Comparison of Human Long-Term Liver Models for Clearance Prediction of Slowly Metabolized Compounds

Lena C. Preiss¹,²,³, Katrin Georgi²,⁴, Volker M. Lauschke¹,⁵,⁶, Carl Petersson², #

¹ Department of Physiology and Pharmacology, Karolinska Institutet, Stockholm, Sweden

² Department of Drug Metabolism and Pharmacokinetics (DMPK), The Healthcare Business of Merck KGaA, Darmstadt, Germany

³ Present address, Pharmaceutical Research and Early Development, F. Hoffman La Roche AG, Basel, Switzerland

⁴ Present address, Department of Drug Metabolism and Pharmacokinetics (DMPK), Bayer Healthcare, Wuppertal, Germany

⁵ Dr Margarete Fischer-Bosch Institute of Clinical Pharmacology, Stuttgart, Germany

⁶ University of Tuebingen, Tuebingen, Germany

# Corresponding author: Carl Petersson, PhD

e-mail: carl.petersson@merckgroup.com

NCE DMPK, Discovery & Development Technologies, Merck Healthcare KGaA, Frankfurter Str. 250, Postcode D50/902, 64293 Darmstadt, Germany
Running title: Comparison of Human Long-Term Liver Models for Clearance

Key words: 3D cell culture; cytochrome P450; IVIVE; low clearance; pharmacokinetics; primary human hepatocytes

Number of:
Text Pages: 23
Tables: 3
Figures: 4
References: 41
Words in Abstract: 231
Words in Introduction: 537
Words in Discussion: 1276

Abbreviations:
ADME, absorption, distribution, metabolism, excretion; CL_{int}, intrinsic clearance; CL_{int,obs}, observed intrinsic clearance; CL_{int,pred}, predicted intrinsic clearance; CL_{b}, in vivo clearance; CL_{b,obs}, observed in vivo clearance; CL_{b,pred}, predicted in vivo clearance; CYP, cytochrome P450; FBS, fetal bovine serum; fu_{b}, fraction unbound in blood; fu_{inc}, fraction unbound in the incubation; fu_{p}, fraction unbound in plasma; IVIVE, in vitro-in vivo extrapolation; LOOCV, leave one out cross-validation; LOQ, limit of quantification; MPCCs, Micropattern co-cultures; PHH, primary human hepatocytes; pl, plasma; Q_{h},
liver blood flow; R_b, blood-to-plasma ration; SCC, Stochastic co-culture; WSM, well-stirred model.

**ABSTRACT**

The accurate prediction of human clearance is an important task during drug development. The proportion of low clearance compounds has increased in drug development pipelines across the industry since such compounds may be dosed in lower amounts and at lower frequency. Such compounds present new challenges to in vitro systems used for clearance extrapolation. In this study we compared the accuracy of clearance predictions of suspension culture to four different long-term stable in vitro liver models, including HepaRG sandwich culture, the Hµrel stochastic co-culture, the Hepatopac micropatterned co-culture (MPCC) and a micro-array spheroid culture. Hepatocytes in long-term stable systems remained viable and active over several days of incubation. Although intrinsic clearance values were generally high in suspension culture, clearance of low turnover compounds could frequently not be determined using this method. Metabolic activity and intrinsic clearance values from HepaRG cultures were low and, consequently, many compounds with low turnover did not show significant decline despite long incubation times. Similarly, stochastic co-cultures occasionally failed to show significant turnover for multiple low and medium turnover compounds. Among the different methods, MPCCs and spheroids provided the most consistent measurements. Notably, all culture methods resulted in underprediction of clearance, this could however be compensated for by regression correction. Combined, the results indicate that spheroid culture as well as the MPCC system provide adequate in vitro tools for human extrapolation for compounds with low metabolic turnover.
SIGNIFICANT STATEMENT

In this study, we compared suspension cultures, HepaRG sandwich cultures, the Hürel liver stochastic co-cultures, the Hepatopac micropatterned co-cultures (MPCC) and micro-array spheroid cultures for low clearance determination and prediction. Overall, HepaRG and suspension cultures showed modest value for the low determination and prediction of clearance compounds. The micro-array spheroid culture resulted in the most robust clearance measurements, whereas using the MPCC resulted in the most accurate prediction for low clearance compounds.

INTRODUCTION

The prediction of human pharmacokinetics and clinical doses is an important task in drug discovery with metabolic clearance evaluation being one of the important parameters. The most commonly applied method to measure the in vitro intrinsic clearance (CL\text{int}) is substrate turnover in liver microsomes or primary hepatocytes in suspension cultures (Obach, 1999; Yadav et al., 2021). However, the activity of drug metabolizing enzymes in these culture systems decrease rapidly (Di et al., 2012; Di & Obach, 2015; Smith et al., 2012; Stringer et al., 2008, Zanelli et al 2017). This limits their utility for slowly metabolized compounds, which are increasingly prevalent in drug development programs.
To overcome this limitation, advanced liver cell cultures systems in which cells remain viable and stable for longer time have been suggested. Hepatic cell lines and induced pluripotent stem cell (iPSC)-derived hepatocyte-like cells are readily available. However, the activity of drug metabolizing enzymes in these cells has been shown to be significantly lower compared to primary human hepatocytes (Kratochwil et al., 2017). Among the cell lines, HepaRG cells exhibit the highest metabolic activity, but still differ considerably in their overall transcriptomic signature and functionality from primary human hepatocytes (PHH) (Bell et al., 2017; Berger et al., 2016; Hart et al., 2010). Recently, different liver cell cultures models have been developed that maintain physiologically relevant metabolic activity in hepatocytes for extended time periods. These models have also been applied for extrapolation of clearance (Lauschke et al., 2016, 2019; Sodhi & Benet, 2021; Yadav et al., 2021). A co-culture of primary hepatocytes with stromal cells was shown to stabilize hepatic phenotypes for multiple weeks in culture and two different variations, a stochastic co-culture and a micropatterned co-culture, were introduced for low clearance determinations (Khetani & Bhatia, 2008; Novik et al., 2017; Wang et al., 2010).

