1521-0103/360/1/95–105$25.00
THE JOURNAL OF PHARMACOLOGY AND EXPERIMENTAL THERAPEUTICS

http://dx.doi.org/10.1124/jpet.116.237495
J Pharmacol Exp Ther 360:95–105, January 2017

Quantitative Assessment of Population Variability in Hepatic Drug Metabolism Using a Perfused Three-Dimensional Human Liver Microphysiological System[S]

N. Tsamandouras,[1] T. Kostrzewski, C. L. Stokes, L. G. Griffith, D. J. Hughes, and M. Cirit

Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts (N.T., L.G.G., M.C.); CN Bio Innovations, Hertfordshire, United Kingdom (T.K., D.J.H.); and Stokes Consulting, Redwood City, California (C.L.S.)

Received August 30, 2016; accepted October 17, 2016

ABSTRACT

In this work, we first describe the population variability in hepatic drug metabolism using cryopreserved hepatocytes from five different donors cultured in a perfused three-dimensional human liver microphysiological system, and then show how the resulting data can be integrated with a modeling and simulation framework to accomplish in vitro–in vivo translation. For each donor, metabolic depletion profiles of six compounds (phenacetin, diclofenac, lidocaine, ibuprofen, propranolol, and prednisolone) were measured, along with metabolite formation, mRNA levels of 90 metabolism-related genes, and markers of functional viability [lactate dehydrogenase (LDH) release, albumin, and urea production]. Drug depletion data were analyzed with mixed-effects modeling. Substantial interdonor variability was observed with respect to gene expression levels, drug metabolism, and other measured hepatocyte functions. Specifically, interdonor variability in intrinsic metabolic clearance ranged from 24.1% for phenacetin to 66.8% for propranolol (expressed as coefficient of variation). Albumin, urea, LDH, and cytochrome P450 mRNA levels were identified as significant predictors of in vitro metabolic clearance. Predicted clearance values from the liver microphysiological system were correlated with the observed in vivo values. A population physiologically based pharmacokinetic model was developed for lidocaine to illustrate the translation of the in vitro output to the observed pharmacokinetic variability in vivo. Stochastic simulations with this model successfully predicted the observed clinical concentration-time profiles and the associated population variability. This is the first study of population variability in drug metabolism in the context of a microphysiological system and has important implications for the use of these systems during the drug development process.

Introduction

During preclinical drug development, prediction of hepatic clearance is of significant importance to set the first human dose and guide the selection of dosage regimens that achieve drug concentrations within the therapeutic window. However, to efficiently design clinical studies, it is crucial to predict hepatic drug metabolism and human pharmacokinetics, not only at the level of an "average individual" but also accounting for the associated population variability (Jamei et al., 2009).

To investigate interindividual variability in vitro, drug-metabolism assays must be performed independently for hepatocytes obtained from different donors and with an appropriate statistical analysis of the obtained data to estimate the interdonor variability in intrinsic clearance disentangled from measurement error/uncertainty and any other types of variability (e.g., interwell). Additionally, the estimate of interindividual variability in intrinsic metabolic clearance, as assessed in vitro, needs to be coupled with the population variability associated with other physiologic processes in vivo (e.g., hepatic blood flow, drug binding to plasma proteins, etc.). Thus, to perform in vivo predictions at the population level, a systems pharmacology approach (Trame et al., 2016) is desirable, where the in vitro results are integrated into population physiologically based pharmacokinetic (PBPK) models (Rostami-Hodjegan, 2012; Jones and Rowland-Yeo, 2013; Tsamandouras et al., 2015a,b).

Several in vitro systems have been traditionally applied to study drug metabolism, including human liver microsomes (Obach, 1999) and cryopreserved human hepatocyte suspensions (Brown et al., 2007). Although these systems have been very valuable in drug development (Di et al., 2012) and are easy to use, they lose metabolic activity over time, hence the study of low-clearance compounds is challenging (Di and Obach, 2015; Hutzler et al., 2015). In addition, overall these in vitro systems tend to underpredict in vivo clearance (Hallifax et al., 2010). Finally, drug-metabolism studies are often performed in human liver microsomes or hepatocytes pooled from several donors. Thus, the predicted intrinsic clearance refers to an "average individual," and the associated interindividual variability is not obtained.

This work was supported by the DARPA Microphysiological Systems Program [Grant W911NF-12-2-0039] and the National Institutes of Health Microphysiological Systems Program [Grant 4-UH3-TR000496-03].
[1]N.T. and T.K. contributed equally to this work.
dx.doi.org/10.1124/jpet.116.237495.
[S] This article has supplemental material available at jpet.aspetjournals.org.

ABBREVIATIONS: CV, coefficient of variation; 3D, three-dimensional; IDV, interdonor variability; IWV, interwell variability; LDH, lactate dehydrogenase; MPS, microphysiological system; P450, cytochrome P450; PBPK, physiologically based pharmacokinetic; PCR, polymerase chain reaction; PT, parallel tube; WS, well stirred.

Recently, new hepatic in vitro culture models have emerged to improve physiologic responses and mitigate the rapid loss of metabolic function typically observed in culture, thus offering opportunity to improve predictions of human drug clearance, especially for low-clearance compounds. Whereas static two-dimensional cocultures of hepatocytes with other cell types show some stabilization of function, a variety of three-dimensional (3D) culture models incorporating perfusion flow exhibit prolonged viability and function under serum-free conditions (Ebrahimkhani et al., 2014). Such 3D perfused models of liver and other tissues, where microfluidic or microscale reactors are used to control the flow of culture medium, are often termed "microphysiological systems" (MPS) or "organs on chips." In this work, we use a particular well developed, commercially available microreactor system for 3D perfused liver culture, the LiverChip (CN Bio Innovations, Hertfordshire, UK) (Dash et al., 2009; Domansky et al., 2010; Sarkar et al., 2015; Vivares et al., 2015). The LiverChip comprises a scaffold that fosters formation of an array of ~0.2-mm 3D tissue structures from primary human liver cells, and an on-board microfluidic pumping system, driven by pneumatics, that precisely perfuses the scaffold with culture medium to control oxygenation and shear stress on the tissue, enabling long-term culture with retention of physiological responses (Dash et al., 2009; Domansky et al., 2010; Vivares et al., 2015). The LiverChip platform, seeded with hepatocytes or with mixtures of hepatocytes and nonparenchymal cells, has been applied to analyze drug metabolism, inflammatory effects, drug-drug interactions, and as a model of breast cancer metastasis to the liver (Wheeler et al., 2014; Sarkar et al., 2015; Vivares et al., 2015; Long et al., 2016).

The current work first focuses on the in vitro assessment of population variability in drug metabolism and the relationship of that with variability in other phenotypic metrics, such as production of liver-specific factors (e.g., albumin, urea) and expression of metabolism-related genes. In vitro experiments were performed in a liver microphysiological system, housed in the LiverChip platform, utilizing hepatocytes from different donors. The ability of pooled hepatocytes to recapitulate the average phenotype across the different donors was also assessed. Since the purpose of investigating metabolism in vitro is to predict human pharmacokinetics, we follow up by integrating the generated data with a computational modeling and simulation framework to attempt in vitro to in vivo translation of population variability in drug metabolism.

