RESEARCH ARTICLE

Comparative Assessment of *In Vitro–In Vivo* Extrapolation Methods used for Predicting Hepatic Metabolic Clearance of Drugs

PATRICK POULIN,[1] CORNELIS E. C. A. HOP,[2] QUYNH HO,[2] JASON S. HALLADAY,[2] SAMI HADDAD,[3] JANE R. KENNY[2]

[1]Consultant, 4009 Sylvia Daoust, Québec City, Québec G1X 0A6, Canada

[2]DMPK, Genentech Inc., South San Francisco, California 94080

[3]Département de Santé Environnementale et Santé au Travail, IRSPUM, Faculté de Médecine, Université de Montréal, Montréal, Québec H3T 1J4, Canada

Received 23 May 2012; revised 26 June 2012; accepted 17 July 2012

Published online in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/jps.23288

ABSTRACT: The purpose of this study was to perform a comparative analysis of various *in vitro--in vivo* extrapolation (IVIVE) methods used for predicting hepatic metabolic clearance (CL) of drugs on the basis of intrinsic CL data determined in microsomes. Five IVIVE methods were evaluated: the "conventional and conventional bias-corrected methods" using the unbound fraction in plasma (fu_p), the "Berezhkovskiy method" in which the fu_p is adjusted for drug ionization, the "Poulin et al. method" using the unbound fraction in liver (fu_{liver}), and the "direct scaling method," which does not consider any binding corrections. We investigated the effects of the following scenarios on the prediction of CL: the use of preclinical or human datasets, the extent of plasma protein binding, the magnitude of CL *in vivo*, and the extent of drug disposition based on biopharmaceutics drug disposition classification system (BDDCS) categorization. A large and diverse dataset of 139 compounds was collected, including those from the literature and in house from Genentech. The results of this study confirm that the Poulin et al. method is robust and showed the greatest accuracy as compared with the other IVIVE methods in the majority of prediction scenarios studied here. The difference across the prediction methods is most pronounced for (a) albumin-bound drugs, (b) highly bound drugs, and (c) low CL drugs. Predictions of CL showed relevant interspecies differences for BDDCS class 2 compounds; the direct scaling method showed the greatest predictivity for these compounds, particularly for a reduced dataset in rat that have unexpectedly high CL *in vivo*. This result is a reflection of the direct scaling method's natural tendency to overpredict the true metabolic CL. Overall, this study should facilitate the use of IVIVE correlation methods in physiologically based pharmacokinetics (PBPK) model. © 2012 Wiley Periodicals, Inc. and the American Pharmacists Association J Pharm Sci

Keywords: disposition; microsomes; clearance; unbound fraction; computational ADME; *in vitro–in vivo* extrapolation; *in vitro–in vivo* correlation; IVIVE; pharmacokinetics; PBPK modeling

INTRODUCTION

Various methods are available to predict human pharmacokinetics with some based on preclinical *in vivo* data and others utilizing human *in vitro* data. *in vitro* methods are most convenient because they require minimal amount of compound and do not necessitate animal studies. However, the predictivity of *in vitro* methods depends on the model used and the input

Abbreviations used: AAG, alpha1-acid glycoprotein; AFE, average-fold error; AL, albumin; BCS, bioclassification system; BDDCS, biopharmaceutics drug disposition classification system; CCC, concordance correlation coefficient; CL, clearance; CL_{int}, intrinsic CL; fu_{inc}, unbound fraction in incubation; fu_{liver}, unbound fraction in liver; fu_p, unbound fraction in plasma; $fu_{p\text{-}app}$, apparent unbound fraction in plasma; IVIVE, *in vitro–in vivo* extrapolation; K_m, Michaelis–Menten constant; PhRMA, Pharmaceutical Research and Manufacturers of America; PLR, plasma-to-whole liver concentration ratio of albumin; Q_{liver}, blood flow rate to liver; R_{BP}, blood-to-plasma concentration ratio; RMSE, root-mean-squared error.