Next to these advanced 2D liver cell culture models, 3D culture models have been developed for pharmacokinetic assessments (Lauschke et al., 2016; Underhill & Khetani, 2018), of which spheroid cultures of primary human hepatocytes are most widely used (Cox et al., 2022a; Kanebratt et al., 2021; Preiss et al., 2022; Riede et al., 2021). Spheroids are typically cultured in ultra-low attachment plates in 96- or 384-well format with one spheroid per well. However, due to the low number of cells per well (mostly...
1,500 to 5,000), these cultures are not suitable to determine intrinsic clearance of slowly metabolized compounds (Preiss et al., 2022). Micro-array cultures in which many spatially separated spheroids are cultured in each well present a promising option to mitigate these shortcomings (Cox et al., 2022a; Preiss et al., 2022).

While there is increasing diversity in the landscape of model systems that are suitable for low clearance determination, a direct side-by-side comparison has been lacking to date. In this study, we compared suspension culture, HepaRG sandwich culture, the Hµrel liver stochastic co-culture (SCC), the Hepatopac micropatterned co-culture (MPCC) and micro-array spheroid culture (henceforth referred to as “multi-spheroids”) using a panel of 21 compounds with hepatic metabolism as the major clearance pathway. Where possible, the same hepatocyte donor was used to mitigate inter-donor variability and to facilitate the direct benchmarking of culture systems. Our results provide the first head-to-head comparison of these systems suitability for low clearance determination.

**MATERIAL AND METHODS**

**Materials**

William’s E medium, L-glutamine-penicillin-streptomycin solution, midazolam, 1’OH-midazolam, 4’OH-diclofenac, dextrophan, dextromethorphan, amodiaquine, N-desethyl-amodiaquine, dexamethasone, 7’OH-coumarin, 7’OH-coumarin glucuronide, trypan blue
and acetonitrile were purchased from Sigma-Aldrich (St. Louis, MO, USA). Elplasia 96-well round bottom ultra-low attachment (ULA) plates, and 96-well white round bottom polystyrene microplates were purchased from Corning (Glendale, AZ, USA). Fetal bovine serum was obtained from Cytiva (Marlborough, MA, USA). Insulin-transferrin-selenium solution, Matrigel, HepaRG cells, GlutaMAX, HepaRG Maintenance/Metabolism Medium Supplement, HepaRG Thaw, Plate, & General Purpose Medium Supplement and rapid equilibrium dialysis plates were obtained from Thermo Fisher Scientific (Waltham, MA, USA). Cryopreserved primary human hepatocytes (Lot HJK, CRT, 30-year-old Female), INVITROGRO Krebs-Henseleit Buffer, INVITROGRO HT medium, INVITROGRO CP medium and Hepatopac Kits (Donor CRT) were purchased from BioreclamationIVT (NY, USA). Hµrel micro liver kits (Lot 1065HUP, 5-donor pool, 3 males, 2 females, age 6-65) were obtained from Visikol, Inc. (Hampton, NJ, USA). All test compounds were provided by the compound management department of the Healthcare Business of Merck KGaA (Darmstadt, Germany).

Methods

**Cell Culture and incubation of test substances**

All incubations were carried out at 37°C and 5% CO₂. Test substance concentration was 1µM, unless otherwise stated. Test compounds were formulated as a solution containing DMSO, the final concentration was less than 0.2% in all test systems. Technical duplicates were performed for each biological replicate experiment.
For suspension culture, PHH were thawed, cells were incubated directly in 100µL of Krebs-Henseleit buffer (pH 7.4) at a density of 0.5 million cells per mL as previously described (Zanelli et al., 2018). Plates were gently agitated during the incubation. Aliquots were collected at 8 time points. The time points were distributed depending on the expected depletion rate up to 110 min. The experiment was performed in duplicate (n=2).

For HepaRG cultures, the cells were quickly thawed in William’s E supplemented with Thaw, Plate & General Purpose Medium Supplement (thawing medium). After centrifugation at 500xg for 3min, the supernatant was removed, and the cell pellet was resuspended in 5mL thawing medium. HepaRG cells were seeded into collagen I-coated 96-well plates at 72,000 cells per well. After seeding, the cells were kept at 37°C and 5% CO₂ overnight. On the next day, thawing medium was removed and 100µL of Matrigel working solution (0.25mg/mL in ice cold William’s E supplemented with HepaRG Maintenance/Metabolism Medium Supplement (HepaRG maintenance medium)) was added to each well. HepaRG maintenance medium was renewed every two to three days. On day eight after seeding, cells were incubated with test substances in HepaRG maintenance medium. The final concentration of HepaRG cells was 0.7 million cells per mL. Approx. half were biliary like cells (not hepatocyte like), hence the concentration of 0.3 million cells per mL was used for any further calculations (Cerec et al 2007). Aliquots were collected at 8 time points. The time points were distributed depending on the expected depletion rate up to 72 hrs. The experiment was performed in quadruplicate (n=4).
SCC were obtained from Hµrel micro liver kits shipped from Visikol, Inc. Cells were shipped plated (96-well format) in a shipping medium at 37°C. After arrival, shipping medium was replaced with maintenance medium and cells were incubated at 37°C and 5% CO₂. For the incubation maintenance medium was replaced with test substances dissolved in dosing medium as described by the manufacturer. The final concentration of hepatocytes were 0.4 million cells per mL. Aliquots were collected at 11 time points. The time points were distributed depending on the expected depletion rate up to 72 hrs. The experiment was performed in quadruplicate (n=4).

Hepatopac kits were obtained from BioIVT. Cells were shipped plated and in shipping medium at 37°C. After arrival, shipping medium was replaced with Hepatopac maintenance medium and plates were recovered at 37°C and 10% CO₂ for four days. Maintenance medium was renewed every two days. After recovery, maintenance medium was replaced with Metabolic Stability Application Medium, and cell were pre-incubated at 37°C and 5% CO₂ for two hours. To start the incubation, 50% of Metabolic Stability Application Medium was renewed with 2X test substances dissolved in Metabolic Stability Application Medium as described by the manufacturer. The final concentration of hepatocytes were 0.05 million hepatocytes per mL. Aliquots were collected at 8 time points. The time points were distributed depending on the expected depletion rate up to 72 hrs. The experiment was performed in duplicate (n=2).