Materials and Methods

3D Hepatocyte Tissue Culture. All sources of chemicals and reagents used are reported in the Supplemental Material (section 1.1). Cryopreserved human primary hepatocytes from five different donors were purchased from Life Technologies (Paisley UK). All donors (three males, two females) were Caucasians spanning a 21–72 years age range. Cells were recovered according to the supplier's instructions. Viability was assessed using trypan blue exclusion and was >85% for all lots. Hepatocyte suspensions were seeded (6×10^5 cells per scaffold) into scaffolds housed in the LiverChip in a total volume of 1.6 ml per compartment. Primary human hepatocytes from the five different donors were cultured, as well as a pooled hepatocyte sample, which contained equal numbers of cells from each of the five donors. In total, 21 wells were seeded for each donor and the pooled sample, from which 18 were intended for the drug-metabolism study (6 compounds × 3 replicate wells) and three were intended to be sacrificed at day 6 for RNA analysis. After the initial attachment period, cells undergo morphogenesis to form an array of 3D microtissues within the channels of the scaffold over a period of 3 days. Cells were maintained in Williams' E medium containing primary hepatocyte thawing and plating supplements (Life Technologies) for the first day of culture. Maintenance supplements (Life Technologies), which are serum-free, were used thereafter. All cultures were maintained in a standard humidified atmosphere at 37°C with 5% CO_2 and had a first complete medium change at 24 hours, then after a further 72 hours.

Hepatocyte Culture Phenotypic Characterization. Albumin and urea production as well as lactate dehydrogenase (LDH) release were measured before the drug-metabolism study (4 days postseeding). Albumin production was measured in supernatant using a human albumin enzyme-linked immunosorbent assay (Assay Pro, St. Charles, MO). Urea was quantified with a colorimetric assay kit (BioAssay Systems, Hayward, CA) and LDH secretion was measured using the CytoTox 96 nonradioactive cytotoxicity assay (Promega, Southampton, UK). Albumin, urea, and LDH were also measured postdose at the end of the drug-metabolism study (day 5 for wells treated with phenacetin; day 6 for wells treated with diclofenac, propranolol, lidocaine, and ibuprofen; and day 7 for wells treated with prednisolone). At the end of the experiment, the scaffolds/tissues were removed and washed with phosphate-buffered saline. Bright field images were taken using an inverted light microscope (Leica, Milton Kaynes, UK).

RNA Isolation and Gene Expression Analysis. Total RNA was extracted from freshly thawed hepatocytes or from LiverChip scaffolds cultured for 6 days (run in parallel to the drug-metabolism study without the addition of any drug) using TRIzol Reagent (Ambion, Loughborough, UK) and a chloroform phase separation. Quantitative polymerase chain reaction (PCR) was performed using SYBR Green PCR Master Mix (Applied Biosystems, Loughborough, UK) and primers designed against transcripts related to hepatic genes of specific interest (Supplemental Table S1). Samples were analyzed using a QuantStudio 6 real-time PCR system (Applied Biosystems, UK). C_t values from samples were compared and normalized to GAPDH expression. Samples were also analyzed by RT^2 Profiler PCR Arrays (Qiagen, Manchester, UK). Reverse transcription was performed using RT^2 First Strand Kit, and cDNA was analyzed by Human Drug Metabolism (PAHS-002ZC-12) RT^2 Profiler PCR Arrays. C_t values from samples were compared and normalized to the average expression across five different housekeeping genes (ACTB, B2M, GAPDH, HPRT1, and RPLP0). The quantitative PCR and super-array data (referring to 6 and 84 genes, respectively) were merged to generate a data set of 90 genes, the expression of which was investigated. The complete methods regarding RNA isolation and the gene-expression analysis are provided in the Supplemental Material (section 1.2).

Drug-Metabolism Study. At day 4, six different compounds (phenacetin, diclofenac, lidocaine, ibuprofen, propranolol, and prednisolone) were added as a bolus dose to the microtissues at an initial concentration of 1 μM during a full medium change (final volume per well was 1.8 ml), and the final solvent (dimethylsulfoxide) concentration never exceeded 0.1% (v/v). Supernatant samples (60 μl) (i.e., corresponding to the extracellular compartment) were taken at predetermined postdose sampling times (0, 1, 4, 6, and 24 hours for phenacetin; 0, 1, 4, 24, and 48 hours for diclofenac, propranolol, lidocaine, and ibuprofen; and 0, 4, 24, 48, and 72 hours for prednisolone). For the time 0 measurements, a sample was taken from each culture well immediately (instantaneously) after the addition of the drug-containing medium to accurately evaluate the initial drug concentration in each well. All samples were analyzed for the presence of the dosed compound. Additionally, metabolite formation was measured for phenacetin (acetaminophen), diclofenac (4-OH-diclofenac), ibuprofen (2-OH-ibuprofen), and prednisolone (6β-OH-prednisolone). Extensive details regarding sample preparation and the liquid chromatography–tandem mass spectrometry analysis are provided in the Supplemental Material (section 1.3).

Investigation of Drug Binding to the Platform and Tissue Culture Medium. The compounds were also added to LiverChip wells containing no hepatocytes to analyze nonspecific binding to plate components. Each well was treated in the same way as the wells containing microtissues. Sixty-microliter samples were taken from these plates at 0, 1, and 48 hours postdosing, and samples were treated and analyzed for the presence of the dosed compound as described in the Supplemental Material (section 1.3).

Unbound drug fraction in the bovine serum albumin-containing cell culture medium was quantified by rapid equilibrium dialysis. Details are provided in the Supplemental Material (section 1.4).

Pharmacokinetic Analysis of the Drug Depletion Data. The drug depletion data corresponding to individual donors were analyzed with a population pharmacokinetic modeling approach using nonlinear mixed-effects modeling software (NONMEM 7.3; ICON Development Solutions, Ellicott City, MD) and the first-order conditional estimation method with interaction. A one-compartment pharmacokinetic model (model of monoexponential decay) was used to fit the drug depletion data (eq. 1):

$$C^{ijk} = C_0^{ij} \cdot e^{-\left(\frac{fu_{med} \cdot CL_{int(u)}^{ij}}{V_{med}}\right) \cdot t_{ijk}} \quad (1)$$

where C^{ijk} is the model prediction for the kth observed concentration regarding the ith donor and the jth well, sampled at time t_{ijk};C_0^{ij} is the substrate concentration in the medium at time 0 regarding the ith donor and the jth well (experimentally measured by sampling each well instantaneously after the addition of the drug-containing medium); V_{med} is the volume of the medium during the substrate depletion experiment (1.8 ml); fu_{med} is the fraction of drug which is unbound in the medium and thus available for metabolism (experimentally determined with equilibrium dialysis); and $CL_{int(u)}^{ij}$ is the unbound intrinsic clearance regarding the ith donor and the jth well. Both interdonor variability (IDV) and interwell variability (IWV) were taken into account during the estimation of unbound intrinsic clearance using an exponential relationship (see eq. 2), which assumes that clearance is log-normally distributed:

$$CL_{int(u)}^{ij} = CL_{int(u)} \cdot e^{\eta_i + \kappa_{ij}} \quad (2)$$

where $CL_{int(u)}$ is the typical (median) value of unbound intrinsic clearance across all donors/wells; η_i is the random effect referring to between-donor differences (thus IDV); and κ_{ij} is the random effect referring to between-well differences within a donor (thus IWV). Both η and κ are assumed to be independently normally distributed with mean 0 and variance $\omega^2$and π^2, respectively. A common variance was assumed for all wells using the SAME option in NONMEM. Finally, an additional level of variability was taken into account, termed residual variability, using an additive error model on the scale of the log-transformed observations/predictions (see eq. 3):

$$\ln\left(C_{obs}^{ijk}\right) = \ln\left(C^{ijk}\right) + \varepsilon_{ijk} \quad (3)$$

where C_{obs}^{ijk} is the kth observed concentration regarding the ith donor and the jth well, sampled at time t_{ijk};C^{ijk} is the respective model prediction (see eq. 1);$\varepsilon_{ijk}$is the random effect referring to residual variability and thus the differences between the observed concentrations and the model predictions due to unexplained factors (e.g., measurement/assay error, model misspecification, etc.); and ε is assumed to be normally distributed with mean 0 and variance σ^2. Typical goodness-of-fit plots (e.g., observations versus predictions, residuals versus time/predictions) and simulation-based diagnostics (e.g., visual predictive checks) were used to detect the adequacy of the developed mixed-effects models (Karlsson and Savic, 2007).