Correspondence to: Dr. Patrick Poulin (Telephone: +418-802-3985; E-mail: patrick-poulin@videotron.ca)

parameters. Factors influencing the predictive performance of an *in vitro–in vivo* extrapolation (IVIVE) method for hepatic metabolic clearance (CL) are related to several input parameters; namely, binding terms such as the unbound fraction in plasma (fu_p) and in incubation medium (fu_{inc}) as well as the intrinsic CL (CL_{int}) and liver blood flow rate (Q_{liver}).[1–11] Wan et al.[1] studied the impact of these input parameters on the CL estimate in rat and human datasets. The authors concluded that the simplified IVIVE method, disregarding binding data (i.e., direct scaling), might be sufficiently good for IVIVE evaluations. Obach[2] also suggested disregarding all binding data to predict human CL for basic and neutral compounds, whereas for acidic compounds, he suggested including all binding terms (i.e., fu_p/fu_{inc}). Recently, Berezhkovskiy et al.[12,13] and Poulin et al.,[14] both of whom also studied the impact of the binding terms on CL estimates, presented two novel IVIVE methods. Berezhkovskiy's method consists of replacing fu_p with an apparent fu_p ($fu_{p\text{-}app}$) that considers drug ionization differences between the plasma and liver cells. Poulin et al.[14] further analyzed the concept of binding terms and suggested converting the value of $fu_{p\text{-}app}$ to an unbound fraction in the liver (fu_{liver}) to take also into account the role of extracellular binding proteins on the passive uptake of drugs in hepatocytes. Using a dataset of 25 drugs, the Poulin et al.[14] method showed the greatest accuracy as compared with other IVIVE methods on the basis of several statistical parameters.[14] Recently, Halifax and Houston[15] used a larger dataset and demonstrated superior precision and lower bias in the majority of cases for the novel method of Poulin et al.; however, these authors are not in total agreement on the mechanistic justification of the method advocated by Poulin et al.[14] Instead, Halifax and Houston[15] proposed an empirical scaling method involving a conventional model, but corrected for the average-fold error (AFE) (i.e., the conventional bias-corrected method). Therefore, a consensus on the use of IVIVE methods could not be agreed upon, and hence, further testing is needed.

The purpose of this study was to further investigate the published IVIVE methods by using large and diverse datasets from human, monkey, dog, and rat. This study might help to identify potential outlier drugs and apply further refined IVIVE methods to identify the strengths and limitations of these methods.

METHODS

The overall strategy consisted of evaluating the effect of the following scenarios on predictive performance of various IVIVE methods for CL based on microsomal data: (a) the preclinical and human datasets, (b) the extent of plasma protein binding [i.e., drugs bound to albumin (AL), drugs bound to alpha1-acid glycoprotein (AAG), and drugs highly bound in plasma], (c) the magnitude of CL under *in vivo* conditions (i.e., very low, low, medium, and high CL), and (d) the extent of drug disposition based on the biopharmaceutics drug disposition classification system (BDDCS) and/or bioclassification system (BCS) categorizations.[16] Furthermore, we explored the effect of hepatic uptake on CL estimations by using the current IVIVE methods for a reduced dataset of drugs in rats. For this dataset, the Poulin et al.[14] method was compared with the direct scaling method. We theorized that the direct scaling method may be advantageous when CL *in vivo* is unexpectedly high because this method naturally overpredicts the true metabolic CL, as is reported in the literature.[1,2,7,13] Finally, we present a sensitivity analysis to demonstrate how the different IVIVE methods vary with input parameters related to drug ionization, plasma protein binding, and/or CL_{int}.

Comparative Analysis of IVIVE Methods

Five IVIVE methods that have undergone previous comparative assessments were the focus of further evaluation in this study.[1,2,14,15] These IVIVE methods are (a) the "conventional" and "conventional bias-corrected" methods using the fu_p, (b) the "Berezhkovskiy method" in which the fu_p is adjusted for drug ionization on either side of the plasma membrane on the basis on pH differences, (c) the "Poulin et al.[14] method" using the fu_{liver} to adjust in addition for protein-facilitated uptake because of the potential ionic interactions between the plasma-protein-bound-drug complex and the cell surface of the hepatocytes, and (d) the "direct scaling method" that does not consider any binding corrections. Table 1 summarizes all equations related to these IVIVE methods. Recently, Halifax and Houston[15] reported an empirical method, the "conventional bias-corrected method", which involves multiplying the predicted CL values from the conventional method with the corresponding average bias of underprediction to reduce the underprediction. The average bias of underprediction was obtained from the AFE observed for each dataset (i.e., each prediction scenario) of this study. This empirical method was also evaluated in this study. The well-stirred model was considered for the purpose of this study. Furthermore, the parallel tube model was also used for high CL compounds for all IVIVE methods tested because it is expected that the prediction accuracy for these drugs will increase with the parallel tube model.[2]

Estimation of the Input Parameters

The five IVIVE methods scale CL_{int} determined in microsomes from *in vitro*-to-*in vivo* conditions by using a physiologically based scaling factor based on