Multi-spheroids were cultured as previously described (Preiss et al., 2022). In brief, 40,000 PHH were seeded in PHH medium (William’s E medium supplemented with 2mM
L-glutamine, 100 units/mL penicillin, 100 μg/mL streptomycin, 10 μg/mL insulin, 5.5 μg/mL transferrin, 6.7 ng/mL sodium selenite, and 100 nM dexamethasone) supplemented with 10% fetal bovine serum (FBS) in plates with ~80 microwells per macrowell. After 5 days, spheroids were formed, and 75% of medium was renewed with serum-free PHH medium. On day 6, incubations were started by replacing 50% of medium with 2X test substances dissolved in serum-free PHH medium resulting in an incubation medium with 1.7% serum. Incubations were carried out in 100 μL. The final concentration of hepatocytes were 0.4 million hepatocytes per mL. Aliquots were collected at 8 time points. The time points were distributed depending on the expected depletion rate up to 72 hrs. The experiment was performed in triplicate (n=3).

**LC-MS Analysis**

Supernatant samples were taken at different time points and stopped with double the volume of ice-cold acetonitrile containing the internal standard pruvanserin. Samples were kept at -20°C until further use. After centrifugation for 30min at 3,500xg and 4°C, supernatant was diluted 1:10 with 66% acetonitrile. An AB Sciex API 6500+ triple quadrupole (AB Sciex LLC, MA, USA) coupled with a Waters Acquity I-Class UPLC (Waters Corporation, MA, USA) were used for LC-MS/MS analysis. The system was controlled using the software Analyst v1.7 (AB Sciex LLC). LC was performed using an Acquity UPLC BEH C18 1.7 μm, 2.1mm x 50mm column (AB Sciex LLC), 70mM ammonium formate buffer containing 0.1% (v/v) formic acid as mobile phase A and acetonitrile as mobile phase B. Aliquots (4 μL) of the diluted samples were injected. The gradient conditions were the following: 0–0.1 min 0% B, 0.1–0.8 min 0–100% B, 0.8–1.0
min 100% B, 1.0–1.4 min 0% B. The flow rate was 0.8 mL/min and data collection occurred between 0.3- and 1.4-min. Samples were measured using electrospray ionization in positive or negative mode with multiple reaction monitoring (Appendix 1).

**Determination of fraction unbound in plasma**

The unbound fraction in plasma ($f_{u_p}$) was measured using a rapid equilibrium device (RED, Thermo Fisher Scientific) with a compound concentration of 1µM. Human plasma was dialyzed against phosphate buffer at pH 7.4 for 4h at 37°C. These conditions are known to yield equilibrium in the device used (van Liempd et al., 2011). Samples were measured using LC-MS/MS. By dividing the drug concentration in the dialysate by the concentration in the plasma, the $f_{u_p}$ was determined. By dividing the obtained $f_{u_p}$ by the blood-to-plasma ratio (Rb), the fraction unbound in blood ($f_{u_b}$) was calculated. Rb was obtained from the literature (Table 1).

**Data analysis**

*In Vitro* $CL_{int}$ values were calculated from parent compound loss using eq.1:

\[
In\, Vitro\, CL_{int} = -k_e \frac{incubation\, volume}{seeded\, hepatocytes} \mu L * min^{-1} * 10^6 cells^{-1} \quad (1)
\]

Where $k_e$ is the elimination rate constant representing the slope of the linear regression of the natural logarithm of the percentage remaining parent compound versus incubation time. Significant drug turnover was defined as a slope significantly different from zero. Deviation of the slope from zero was tested using a F-test followed by a Bonferroni-
Dunn correction. No further calculations were performed in absence of significant turnover. The limit of quantification (LOQ) for the In Vitro CL$_{\text{int}}$ was used for further calculations when one or more experiments showed significant turnover. The LOQ was defined as parent compound loss of more than 15% over the cause of the experiment. CL$_{\text{int}}$ values corresponding to LOQ were 0.1, 0.1, 3, 0.8, 0.1 µL/min/10$^6$ hepatocytes for HepaRG, stochastic co-culture, suspension, MPCC and spheroid cultures, respectively.

The scaled In Vivo CL$_{\text{int}}$ was calculated from in vitro CL$_{\text{int}}$ by taking physiological scaling factors (25.7 g liver/kg body weight, 120×10$^6$ cells/g liver), fraction unbound in the incubation (fu$_{\text{inc}}$) and blood (fu$_{b}$) into account (eq.2).

$$\text{Scaled In Vivo CL}_{\text{int}} = \text{In Vitro CL}_{\text{int}} \cdot \frac{\text{liver weight}}{\text{standard body weight}} \cdot \frac{\text{hepatocytes}}{\text{gram of liver}} \cdot \frac{f_{u_{b}}}{f_{u_{\text{inc}}}}$$  \hspace{1cm} (2)

Using logD$_{7.4}$ for acidic and neutral compounds or logP for basic compounds, fu$_{\text{inc}}$ was calculated as previously described (Kilford et al, 2008). Additionally, a fu$_{\text{inc}}$ due to plasma protein binding based the serum in the assay was compensated for (applicable for multi well spheroids only) as described previously (Yamagata et al, 2017).

For in vivo CL$_{\text{int}}$ predictions (In Vivo CL$_{\text{int,pred}}$), a regression line correction was used as described previously (Sohlenius-Sternbeck et al, 2012). Using the well-stirred model (WSM) and the corrected predicted In Vivo CL$_{\text{int}}$, the predicted in vivo clearance (CL predicted) was determined (eq.3) with $Q_h$ being the liver blood flow (20 mL/min/kg).

$$\text{CL predicted} = \frac{\text{In Vivo CL}_{\text{int,pred}} * Q_h}{Q_h + \text{In Vivo CL}_{\text{int,pred}}}$$  \hspace{1cm} (3)

The observed In Vivo hepatic clearance (CL observed) was obtained from the In Vivo Plasma Clearance (In Vivo CL$_{\text{pl,obs}}$) as described in eq 4.
\[ \text{CL observed} = \frac{\text{In Vivo } \text{CL}_{\text{pl,obs}}}{R_b} \quad (4) \]