The drug depletion data corresponding to the pooled hepatocytes where subsequently analyzed in a similar framework, with the difference that, in the pooled hepatocytes data, the only level of variability in $CL_{int(u)}$ is the IWV.

Identification of In Vitro Clearance Predictors. The normalized (see Supplemental Material, section 1.5) values of intrinsic clearance obtained across different donors/wells were investigated in relation to the values of the respective predose phenotypic metrics (albumin/urea production, LDH release) and donor-specific mRNA levels of the primary for the metabolism of each compound cytochrome P450 (P450). All methodological details regarding this investigation are provided in the Supplemental Material (section 1.5).

Prediction of In Vivo Hepatic Clearance. The typical values of unbound intrinsic clearance [$CL_{int(u)}$] determined for each compound from the pharmacokinetic analysis of the individual-donor in vitro data were subsequently used to derive a prediction with regard to the in vivo hepatic clearance. These predicted hepatic clearance values were then compared with clinically observed values, and the overall agreement was determined by the calculation of the average fold-error across all compounds. All of the related methodological details and the complete procedure (equations) used for in vitro–in vivo extrapolation of clearance are reported in the Supplemental Material (section 1.6).

Population In Vitro–In Vivo Translation with the Aid of PBPK Modeling. A population PBPK model for lidocaine was developed to illustrate the framework under which the liver MPS data can be translated to predictions of in vivo concentration-time profiles at the population level. The rationale for selection of lidocaine among the other in vitro evaluated compounds is described in the Supplemental Material (section 1.7). Each tissue/organ of the developed PBPK model was assumed to be a well stirred compartment with perfusion-limited kinetics, and the liver was considered as the only site of elimination (Supplemental Fig. S1). Renal clearance was assumed to be negligible, as only around 8% of the drug is excreted unchanged in urine (Benet et al., 2011). An empirical scaling factor (determined across all compounds evaluated in this work) was incorporated on the in vitro–determined hepatic clearance of lidocaine with the aim of correcting for any systematic underprediction of in vivo clearance with the used in vitro system (Supplemental Material, sections 1.6 and 1.10). The PBPK model was mathematically described with a system of 14 mass balance differential equations (Supplemental Material, section 1.8), which were solved in MATLAB R2015b (MathWorks, Inc., Natick, MA). Model simulations for 1000 "virtual individuals" were performed, taking into account population variability in both the system- and drug-related parameters of the model. All model parameters along with the exact methodology for the generation of the respective population distributions are described in the Supplemental Material (sections 1.9 and 1.10 for system- and drug-related parameters, respectively). This approach allowed the generation of 95% population prediction intervals associated with lidocaine concentration-time profiles in arterial plasma after a constant-rate i.v. infusion of lidocaine HCl, 3 mg/kg, over a 3-minute period. Model predictions were then compared with clinically observed arterial concentration-time profiles (Tucker and Boas, 1971) obtained under the same dosage regimen.

Results

Variability in Cell Culture Phenotype. Human serum albumin production, urea production, and LDH release to extracellular medium in the liver MPS were quantified to assess baseline cell health and phenotypic variability among the donors and wells prior to drug exposure (predose, day 4). Substantial and statistically significant differences in albumin production, urea production, and LDH release were observed across hepatocyte cultures from the five different donors (Fig. 1). These phenotypic metrics were also merged across the five different donors and compared with the equivalent metrics from the pooled hepatocytes. Although no difference was observed in albumin production, the pooled hepatocytes were associated with higher urea production and higher LDH levels (Fig. 1).

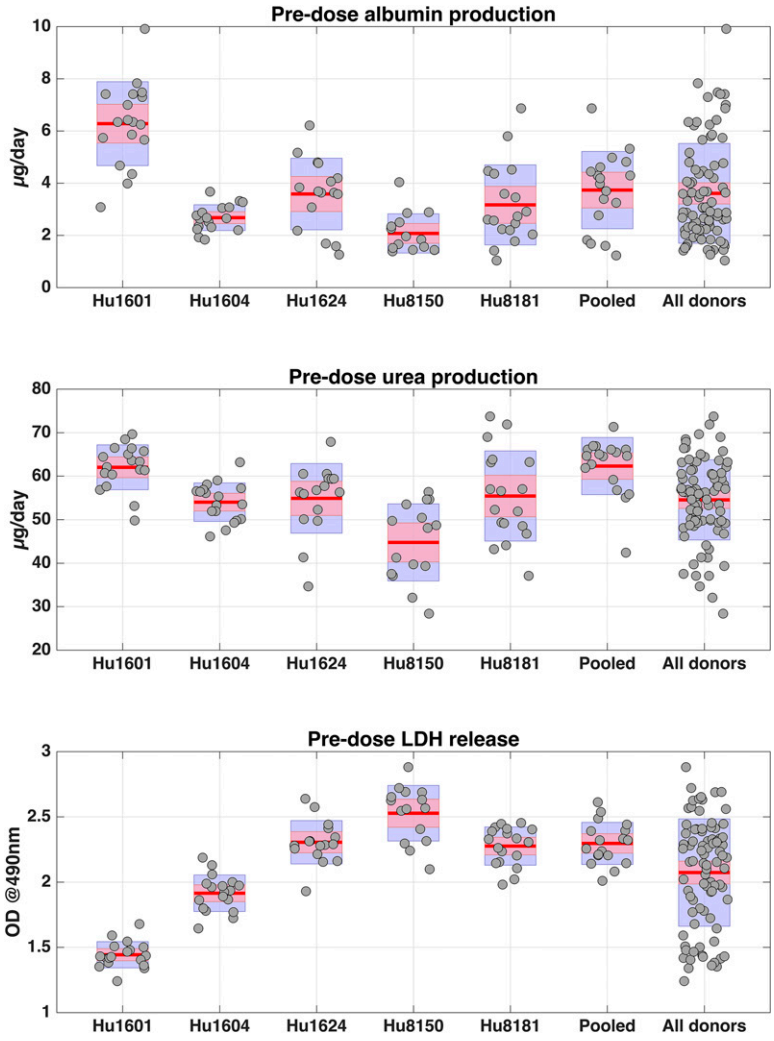

Fig. 1. Predose (measured at day 4) albumin, urea, and LDH levels stratified across different donors. Hu1601, Hu1604, Hu1624, Hu8150, and Hu8181 are lot numbers corresponding to five different donors. "Pooled" refers to the pool of hepatocytes from the five donors, and "All donors" refers to the data from all five donors merged together. Red lines correspond to the mean of the data, purple boxes extend the mean by ± 1 S.D., and pink boxes correspond to 95% confidence intervals around the mean. LDH levels are expressed in optical density (OD) units at 490 nm.

