.2. IDEA™ 2.0

IDEA™ simulates human physiology and accounts for regional variations in intestinal permeability, solubility, surface area and fluid movement. The IDEA™ absorption module runs as a web application on the corporate Intranet and is aimed specifically at facilitating lead selection early in drug discovery.

The simulation model underlying IDEA™ is based upon published work by Grass (1997). Subsequently work was done within Lion to develop and train the model with input from a consortium of pharmaceutical companies. Consortium members supplied oral and intravenous clinical data while Lion generated a database of in vitro data for a training set of around 70 nonmetabolized drugs (Norris et al., 2000). Many model details are kept proprietary and only briefly described in the IDEA™ reference manual.

In addition to the physiologically based absorption model IDEA™ 2.0 includes a structure-based model for in silico prediction of absorption class. This model uses statistical pattern recognition techniques trained on in vitro and clinical pharmacokinetic data for 121 drug compounds. A separate model for prediction of Caco-2 permeability was trained on Lion data for 250 drugs.

Detailed assay protocols for generation of in vitro data for use with the absorption model are described in the

[1]Simulations Plus, I., 1220 W. Avenue J, Lancaster, California 93534-2902, USA. http://www.simulations-plus.com/.

IDEA™ reference manual. In addition, a set of data for ten example drugs is provided so that results of company assays may be compared to the data of Lion.

In a previous evaluation of IDEA™ (Leesman et al., 2000) Roche provided eight drug samples as an external validation set for a blinded study and the predictions based upon the CACO-2 permeability model were within previously agreed acceptance criteria for seven of the eight compounds.

Version 2.0 of IDEA™ was evaluated. Version 2.2 is due for release in Q4 of 2002 [Lion, personal communication].

2.3. Input data used

Three different combinations of input data were assessed

(i) Pure in silico input for solubility and permeability (i.e. chemical structure alone as input)
(ii) thermodynamic solubility and in silico permeability and
(iii) thermodynamic solubility and measured CACO-2 permeability.

For a set for 28 drugs, data as above was generated in the standard screening assays used in our drug discovery projects. Additional common input was the clinical dose levels corresponding to the fraction absorbed reported in the literature (Table 1).

The drugs were selected for diversity in physicochemical properties and to cover the full range of fraction absorbed in man (Fig. 1). Practical considerations related to the availability of compound samples necessitated a bias towards well-absorbed drugs. For the 28 drugs the solubility at pH 6.5 covered the range from 0.005 to over >1000 mg/ml. The CACO-2 permeability covered a range from 0.2×10^{-6} to $>60\times10^{-6}$ cm/s.

At this point some explanation of the input requirements of the two software as well as the differences in the modes of operation is necessary. Table 2 compares the input data requirements while Figs. 2 and 3 illustrate the operation modes for pure in silico prediction and prediction where measured data for permeability and solubility are used.

Regarding the pure in silico predictions, one required input where estimates are not currently available from QMPRPLUS™ is pK_a. Therefore to complete the in silico data set for GASTROPLUS™ we used a separate tool pKalc (CompuDrug[2]). For IDEA™ the pure in silico capability, consisting of a classification into low, medium or high (≤33%, 33–66%, ≥66%), requires only chemical structure as input.

To facilitate a direct comparison it was necessary to ensure identical input to both tools. Since IDEA™ requires the input of solubility values measured at several pH values, the ability of GASTROPLUS™ to generate solubility versus pH profiles from a single measured value and a table of pK_a was not used. Rather solubility values at 5 pH values ranging from 1.5 to 7.5 were used for both softwares.

Concerning dose, GASTROPLUS™ takes a number of inputs describing dosage form and formulation whereas IDEA™ accepts just a single value (Table 2). In this evaluation we took the GASTROPLUS™ defaults namely

1. Dosage form: immediate release tablet
2. Dose volume: 250 ml
3. Drug particle density: 1.2 g/ml
4. Effective particle radius: 25 μm

For predictions based upon measured permeability, the GASTROPLUS™ model has been trained to accept values for human jejunal permeability as input. Thus a preliminary step when using CACO-2 data is transformation based upon a correlation. In this study we built a correlation of log human permeability against log CACO-2 permeability that included 20 drugs. The correlation coefficient was

Table 1
Clinical fraction absorbed data

Drug	Dose (mg)	Fraction absorbed (%)
Aciclovir	350	23
Amiloride	10	50
Antipyrine	600	97
Atenolol	50	50
Carbamazepine	200	70
Chloramphenicol	250	90
Desipramine	150	100
Diazepam	5	100
Diltiazem	90	90
Etoposide	350	50
Furosemide	80	61
Ganciclovir	75	3
Hydrochlorothiazide	50	69
Ketoprofen	75	92
Metoprolol	100	95
Naproxen	500	99
Penicillin V	200	38
Pirenzepine	50	27
Piroxicam	20	100
Progesterone	2.5	100
Propranolol	240	99
Ranitidine	60	63
Saquinavir	600	30
Sulpiride	200	44
Terbutaline	10	62
Theophylline	200	100
Verapamil	120	100
Warfarin	5	98

Where not otherwise indicated fraction absorbed is from Zhao et al. (2001). Diltiazem, etopside are from Dollery (1998), carbamazepine is from Goodman et al. (1999), pirenzepine is from Vergin (1986), saquinavir is from Hoffmann LaRoche (data on file).