The observed In Vivo CL\text{int} (\text{In Vivo } \text{CL}_{\text{int,obs}}) was derived from the CL observed using the WSM (eq. 5).

\[ \text{In Vivo } \text{CL}_{\text{int,obs}} = \frac{\text{CL observed} \times Q_h}{Q_h - \text{CL observed}} \quad (5) \]

RESULTS

**In Vitro Clearance Assessment: Comparison between Long-Term In Vitro Liver Model and Suspension Cultures**

In this study, five different hepatic cell systems were compared regarding their ability to accurately determine hepatic clearance (Figure 1). Three of the studied systems were evaluated using the same donor; suspension cultures, multi-spheroids and MPCCs. In total, the \textit{in vitro} clearance of 21 compounds, spanning a wide range from low to high clearance was determined. The set of compounds was selected to reflect marketed compounds in terms of metabolic pathways and physicochemical properties (Table 1). Eight compounds of the set were defined as low CL compounds (Table 2).

Four compounds could not be measured in the HepaRG culture due to instrument failure during the analysis or compound availability. From the 17 measured compounds,
significant turnover in all repeats could be measured for eight compounds. Naloxone and theophylline, a low clearance compound, showed no significant turnover in HepaRG cells in any of the repeats. The remaining seven compounds showed a measurable intrinsic clearance in some but not all replicates. A similar success rate in turnover measurement was achieved using suspension-cultured hepatocytes. 12 of 21 (57%) compounds resulted in significant turnover in all repeats. Especially compounds with low hepatic turnover were not consistently quantified, three low clearance compounds showed no significant turnover in any experiment. In contrast, significant turnover was obtained in both co-culture systems and the multi-spheroids for all compounds in at least one experiment (Figure 2A). Multi-well spheroid showed significant turnover in all repeats for 20 of 21 compounds (95%) while the corresponding value was about 80% for both SCC and MPCC (Table 2).

Overall, the lowest CLint values were obtained when using the HepaRG culture (Figure 3, Table 2). Interestingly, differences in CLint values were also observed in donor-matched systems. The CLint values as well as the regression offset, which is a function thereof, indicated that MPCC and suspension culture had similar activity while the metabolic activity was lower in multi-spheroids (Figure 3, Table 2). The multi-spheroid model showed the narrowest confidence interval for individual measurements, whereas the HepaRG culture showed the widest. The reproducibility was highest for suspension cultures, followed by multi-spheroids and MPCCs for medium-to-high clearance compounds (Figure 2D). The variation was lowest for multi-spheroids compared to the co-culture systems when considering low clearance compounds (Figure 2E).
There was a noticeable underprediction of human in vivo CLint across all model systems (Figure 3). This was most pronounced for high clearance compounds. Overall, the underprediction was largest for HepaRG cells with a Root Mean Squared Error (RMSE) of 1.65 while it was lowest for the suspension model and MPCC with values close to 1 when compared to the line of unity (Figure 3).

**Total Clearance Prediction after Regression Correction**

As the uncorrected intrinsic clearance prediction showed underprediction for all tested systems, a regression correction was performed for total clearance prediction. The predicted clearance was within 2-fold for 76, 95, 67, 81 and 86% of the compounds in HepaRG, stochastic co-culture, suspension, MPCC or multi-spheroid system, respectively. The low clearance compound subset (N=7-8) resulted in 57, 86, 25, 50 and 63% within 2-fold when using the HepaRG, stochastic co-culture, suspension, MPCC or multi-spheroid system, respectively. The subset with higher CLint values (N=10-13) resulted in more than 90% of the compounds within two-fold for all the tested systems (Figure 4, Table 3).

**DISCUSSION**

Drugs with a low clearance are associated with lower dose levels and frequently show a long half-life, allowing for prolonged target interaction. Such drugs therefore are
frequently compatible with longer dosing intervals. These attractive properties have made them increasingly common in drug discovery pipelines. However, their human dose prediction is however challenging, as standard in vitro models like human liver microsomes and hepatocytes in suspension culture show rapid loss of metabolizing enzyme activity, thereby limiting the incubation time. In fact, significant turnover of such slowly metabolized compounds is frequently not detectable using standard systems. Therefore, new human hepatic models using different cell models or different culture systems have been developed in which liver cells are stable for longer periods of time.

Determination of intrinsic clearance using HepaRG cells was challenging. The reproducibility of the data was low and significant turnover was often not measurable. In addition, CLint values were generally low with pronounced underprediction compared to other systems and especially low clearance compounds were not predicted well. This is in line with previous studies, Bonn et al.

This result was expected as the HepaRG cells were derived from an individual that is a poor CYP2D6 metabolizer (Guillouzo et al. 2007).

The suspension culture showed good metabolic activity and high CLint values compared to other systems. Determination of intrinsic clearance for the medium to high clearance compounds was straight forward and resulted in robust data. However, significant
turnover was frequently not observed for the low clearance compounds. This underlines the limitation of suspension cultures for low clearance determination as reported previously (Di & Obach, 2015; Hutzler et al., 2015; Smith et al., 2012).

The intercept of the regression lines was more than half a log unit lower for the stochastic co-culture compared to MPCCs suggesting that the CLint values were systematically lower (Figure 3). This is in line with previous reports (Kratochwil et al., 2017). However, a direct comparison of the culture systems per se remains difficult as both in the study by Kratochwil et al. and our work, used different donors or donor-pools in the two systems. In general, both systems can be used for accurate and robust clearance prediction of low but also medium to high clearance compounds in agreement with previous studies (Docci et al., 2019a; Hultman et al., 2016; Umehara et al., 2020). Both systems are commercially available as ready-to-use, which decreases hands-on-time and makes them less technically challenging to work with. However, the shipments are time critical and need to be planned well in advance. This makes these systems less flexible compared to models which can be run completely in house such as suspension and spheroid cultures.