The same phenotypic metrics were also measured at the end of the drug-metabolism study (postdose, days 5–7). Hepatocyte cultures were clearly functional throughout the study, as demonstrated by continued albumin and urea production. More specifically, across all treatments, postdose albumin production was significantly increased and LDH release was significantly decreased compared with the equivalent predose levels (Fig. 2). Urea production also exhibited an overall trend of increase at postdose measurements for all treatments, except prednisolone. Further analysis supported that the aforementioned differences are likely due to the increased period in culture rather than due to a treatment (compound) effect, as the postdose metrics were not significantly different between treatments (see Supplemental Material, section 2.1). The stratification of these postdose phenotypic metrics, not only across treatments but also across different donors (Supplemental Figs. S2–S4), indicates continued significant interdonor variability during the drug-metabolism study. Last, a correlation matrix plot of all pre- and postdose phenotypic metrics measured across all different donors (or pool of donors) and wells illustrated that strong pairwise correlations may occur across these metrics (Supplemental Fig. S5).

The three-dimensional microtissue structures were visualized at the end of the experiment. The results show that the tissue formation was consistent/comparable across the different donors and was maintained throughout the culture period (Supplemental Fig. S6). All of the results regarding the quantitative/statistical analysis of the phenotypic metrics are described in the Supplemental Material (section 2.1).

Gene Expression. A comparison of gene expression (drug-metabolism related) signature between freshly thawed hepatocytes and the liver MPS showed a statistically significant difference for only 10 of the 90 investigated genes (four downregulated and six upregulated in the liver MPS; see Fig. 3). Thus, for the vast majority of the investigated metabolism-related genes, hepatocyte microtissues retain gene expression (6 days after seeding) at levels comparable to those in freshly thawed hepatocytes. The list of all genes investigated in this work, the average fold-changes in expression, and the associated statistical significance are reported in Supplemental Table S2. To visualize the changes in gene expression signature separately at the level of each donor (or pool of donors) and identify clusters of genes that are jointly up- or downregulated, agglomerative hierarchical clustering was performed (Supplemental Fig. S7). A number of genes were consistently down- or upregulated in the liver MPS compared with freshly thawed hepatocytes for all different donors; however, there were also genes that were diversely regulated across donors [some genes

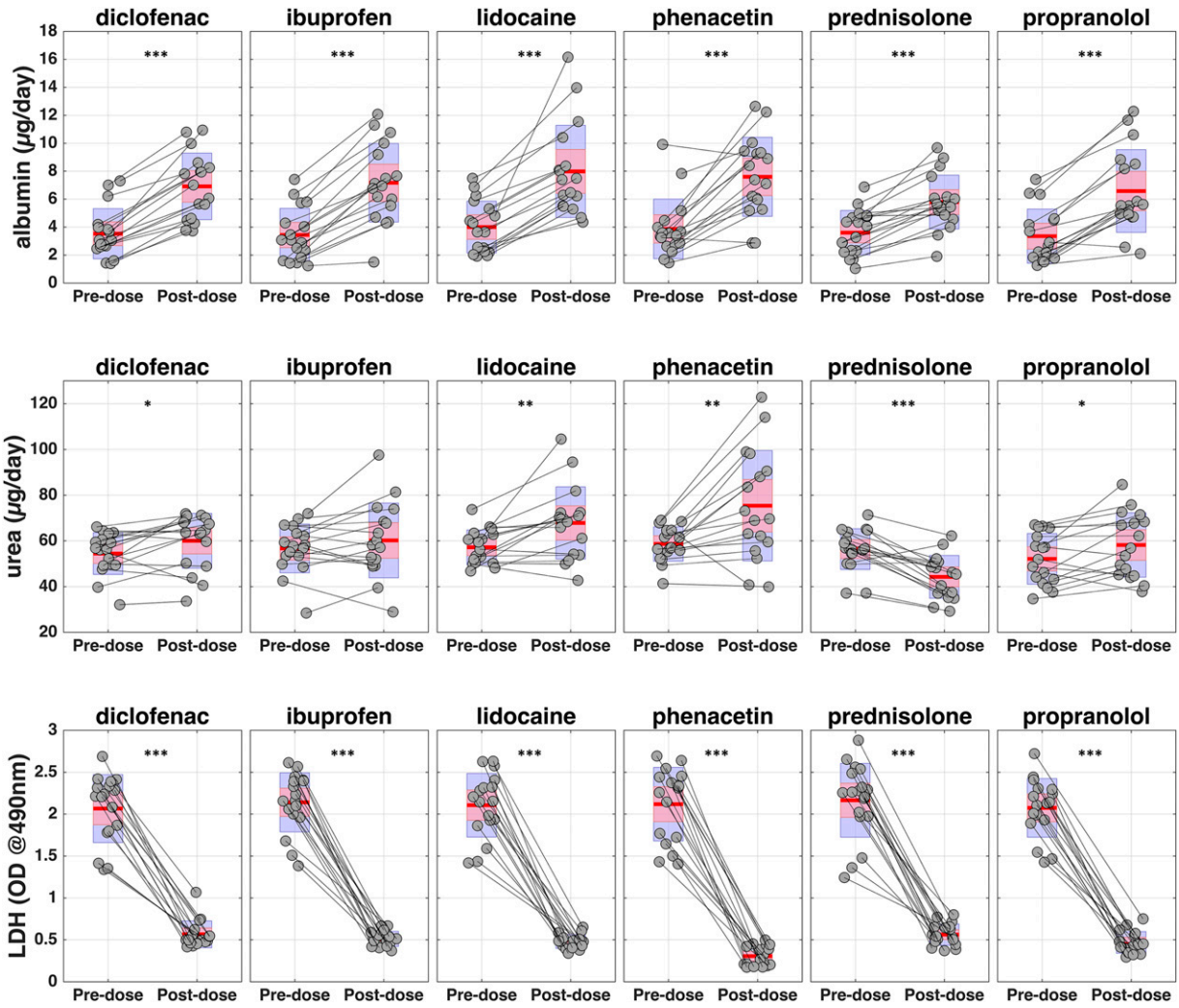

Fig. 2. Comparison between the predose (measured at day 4) and the postdose (measured at day 6 for diclofenac, propranolol, lidocaine, and ibuprofen and days 5 and 7 for phenacetin and prednisolone, respectively) albumin, urea, and LDH levels stratified across different treatments. Data from both the five donors and the pooled hepatocytes are shown. Red lines correspond to the mean of the data, purple boxes extend the mean by ± 1 S.D., and pink boxes correspond to 95% confidence intervals around the mean. Thin black lines connect the pre- and postdose levels in a given donor (or pool of donors) and well. Asterisks inside each subplot indicate significant differences between pre- and postdose levels (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$). LDH levels are expressed in optical density (OD) units at 490 nm.

(e.g., CYP3A5) were strongly upregulated in donor Hu8181 but downregulated in all other donors]. This further highlights the presence of interdonor variability.

Additionally, it was observed (Supplemental Fig. S8) that for the vast majority of genes, the magnitude of interdonor variability in mRNA expression levels was significantly larger in freshly thawed hepatocytes [average coefficient of variation (CV) across all genes was 72%] compared with the liver MPS (average CV across all genes was 34%).

Finally, of the 90 investigated genes, only five in the liver MPS and none in the freshly thawed hepatocytes were significantly differentially expressed in the pooled hepatocyte samples compared with the average values observed across the different donors (Supplemental Fig. S9). Thus, overall, there was no evidence against the argument that the mRNA expression levels obtained from the pooled hepatocytes can be considered as representative of the average mRNA expression obtained across the different donors. However, for the majority of genes in the liver MPS, there was a nonsignificant trend that the pooled hepatocytes marginally overpredict the average mRNA expression obtained across the different donors, whereas for the majority of genes in the freshly thawed hepatocytes, the opposite nonsignificant trend was observed (Supplemental Fig. S9).