[2]CompuDrug International, I., P.O.B. 160, Budapest, 1255, Hungary.

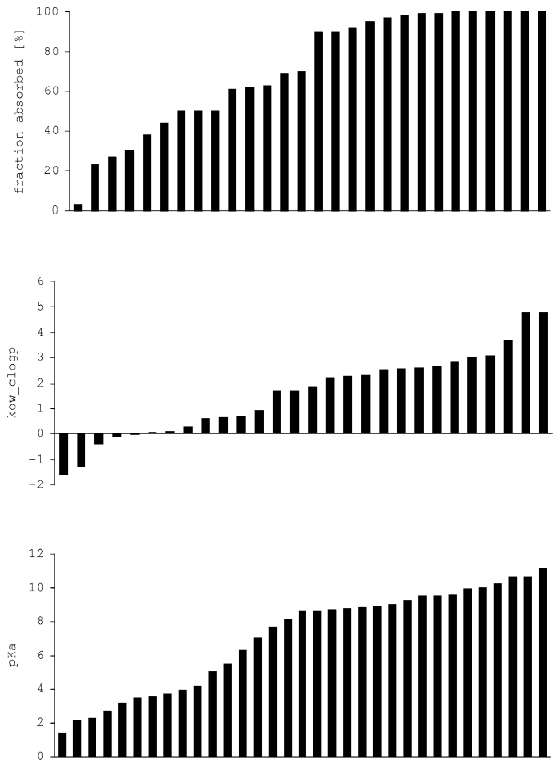

Fig. 1. Distribution of fraction absorbed in human, calculated log P and measured pK_a for the 28 drugs studied.

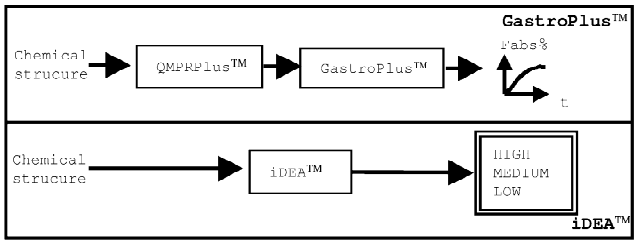

Fig. 2. Operation modes for pure in silico prediction.

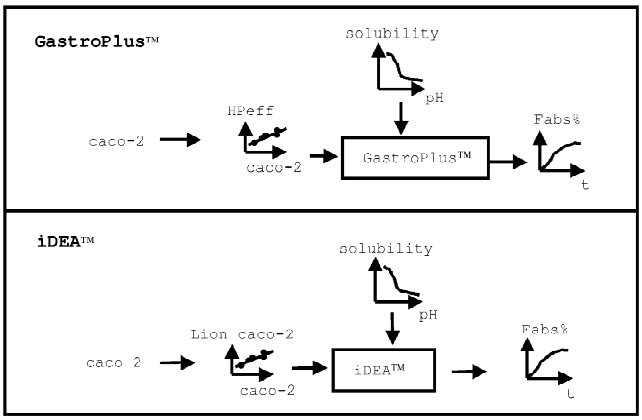

Fig. 3. Operation modes for prediction when measured data for permeability and solubility are used.

Table 2
Input data for GASTROPLUS™ and IDEA™

	GASTROPLUS™	IDEA™
Chemical structure	ISIS structure (or SMILES)	SMILES structure (or ChemDraw structure)
Dose and formulation	Initial dose (mg) Subsequent doses (mg) Dosing interval (h) Dose volume (ml) Drug particle density (g/ml) Effective particle radius (microns) Dosage form (selection from a list of options)	Dose (mg)
Solubility	Solubility at different pH values in the range 1.5–7.5 or solubility at one known pH plus a table of pK_a values	Solubility at different pH values in range 1.5–7.5
Permeability	Permeability measure that is transformed based on a correlation to human permeability or in silico estimate of human P_{eff}	Caco-2 (may need to be transformed using a correlation to Lion data) or rabbit intestinal tissue (four segments) plus permeability efflux ratio (optional)
pK_a	Table of pK_a values (used in the log D model of regional permeability)	
Lipophilicity	log D at known pH or log P	(Lipophilicity is predicted internally—no input is possible)

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.