The multi-spheroid culture provides a higher flexibility but requires considerable manual work. In addition, so far only single donors can be used for spheroid formation. When choosing a donor with good activity of drug metabolizing enzymes, accurate predictions can be obtained as shown in this study. The underprediction was similar to the 5-donor pool of the stochastic co-culture. However, it was more pronounced compared to the
MPCC culture using the same donor. A potential reason for this observation is that the number of hepatocytes in the incubation was overestimated for the spheroid culture as not all seeded cells were incorporated in the spheroids. Therefore, a precise measurement of cells after spheroid formation would be desired. Nevertheless, in direct comparisons CLint determination of low clearance compounds was most successful using the multi-spheroid culture. This may be attributed to the fact that the multi-spheroid culture does not only allow for a long incubation time but also has the highest cell concentration of the advanced cell culture systems based on hepatocytes. In addition, compounds with metabolic turnover could be measured with high precision in line with previous findings (Preiss et al., 2022).

The inter experimental variability was lowest in the suspension culture for high clearance compounds. The current study adopted the time points based on the expected CLint values. Hence, the higher variability in the other systems was not a product of an insufficient number of data points defining the slope. Other factors explaining the observed differences needs to be sought. Long-term cultures have undergone culturing prior to the experiment which may cause variability in enzyme expression/activity. An alternate explanation could be that longer incubation times may affect the volume of the incubation through evaporation. The degree of evaporation may also vary with the position on the plate. Such effect may be addressed by a metabolically inert internal standard in the incubation as suggested by Zanelli et al (Zanelli et al 2017). The variability was lowest for the multi-spheroids for the low clearance group. This system showed the highest turnover of all systems. Higher turnover will increase the precision.
and reproducibility in the determination slope used to derive the CLint values in cases where the slope is shallow (approaching the LOQ). The high turnover of the multi-spheroids is most likely related to the relatively high cell concentration (0.4 million cells per mL) combined with a good activity per cell. Overall, this indicates that consistent turnover is a critical factor to obtain reproducible results for low clearance compounds.

Incorporating a regression correction factor for total clearance prediction, yielded good accuracy for all the tested systems. This suggests that empirical correction may compensate for suboptimal donor selection and possibly also for shortcomings in the estimation of cells in the cultures used. The validity of the regression offset approach has been confirmed by external validation sets or by leave-one-out cross validation in multiple publications (Preiss et al., 2022, Sohlenius-Sternbeck el al., 2012, Yamagata et al., 2017). Regression correction is however only suitable for correcting systematic underprediction. Further, the approach is dependent on the composition of the calibration set, containing compounds eliminated by the most important drug metabolizing enzymes and with diverse physicochemical properties. The compounds used in the present study fulfill these requirements (Table 1). Compounds eliminated via some less common drug metabolizing enzymes such as N-acetyltransferase-1 were however not present in the calibration set used in this study. Thus, we cannot make statements about the accuracy of the established regression correction for such compounds and further data is required to accurately extrapolate their human clearance.
In conclusion, the prediction of clearance of low turnover compounds strongly benefits from the use of advanced \textit{in vitro} systems as hepatocyte suspension. HepaRG cells were found to perform relatively poorly due to overall lower activity of drug metabolizing enzymes and very low activity of CYP2D6. 2D hepatocyte co-cultures formats and multi-spheroid cultures gave robust measurements for low clearance compounds and good prediction accuracies. Spheroid cultures performed best in low turnover measurements, whereas the underprediction was lowest using the MPCC and suspension.

This study showed that suspension culture of human hepatocytes remains a robust and simple tool to predict clearance of medium to high clearance compounds. This culture system does however suffer from shortcomings dealing with compounds with low metabolic turnover. More sophisticated cultures like co-cultures or spheroid cultures present a valuable alternative for further clearance determination and lead candidate selection. In addition, literature suggest that organ-on-a-chip culture systems may be of value for low clearance measurements and determination of fraction metabolized. However, to date, these tools are limited in their throughput and a correction for medium evaporation is often required (Cox et al., 2022; Docci et al., 2019, Rajan et al.,2023).

\textbf{Acknowledgments}

This study received financial support from the healthcare business of Merck KGaA (Darmstadt, Germany).
Data Availability Statement Samples

The authors declare that all the data supporting the findings of this study are available within the paper and its Supplemental Data.

Conflict of Interest Statement

Lena C. Preiss, Katrin Georgi and Carl Petersson were employed by the healthcare business of Merck KGaA (Darmstadt, Germany) when this study was conducted. Volker M. Lauschke is co-founder, CEO, and shareholder of HepaPredict AB, and as well as co-founder and shareholder of Shanghai Biotechnology Ltd.

Authorship Contributions

Participated in research design: Preiss, Petersson, Lauschke, Georgi

Conducted experiments: Preiss

Performed data analysis: Preiss, Petersson, Lauschke

Wrote or contributed to the writing of the manuscript: Preiss, Petersson, Lauschke, Georgi
REFERENCES