Drug Binding to the Hepatocyte-Free LiverChip and Tissue Culture Medium. Nonspecific drug binding to the hepatocyte-free LiverChip platforms was evaluated for the drugs used in metabolism studies. The quantification of each drug showed no evidence of nonspecific drug binding to LiverChip components after 48-hour exposure (Supplemental Fig. S10). Therefore, nonspecific binding of the investigated compounds to the LiverChip materials was treated as negligible in the current work.

The rapid equilibrium dialysis analysis indicated that the extent of binding to cell culture media components (e.g., bovine serum albumin) varies substantially across the investigated compounds. The unbound fraction in media for diclofenac, ibuprofen, lidocaine, prednisolone, propranolol, and phenacetin was determined to be 0.13 (35% CV), 0.31 (10% CV), 0.88 (9% CV), 0.94 (4% CV), 0.98 (2% CV), and 0.98

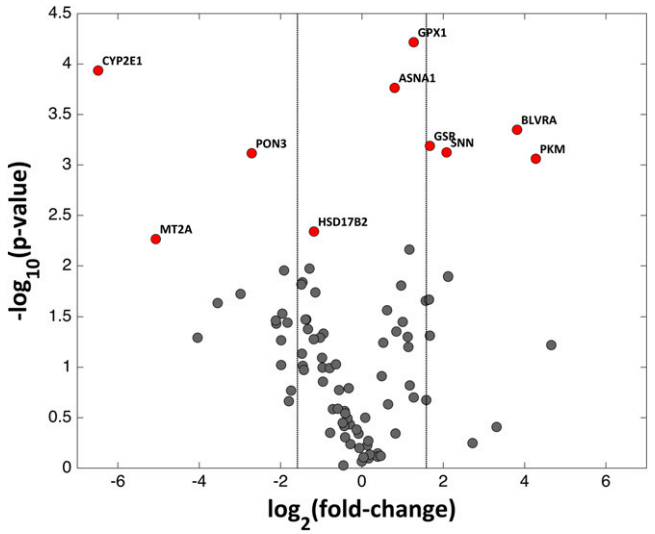

Fig. 3. Volcano plot that illustrates the average fold-change in gene expression between the liver MPS (day 6) and freshly thawed hepatocytes along with the associated statistical significance. The $\log_2$ of the fold-change is plotted on the *x*-axis; thus, positive values indicate upregulation in the liver MPS compared with the freshly thawed hepatocytes, whereas negative values indicate downregulation. Genes outside the two black vertical lines are up- or downregulated more than 3-fold. On the *y*-axis, the $-\log_{10}$ of the *P* value is plotted; thus, the higher values indicate stronger statistical evidence of a significant difference in gene expression between the liver MPS and freshly thawed hepatocytes. The genes for which significant differences were detected after multiple testing correction are highlighted in red, and the respective gene names are reported.

(1% CV), respectively (CV refers to coefficient of variation across triplicate experiments).

Pharmacokinetic Analysis of the Drug Depletion Data. All drug depletion data across different donors and wells available to the pharmacokinetic analysis are presented in Fig. 4 (see also Supplemental Material, section 2.2 for a numerical summary and Supplemental Fig. S11 for averaged concentration-time profiles for each donor across different wells). Substantial interdonor and interwell variability was observed in the metabolic depletion profiles of all compounds.

The results of the mixed-effects modeling of the individual-donor drug depletion data are presented in Table 1. The estimates of the typical intrinsic clearance [$CL_{int(u)}$] for the six investigated compounds ranged from 0.81 μl/min/10^6 cells for prednisolone to 17.8 μl/min/10^6 cells for diclofenac. These parameters were precisely estimated for all compounds with relatively low standard errors. The intrinsic clearance of all compounds was associated with substantial interdonor variability, and the respective CV% ranged from 24.1% for phenacetin to 66.8% for propranolol. Interwell variability (within donor) in intrinsic clearance was less pronounced than interdonor variability for all compounds except phenacetin (marginally higher IWV compared with IDV). The coefficient of variation with respect to the interwell variability in intrinsic clearance ranged from 6% for diclofenac to 32.9% for propranolol. The residual (unexplained) variability of the model regarding the observed concentrations was relatively small for all compounds, ranging from 8.4% CV for lidocaine to 21.6% CV for propranolol.

The developed mixed-effects models adequately reflect not only the average trend in the data but also the observed variability (Fig. 5). Additionally, the ability of this modeling approach to accurately describe the data not only in total but also at the level of each individual donor and well is illustrated in Supplemental Fig. S12 in the case of propranolol (the compound with the highest degree of interdonor and interwell variability). The one-compartment pharmacokinetic model that was assumed for drug depletion provided an adequate description of the data, as the majority of the compounds exhibited monoexponential declines in their concentration-time profiles, with the exception of diclofenac, for which a model of biexponential decline might be more appropriate (Figs. 4 and 5). However, an additional analysis (Supplemental Material, section 2.3) supported that the monoexponential decline assumption for diclofenac does not introduce any substantial bias for the purpose of this work.

The results regarding the analysis of the pooled hepatocyte drug depletion data are also presented in Table 1, and the adequacy of the model to describe the observed data is illustrated in Supplemental Fig. S13. Estimates of the typical intrinsic clearance [$CL_{int(u)}$] for the six investigated compounds in pooled hepatocytes ranged from 0.91 μl/min/10^6 cells for prednisolone to 18.6 μl/min/10^6 cells for diclofenac. A comparison with the equivalent clearance estimates determined from the individual-donor data indicates only minor differences (Table 1) and supports the notion that pooled hepatocytes can provide a relatively unbiased estimate of the average clearance in the donor population. More specifically, the ratio of $CL_{int(u)}$ determined in the individual donor data to the $CL_{int(u)}$ determined in the pooled hepatocyte data ranged from 0.61 for propranolol to 1.42 for ibuprofen, with an average of 0.97 across all compounds (see Table 1). In addition, the 95% confidence intervals associated with this ratio included 1 for all compounds with the sole exception of propranolol, which is consequently the only compound exhibiting marginal evidence of bias in the determination of clearance in the donor population by using pooled hepatocytes.

Metabolite Formation. Metabolite concentration-time profiles were determined for prednisolone, phenacetin, ibuprofen, and diclofenac (Supplemental Fig. S14). A strong correlation was observed between the intrinsic clearance for drug depletion in a given donor/well and the respective metabolite formation levels (Supplemental Fig. S15). More specifically, the linear regression R^2 values were very high for three of these compounds (0.91, 0.82, and 0.77 for prednisolone, ibuprofen, and diclofenac, respectively), whereas the correlation was weaker for phenacetin (R^2 squared = 0.4). These results indicate that the clear interdonor differences observed in drug depletion clearance are also reflected in the metabolite formation levels.