### Table 1. Overview of tested compounds.

<table>
<thead>
<tr>
<th>Compound</th>
<th>Ion class</th>
<th>Main metabolizing enzymes</th>
<th>$R_b$</th>
<th>$fu_p$</th>
<th>$fu_{inc}$</th>
<th>$f_{HepaRG}$</th>
<th>$f_{SCC}$</th>
<th>$f_{Suspension}$</th>
<th>$f_{MPCC}$</th>
<th>$f_{Sph}$</th>
<th>CL,pl [mL/min/kg]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Warfarin</td>
<td>A</td>
<td>2C9, 3A4</td>
<td>0.55$^a$</td>
<td>0.01</td>
<td>0.98</td>
<td>0.98</td>
<td>0.97</td>
<td>1.00</td>
<td>0.46</td>
<td>0.1$^f$</td>
<td></td>
</tr>
<tr>
<td>Tolbutamide</td>
<td>A</td>
<td>2C9, 2C19</td>
<td>0.68$^a$</td>
<td>0.02</td>
<td>0.98</td>
<td>0.98</td>
<td>0.97</td>
<td>1.00</td>
<td>0.56</td>
<td>0.2$^f$</td>
<td></td>
</tr>
<tr>
<td>Diazepam</td>
<td>N</td>
<td>2C19, 3A4</td>
<td>0.78$^a$</td>
<td>0.01</td>
<td>0.93</td>
<td>0.91</td>
<td>0.88</td>
<td>0.99</td>
<td>0.38</td>
<td>0.4$^d$</td>
<td></td>
</tr>
<tr>
<td>Theophylline</td>
<td>N</td>
<td>1A2</td>
<td>0.83$^a$</td>
<td>0.21</td>
<td>0.99</td>
<td>0.98</td>
<td>0.98</td>
<td>1.00</td>
<td>0.94</td>
<td>0.5$^d$</td>
<td></td>
</tr>
<tr>
<td>Glipizide</td>
<td>A</td>
<td>2C9</td>
<td>0.55$^a$</td>
<td>0.01</td>
<td>0.98</td>
<td>0.98</td>
<td>0.97</td>
<td>1.00</td>
<td>0.40</td>
<td>0.6$^d$</td>
<td></td>
</tr>
<tr>
<td>Etodolac</td>
<td>A</td>
<td>2C9, UGT</td>
<td>0.6$^b$</td>
<td>0.01</td>
<td>0.98</td>
<td>0.97</td>
<td>0.97</td>
<td>1.00</td>
<td>0.16</td>
<td>1.3$^d$</td>
<td></td>
</tr>
<tr>
<td>Ketoprofen</td>
<td>A</td>
<td>UGT</td>
<td>0.55$^a$</td>
<td>0.02</td>
<td>0.98</td>
<td>0.98</td>
<td>0.97</td>
<td>1.00</td>
<td>0.50</td>
<td>1.6$^g$</td>
<td></td>
</tr>
<tr>
<td>Gemfibrozil</td>
<td>A</td>
<td>2C8</td>
<td>0.55$^a$</td>
<td>0.01</td>
<td>0.97</td>
<td>0.97</td>
<td>0.96</td>
<td>1.00</td>
<td>0.14</td>
<td>1.7$^d$</td>
<td></td>
</tr>
<tr>
<td>Diclofenac</td>
<td>A</td>
<td>2C9, UGT2B7</td>
<td>0.55$^a$</td>
<td>0.01</td>
<td>0.73</td>
<td>0.68</td>
<td>0.61</td>
<td>0.94</td>
<td>0.32</td>
<td>3.5$^d$</td>
<td></td>
</tr>
<tr>
<td>Ondansetron</td>
<td>B</td>
<td>3A4, 2D6, 1A2</td>
<td>0.83$^a$</td>
<td>0.50</td>
<td>0.95</td>
<td>0.94</td>
<td>0.92</td>
<td>0.99</td>
<td>0.94</td>
<td>5.8$^d$</td>
<td></td>
</tr>
<tr>
<td>Domperidone</td>
<td>B</td>
<td>3A4</td>
<td>1.00$^a$</td>
<td>0.06</td>
<td>0.80</td>
<td>0.77</td>
<td>0.71</td>
<td>0.96</td>
<td>0.75</td>
<td>9.5$^d$</td>
<td></td>
</tr>
<tr>
<td>Desipramine</td>
<td>B</td>
<td>2D6</td>
<td>0.96$^a$</td>
<td>0.13</td>
<td>0.63</td>
<td>0.58</td>
<td>0.51</td>
<td>0.91</td>
<td>0.56</td>
<td>11.0$^a$</td>
<td></td>
</tr>
<tr>
<td>Imipramine</td>
<td>B</td>
<td>2C9,2D6,3A4,1A2</td>
<td>1.07$^a$</td>
<td>0.12</td>
<td>0.55</td>
<td>0.49</td>
<td>0.42</td>
<td>0.88</td>
<td>0.48</td>
<td>13$^a$</td>
<td></td>
</tr>
<tr>
<td>Metoprolol</td>
<td>B</td>
<td>2D6,3A4</td>
<td>1.07$^a$</td>
<td>0.94</td>
<td>0.96</td>
<td>0.95</td>
<td>0.94</td>
<td>0.99</td>
<td>0.95</td>
<td>13$^a$</td>
<td></td>
</tr>
<tr>
<td>Sildenafil</td>
<td>N</td>
<td>3A4</td>
<td>0.62$^c$</td>
<td>0.07</td>
<td>0.93</td>
<td>0.92</td>
<td>0.89</td>
<td>0.99</td>
<td>0.81</td>
<td>9.1$^d$</td>
<td></td>
</tr>
<tr>
<td>Nifedipine</td>
<td>N</td>
<td>3A4</td>
<td>0.59$^a$</td>
<td>0.03</td>
<td>0.93</td>
<td>0.91</td>
<td>0.88</td>
<td>0.99</td>
<td>0.66</td>
<td>7.3$^d$</td>
<td></td>
</tr>
<tr>
<td>Propafenone</td>
<td>B</td>
<td>2D6</td>
<td>1.00$^a$</td>
<td>0.07</td>
<td>0.81</td>
<td>0.77</td>
<td>0.71</td>
<td>0.96</td>
<td>0.76</td>
<td>16.0$^a$</td>
<td></td>
</tr>
<tr>
<td>Chlorpromazine</td>
<td>B</td>
<td>2D6</td>
<td>0.78$^a$</td>
<td>0.01</td>
<td>0.23</td>
<td>0.20</td>
<td>0.15</td>
<td>0.65</td>
<td>0.19</td>
<td>16.0$^a$</td>
<td></td>
</tr>
</tbody>
</table>

$R_b$ = blood-to-plasma ratio; $fu_p$ = fraction unbound in plasma; $fu_{inc}$ = calculated fraction unbound in the incubation; CL,pl = observed human plasma clearance;
<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>UGT</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>f</th>
<th>g</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naloxone</td>
<td>B</td>
<td>UGT</td>
<td>1.22&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.25</td>
<td>0.96</td>
<td>0.95</td>
<td>0.94</td>
<td>0.99</td>
<td>0.95</td>
</tr>
<tr>
<td>Verapamil</td>
<td>B</td>
<td>2C8,3A4,1A2,2C9</td>
<td>1.00&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.13</td>
<td>0.65</td>
<td>0.60</td>
<td>0.53</td>
<td>0.92</td>
<td>0.59</td>
</tr>
<tr>
<td>Propranolol</td>
<td>B</td>
<td>2D6,1A2</td>
<td>0.86&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.20</td>
<td>0.86</td>
<td>0.83</td>
<td>0.79</td>
<td>0.97</td>
<td>0.82</td>
</tr>
</tbody>
</table>


<sup>a</sup> Values obtained from (Brown et al., 2007)
<sup>b</sup> Values obtained from (Paixão et al., 2009)
<sup>c</sup> Values obtained from (Rodgers & Rowland, 2007)
<sup>d</sup> Values obtained from (Sohlenius-Sternbeck et al., 2012)
<sup>e</sup> Estimated $R_b$ values
<sup>f</sup> Values obtained from (Hallifax et al., 2010)
<sup>g</sup> Values obtained from (Obach et al., 2008)
Table 2. *In vitro* clearance (CLint) for the different hepatic cell systems. Data is shown as mean ± S.D., n = 4 for HepaRG and stochastic co-culture, n = 3 for multi-spheroids (Sph) n = 2 for suspension and micropatterned co-cultures (MPCC). ND; not determined, NST; no significant turnover

<table>
<thead>
<tr>
<th>Compound</th>
<th><em>In vitro</em> CLint [µL/min/10^6 hepatocytes]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HepaRG</td>
</tr>
<tr>
<td>Chlorpromazine</td>
<td>21.7 ± 14</td>
</tr>
<tr>
<td>Desipramine</td>
<td>ND</td>
</tr>
<tr>
<td>Diazepam†</td>
<td>0.3 ± 0.2§</td>
</tr>
<tr>
<td>Diclofenac</td>
<td>15.9 ± 6.7</td>
</tr>
<tr>
<td>Domperidone</td>
<td>4.7 ± 3.8§</td>
</tr>
<tr>
<td>Etodolac†</td>
<td>6.2 ± 4.5</td>
</tr>
<tr>
<td>Gemfibrozil</td>
<td>ND</td>
</tr>
<tr>
<td>Glipizide†</td>
<td>0.3 ± 0.4§</td>
</tr>
<tr>
<td>Imipramine</td>
<td>1.4 ± 0.6</td>
</tr>
<tr>
<td>Ketoprofen†</td>
<td>1.6 ± 1.6</td>
</tr>
<tr>
<td>Metoprolol</td>
<td>0.3 ± 0.3§</td>
</tr>
<tr>
<td>Naloxone</td>
<td>NST</td>
</tr>
<tr>
<td>Nifedipine</td>
<td>ND</td>
</tr>
<tr>
<td>Ondansetron‡</td>
<td>0.3 ± 0.3§</td>
</tr>
<tr>
<td>Propafenone</td>
<td>5.9 ± 3.3</td>
</tr>
<tr>
<td>Propranolol</td>
<td>1.6 ± 1.5§</td>
</tr>
<tr>
<td>Sildenafil</td>
<td>6.6 ± 2.1</td>
</tr>
<tr>
<td>Theophylline‡</td>
<td>NST</td>
</tr>
<tr>
<td>Tolbutamide‡</td>
<td>ND</td>
</tr>
<tr>
<td>Verapamil</td>
<td>7.8 ± 7.4</td>
</tr>
<tr>
<td>Warfarin‡</td>
<td>0.3 ± 0.3§</td>
</tr>
</tbody>
</table>

§ Some measurements of intrinsic clearance did not show significant turnover, the calculation of the average value utilized the LOQ; †Low clearance set.
Table 3. Predicted and observed clearance and fold error of the prediction for all tested compounds in the different cell systems. ND; not determined, NA; not applicable, NST; no significant turnover

<table>
<thead>
<tr>
<th>Compound</th>
<th>Predicted CL [mL/min/kg]</th>
<th>Observed CL [mL/min/kg]</th>
<th>Fold error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HeLaRG</td>
<td>SCC</td>
<td>Suspension</td>
</tr>
<tr>
<td>Chlorpromazine</td>
<td>17.24</td>
<td>15.61</td>
<td>12.65</td>
</tr>
<tr>
<td>Desipramine</td>
<td>ND</td>
<td>ND</td>
<td>13.88</td>
</tr>
<tr>
<td>Diazepam</td>
<td>0.11§</td>
<td>0.31</td>
<td>NST</td>
</tr>
<tr>
<td>Diclofenac</td>
<td>12.09</td>
<td>9.82</td>
<td>8.47</td>
</tr>
<tr>
<td>Domperidone</td>
<td>12.79§</td>
<td>8.16</td>
<td>11.84</td>
</tr>
<tr>
<td>Etodolac</td>
<td>1.68</td>
<td>1.52</td>
<td>0.65§</td>
</tr>
<tr>
<td>Gemfibrozil</td>
<td>ND</td>
<td>3.45</td>
<td>1.28§</td>
</tr>
<tr>
<td>Glipizide</td>
<td>0.23§</td>
<td>0.35§</td>
<td>0.85§</td>
</tr>
<tr>
<td>Imipramine</td>
<td>11.88</td>
<td>13.39</td>
<td>15.08</td>
</tr>
<tr>
<td>Ketoprofen</td>
<td>2.49</td>
<td>4.74</td>
<td>1.94</td>
</tr>
<tr>
<td>Metoprolol</td>
<td>12.00§</td>
<td>16.21</td>
<td>17.19</td>
</tr>
<tr>
<td>Naloxone</td>
<td>NST</td>
<td>18.38</td>
<td>17.30</td>
</tr>
<tr>
<td>Nifedipine</td>
<td>ND</td>
<td>13.98</td>
<td>3.86§</td>
</tr>
<tr>
<td>Ondansetron</td>
<td>9.65§</td>
<td>6.90§</td>
<td>12.85§</td>
</tr>
<tr>
<td>Propafenone</td>
<td>14.70</td>
<td>15.19</td>
<td>14.97</td>
</tr>
<tr>
<td>Propranolol</td>
<td>13.81§</td>
<td>15.19§</td>
<td>16.21</td>
</tr>
<tr>
<td>Sildenafil</td>
<td>16.14</td>
<td>11.43§</td>
<td>12.94</td>
</tr>
<tr>
<td>Theophylline</td>
<td>NST</td>
<td>ND</td>
<td>NST</td>
</tr>
<tr>
<td>Tolbutamide</td>
<td>ND</td>
<td>0.55</td>
<td>1.44§</td>
</tr>
<tr>
<td>Verapamil</td>
<td>17.86</td>
<td>17.28</td>
<td>16.30</td>
</tr>
<tr>
<td>Warfarin</td>
<td>0.24§</td>
<td>0.18§</td>
<td>NST</td>
</tr>
</tbody>
</table>