Identification of In Vitro Intrinsic Clearance Predictors. Predose albumin and urea production levels in a given donor/well were positively correlated with the respective intrinsic clearance values that were subsequently obtained from the drug-metabolism study (Supplemental Fig. S16). On the other hand, these intrinsic clearance values were negatively correlated with the predose LDH release levels, whereas P450 mRNA levels had only a marginal positive correlation (Supplemental Fig. S16). A Lasso regression model, in which several covariates are considered simultaneously, identified all of the previously discussed metrics (albumin, urea, LDH, and P450 mRNA levels) as significant predictors of in vitro intrinsic clearance (Supplemental Table S3 and Fig. S17). Although this model was able to account for a substantial part of the observed variability in intrinsic clearance values (R^2 = 0.52), there is still unexplained

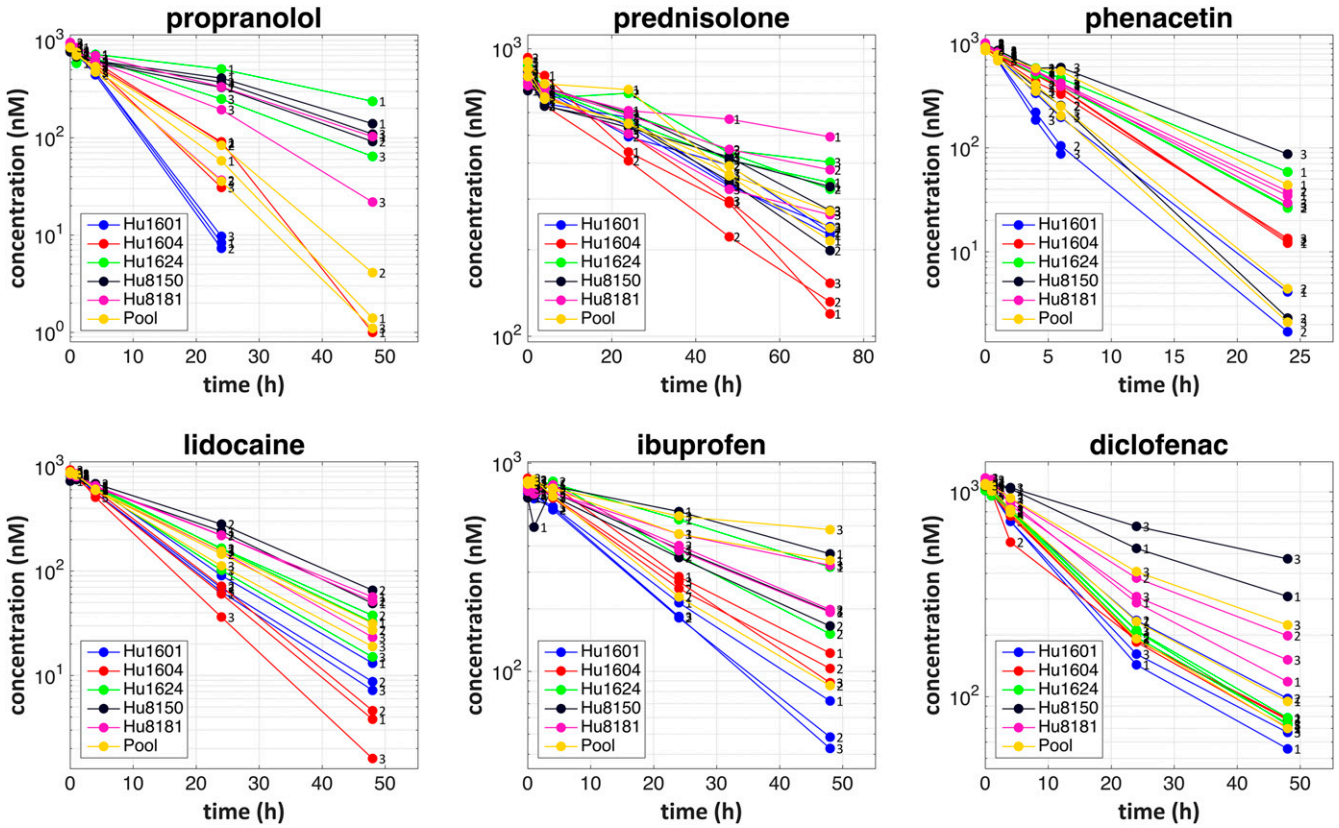

Fig. 4. Drug depletion data available for the pharmacokinetic analysis. Hu1601, Hu1604, Hu1624, Hu8150, and Hu8181 are lot numbers corresponding to five different donors. "Pool" refers to the pool of hepatocytes from the five donors. The small numbers on the right of each concentration point (values of 1, 2, or 3) aim to distinguish different wells across the same donor (or pool of donors).

variability that cannot be captured solely by these four predictors (Supplemental Material, Fig. S18).

Prediction of In Vivo Hepatic Clearance. The results regarding the agreement between the observed hepatic clearances in vivo and the predicted hepatic clearances from the in vitro data are graphically illustrated and numerically summarized in Supplemental Fig. S19 and Table S4, respectively. Predicted clearance values from the liver MPS study were strongly correlated with the observed in vivo values [linear regression R^2 values of 0.75 and 0.77, respectively, when the parallel tube (PT) or the well stirred (WS) liver model was used]. The average fold-error across all compounds (underprediction) was 4.2-fold and 4.5-fold when the PT or the WS liver model was used, respectively. The lowest degree of underprediction was observed for phenacetin (1.7-fold and 2.1-fold for the PT and WS models, respectively), and the highest, for propranolol (8.2-fold and 8.5-fold for the PT and WS models, respectively). By calculating the

TABLE 1
Parameter estimates from the modeling of the drug depletion data

Parameter	Propranolol	Prednisolone	Phenacetin	Lidocaine	Ibuprofen	Diclofenac
Individual-donor data						
$CL_{int(u)}$[a]	3.88 (28.6%)	0.81 (14.7%)	8.91 (12.7%)	4.38 (12.9%)	5.02 (16.2%)	17.80 (16.8%)
IDV[b]	66.8% (40.9%)	29.3% (58.1%)	24.1% (80.4%)	28.5% (38.1%)	32.6% (44.9%)	36.2% (71.5%)
IWV[b]	32.9% (62.7%)	21.5% (70.0%)	26.1% (65.2%)	11.3% (26.5%)	30.7% (55.8%)	6.0% (120.9%)
RV[b]	21.6% (56.6%)	10.3% (25.8%)	14.2% (69.8%)	8.4% (34.0%)	9.6% (29.3%)	20.0% (26.2%)
Pooled hepatocytes data						
$CL_{int(u)}$[a]	6.34 (5.9%)	0.91 (4.0%)	9.67 (17.4%)	4.24 (3.6%)	3.54 (33.3%)	18.60 (14.4%)
IWV[b]	9.6% (50.4%)	–[c]	30.7% (41.5%)	5.9% (43.7%)	62.5% (40.0%)	23.6% (54.4%)
RV[b]	19.7% (41.1%)	11.4% (32.0%)	9.5% (40.4%)	6.9% (14.4%)	7.5% (15.8%)	16.3% (28.6%)
Individual-donor/pooled hepatocytes						
$CL_{int(u)}$ ratio[d]	0.61 (0.27, 0.97)	0.89 (0.63, 1.16)	0.92 (0.61, 1.47)	1.03 (0.77, 1.31)	1.42 (0.75, 4.20)	0.96 (0.60, 1.48)

[a]The typical unbound intrinsic clearance [$CL_{int(u)}$] for each drug is reported in μl/min/10^6 cells.
[b]Interdonor variability in unbound intrinsic clearance (IDV), interwell variability in unbound intrinsic clearance (IWV), and the residual variability in the observed data (RV) are reported in terms of CV%, which was calculated as $\sqrt{(e^{variance}-1)}\cdot 100$, where "variance" is the estimate of ω^2, π^2, and σ^2 for IDV, IWV, and RV, respectively (see *Materials and Methods*). Values in parentheses correspond to relative standard errors calculated as (*standard error*/*estimate*)·100.
[c]Interwell variability could not be estimated and was fixed to 0.
[d]Ratio of $CL_{int(u)}$ determined in the individual donor data to the $CL_{int(u)}$ determined in the pooled hepatocytes data. Values in parentheses correspond to 95% confidence intervals of this ratio, calculated using Fieller's theorem and assuming normality of the $CL_{int(u)}$ estimators. The average $CL_{int(u)}$ ratio across all compounds is 0.97.