§ Some measurements of intrinsic clearance did not show significant turnover, the calculation of the average value utilized the LOQ; §Low clearance set
FIGURES

Figure 1. Overview of the tested hepatic in vitro systems. Suspension culture, a HepaRG sandwich culture, the Hµrel stochastic co-culture, the Hepatopac micropatterned co-culture (MPCC) and a micro-array hepatic spheroid culture were used in this study. Suspension culture was tested with 0.5 million cells/mL, HepaRG with 0.3 million hepatocyte-like cells/mL, stochastic co-culture with 0.4 million PHH/mL and stromal cells and MPCC with 0.05 million PHH/mL and stromal cells. A micro-array culture of PHH spheroids (500 cells per spheroid) was used as a 3D culture model. For the suspension culture, the maximum incubation time was 110 min, whereas for the other cultures it was 72h.

Figure 2. Robustness of CLint determination using the different cell models. (A) Percentage of compounds showing significant turnover in all repeats (blue), some repeats (yellow), or none of the repeats (purple). (B) Raw CLint values of compounds which were measured in all in vitro systems. Measurements without significant turnover indicated by LOQ are highlighted with a purple arrow. (C) Compound turnover plots of Diazepam of all tested in vitro models. (D) Coefficient of variation (CV) of raw CLint values of all tested compounds measured in HepaRG (purple), stochastic co-culture (blue), suspension (red), MPCC (yellow) and multi-spheroids (green). (E) CV of raw CLint values of low clearance compounds measured in the stochastic co-culture (blue), MPCC (yellow) and multi-spheroids (green).
Figure 3. **In vitro** to **in vivo** correlation of CLint for the different cell models. Correlations are shown for HepaRG (A), stochastic co-culture (B), suspension culture (C), MPCC (D) and multi-spheroid culture (E). The regression line (black line) is calculated based on CLint measurements with significant turnover (black triangles). The 95% confidence interval is represented as a dotted line. **In vitro** predictions were additionally made considering compounds without significant turnover by the calculated LOQ value (open triangle). If some but not all measurements showed significant turnover triangles are marked in orange, if all measurements failed to show significant turnover, triangles are marked in purple. The proportion of measurements with significant turnover is depicted in the pie chart (purple part). **In vitro** CLint values are shown as mean ± S.D. The grey line represents the line of unity.

Figure 4. **In vitro** to **in vivo** correlation of CL after regression correction for the different cell models. Correlations are shown for HepaRG (A), stochastic co-culture (B), suspension culture (C), MPCC (D) and multi-spheroid culture (E). Measurements where all repeats showed significant turnover are shown in black. Clearance values (CL) which are based on measurements where at least one repeat failed to show significant turnover are represented with an arrow down, indicating the true CL value being lower. If at least one repeat was showed significant turnover, arrows are shown in orange, if none of the measurements showed significant turnover and the LOQ was used for calculations, arrows are shown in purple. The grey line represents the line of unity. The dashed and dotted black lines represent the 3- and 2-fold prediction bands, respectively.
Figure 1

Suspension culture
- Cells: Primary human hepatocytes
- Donor: Single
- Hepatocyte concentration: 0.5x10^6 cells/mL
- Incubation time: 110 min

HepaRG sandwich culture
- Cells: HepaRG cell line
- Donor: Single
- Hepatocyte concentration: 0.3x10^6 cells/mL
- Incubation time: 72 h

Stochastic co-culture
- Cells: Primary human hepatocytes + stromal cell line
- Donor: 5-donor pool
- Hepatocyte concentration: 0.375x10^6 cells/mL
- Incubation time: 72 h

Micropatterned co-culture
- Cells: Primary human hepatocytes + mouse fibroblasts
- Donor: Single
- Hepatocyte concentration: 0.05x10^6 cells/mL
- Incubation time: 72 h

Microarray spheroid culture
- Cells: HepaRG cell line
- Donor: Single
- Hepatocyte concentration: 0.3x10^6 cells/mL
- Incubation time: 72 h
Figure 2

A

% of measured compounds

No sig. turnover
Occasional sig. turnover
Always sig. turnover

Donor
HepaRG
Stochastic co-culture
Suspension
MPCC
Multi-spheroids

B

Chlorpromazine
Diazepam
Diclofenac

Domperidone
Etodolac
Glipizide

Imipramine
Ketoprofen
Metoprolol

Ondansetron
Propafenone
Sildenafil

Verapamil
Warfarin
Propranolol

C

HepaRG
Stochastic co-culture
Suspension
MPCC
Multi-spheroids

Raw CLint [mL/min/10^6 hepatocytes]

D

CV% of raw CLint

Donor
HepaRG
Stochastic co-culture
Suspension
MPCC
Multi-spheroids

E

CV% of raw CLint

Donor
A
B
B

Stochastic co-culture
MPCC
Multi-spheroids

This article has not been copyedited and formatted. The final version may differ from this version.
Figure 3

A. HepaRG

\[ Y = 0.6X - 1.1 \quad \text{RMSE} = 1.65 \]

B. Stochastic co-culture

\[ Y = 0.7X - 0.9 \quad \text{RMSE} = 1.31 \]

C. Suspension CRT

\[ Y = 0.6X - 0.2 \quad \text{RMSE} = 1.01 \]

D. MPCC CRT

\[ Y = 0.5X - 0.3 \quad \text{RMSE} = 1.06 \]

E. Multi-spheroids

\[ Y = 0.5X - 0.5 \quad \text{RMSE} = 1.31 \]
Figure 4

- A: HepaRG
- B: Co-culture
- C: Suspension CRT
- D: MPCC CRT
- E: MWH CRT

The plots compare the observed and predicted clearance (CL) values for different systems. The data points are scattered along the diagonal line, indicating a good correlation between observed and predicted values.