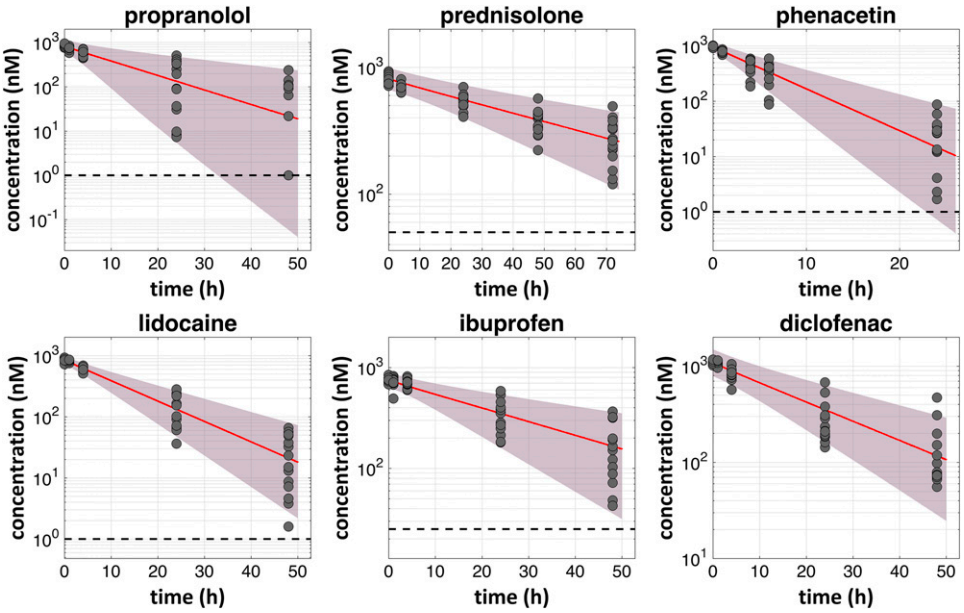

Fig. 5. Visual predictive checks of the developed mixed-effect models with regard to the observed individual-donor drug depletion data. Closed gray circles represent the observed concentrations in medium; highlighted with purple are the areas between the 5th and 95th percentiles of model simulations that take into account the different levels of variability (90% prediction intervals), whereas the red solid line represents their median (median prediction); the horizontal dashed black line represents the limit of quantification.

deviations between observations and predictions at the level of intrinsic clearances (see Supplemental Material, section 1.6), an empirical scaling factor of 5.4 or 8.7 was derived when the PT or the WS liver model was considered, respectively.

Population In Vitro–In Vivo Translation with the Aid of PBPK Modeling. The success of the population PBPK modeling approach used for in vitro–in vivo translation at the "population level" is illustrated in Fig. 6. The model predictions were in close agreement with the clinically observed data (Tucker and Boas, 1971), adequately capturing not only the average trend in the observed clinical data but also the extent of the associated interindividual variability. Minor disagreements between model predictions and observations are considered acceptable, as the model mainly utilizes in vitro/in silico information, and the observed concentration-time data have not been used to fit (estimate) any of the model parameters.

Discussion

Microphysiological systems have not been fully evaluated for quantitative pharmacology applications, such as prediction of hepatic drug metabolism. The current work focuses on the in vitro assessment of population variability in drug metabolism using a liver MPS and the subsequent translation to variability in pharmacokinetics in vivo using computational modeling and simulation methodologies. The overall framework used in this work (see schematic in visual abstract) represents our recommendation with regard to the analysis and the subsequent in vivo translation of in vitro data generated in microphysiological systems.

The generated output in such systems (e.g., drug depletion profiles) is a complex function of the characteristics of the in vitro system (e.g., number of cells, medium volume, and composition) and intrinsic biologic parameters (e.g., unbound intrinsic clearance for a given drug). Through model-based analysis of the in vitro output, we estimated the intrinsic biologic parameter (unbound intrinsic clearance) disentangled to the greatest possible degree from the in vitro system characteristics and any additional processes taking place in the platform (e.g., drug binding to medium components).

To estimate population variability associated with the intrinsic biologic parameter, the study was designed to capture drug depletion data across multiple donors and multiple wells for each donor. The statistical analysis of such multilevel longitudinal data is challenging and can be approached with different methods. However, the most suitable and unbiased method is through nonlinear mixed-effects modeling (Sheiner and Beal, 1981; 1983; Mould and Upton, 2013), as this simultaneously takes into account the different sources and levels of variability. It was demonstrated here that the in vitro–determined metabolic drug clearance varied substantially across hepatocytes from different donors. This highlights that clearance predictions for new compounds should be evaluated carefully when hepatocytes from a single donor only are used. It was also demonstrated that interwell variability in intrinsic clearance was generally lower than the associated interdonor variability, providing further confidence in microphysiological systems for future investigations of population variability in drug metabolism.

The unbound intrinsic clearance along with the associated interdonor variability obtained from the liver MPS can be scaled up and integrated with the characteristics of the in vivo system (hepatic blood flow, organ volumes, etc.) and their respective population variability through the use of PBPK modeling and the performance of stochastic simulations (Jones and Rowland-Yeo, 2013; Tsamandouras et al., 2015b,c). The lidocaine case study illustrated the details of this approach, and to our knowledge, this work is the first to combine experimental liver MPS data with a computational systems pharmacology framework to perform in vivo pharmacokinetic predictions. The accurate prediction of the clinically observed population variability in lidocaine plasma concentration-time profiles provides further confidence in the value of this combined experimental and computational approach.

Interdonor variability was also investigated in the liver MPS with respect to additional phenotypic levels. Specifically, secreted and released biomolecules (albumin, urea, LDH) further

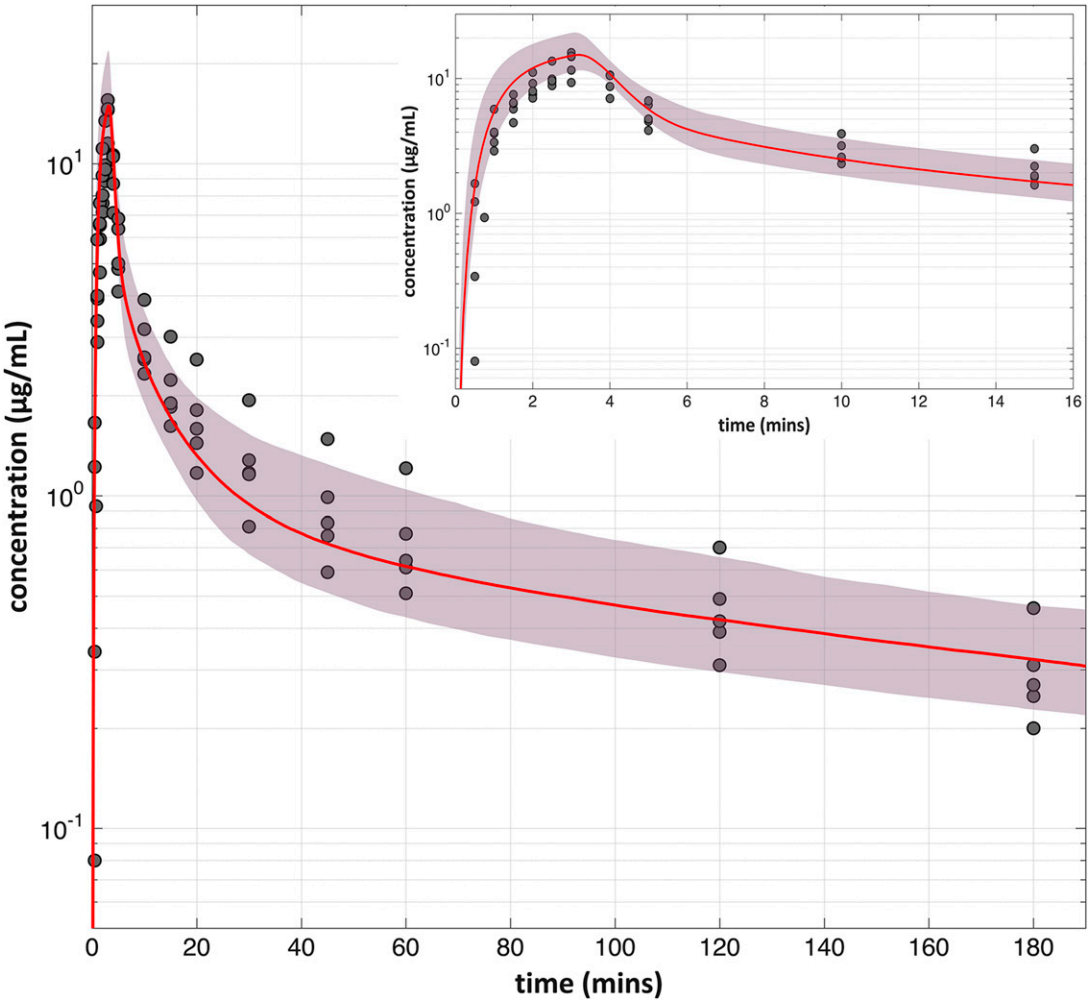

Fig. 6. Population PBPK model prediction of lidocaine arterial plasma concentrations during and after a constant-rate i.v. infusion (lidocaine HCl, 3 mg/kg for 3 minutes). The clinically observed data represented with closed gray circles were extracted from Tucker and Boas (1971) across five different subjects. The shaded area corresponds to the 95% population prediction intervals of the model, and the red line corresponds to the median model prediction. The insert plot magnifies the first 16 minutes for the purpose of clarity.

highlighted the donor variability in terms of culture functionality and viability. Interestingly, we observed a clear correlation between the levels of these biomolecular markers before drug administration and the subsequently determined drug clearance in the respective donors/wells. Although accurate cell number quantification in MPS technologies is a challenge, it is essential for quantitative pharmacology studies. In the current study, visual inspection of phase-contrast images indicated that seeding across different wells was consistent and equally successful across hepatocytes from different donors. In comparison, the extent of the interdonor/interwell differences observed in this work in drug clearance and other biomolecular metrics (albumin, urea, LDH) is much more pronounced, indicating that the aforementioned differences and correlations mainly arise through the MPS biology and are not simply a reflection of differences in attached cell numbers on the scaffolds.

The mRNA expression of metabolism-related genes exhibited substantial diversity across different donors. Interestingly, interdonor variability in mRNA expression levels was significantly lower in liver MPS cultured hepatocytes (6 days after seeding) compared with freshly thawed hepatocytes. We hypothesize that this is due to adaptation to the much more controlled and consistent environment of stimuli/cues (medium composition, flow, oxygen gradient, etc.) present in the liver MPS culture. Finally, although P450 mRNA level was identified in conjunction with other phenotypic metrics (albumin, urea, LDH) to be a predictor of intrinsic metabolic clearance, it accounted for only a very small portion of the clearance variability. Thus, screening mRNA expression of metabolic enzymes across different donors should not be used as a surrogate marker for interdonor variability in metabolic activity.

The retention of hepatocyte viability and functionality in the liver MPS for the entire period of the study (up to 7 days) was also clearly demonstrated. Specifically, at the end of the drug-metabolism study, not only were albumin and urea produced in high levels and LDH secretion was low, but these metrics were also substantially improved compared with predose (day 4) determinations. The decrease in LDH release after a few days in culture is something routinely observed in the investigated liver MPS and is due to the adaptation of the cells in the tissue culture microenvironment. Additionally, the retention of gene expression in the liver MPS was illustrated across an array of 90 different

genes, including several phase I (e.g., P450s) and phase II (e.g., glutathione *S*-transferases) drug-metabolizing enzymes, together with a few important hepatic regulators (e.g., HNF4a) and transporters (e.g., MRP2, BSEP, NTCP. On top of that, additional confidence on the sustained functionality of the system stems from the time-dependent accumulation of drug metabolite levels across several donors/wells, whereas these levels were also highly correlated to the respective intrinsic clearance for the depletion of the parent drug.

In vitro drug-metabolism experiments using hepatocytes that are pooled across different donors (Shibata et al., 2002) have been a common practice to avoid bias arising from interdonor differences. However, the validity of such a practice has not previously been evaluated in MPS technologies. For the vast majority of 90 genes studied, the mRNA expression levels obtained from the pooled hepatocytes were not significantly different from the average mRNA expression levels obtained across the different donors. More importantly, by performing drug depletion studies in the liver MPS in both pooled hepatocytes and individual-donor hepatocytes, it was found that pooled hepatocytes can provide a relatively unbiased estimate of the average metabolic clearance in the donor population. Thus, utilization of pooled hepatocytes to study drug metabolism in the liver MPS is a reliable option as long as the determination of the associated interindividual variability is not of interest.

This study focused intensively on the investigation of interdonor variability and thus included only a small set of compounds ($n = 6$). For this particular set of compounds, using state-of-the-art in vitro–in vivo extrapolation methodologies, we obtained a robust correlation between clinically observed and predicted clearances; however, in absolute values, the predicted clearances were lower than those observed in vivo (average fold-error was 4.2 across all evaluated compounds). This trend of underprediction is similar to that previously observed with other traditionally used in vitro systems (Hallifax et al., 2010), and its origins remain a subject of ongoing research in the drug-metabolism field (Galetin, 2014; Bowman and Benet, 2016). Future studies with a wide and diverse set of compounds are needed to clearly evaluate liver MPS technologies with respect to their clearance prediction capabilities and develop robust empirical relationships that can be used to correct for any underprediction of the in vivo values. Finally, further work is needed on the development of mechanistic model-based methodologies to determine in vitro intrinsic clearance that are particularly focused on liver MPS technologies and their features.

In contrast to the in vitro systems traditionally used to study drug metabolism, liver MPS technologies can be integrated along with MPS of other organ systems, allowing the development of platforms where several organ modules are interacting (Stokes et al., 2015; Yu et al., 2015). The development of such physiome-on-a-chip (or human-on-a-chip) platforms is a novel and exciting research field that holds promise for significant applications in drug development (e.g., screening compounds for efficacy/toxicity) and personalized medicine (e.g., in vitro clinical trials) (Fabre et al., 2014). Since the liver MPS has a central role in these platforms, the current work provides further confidence with respect to their use in pharmacokinetic/pharmacodynamic investigations.

In summary, this is the first study that specifically focuses on the in vitro assessment of interindividual variability in drug metabolism in the context of a microphysiological system. It was clearly illustrated that interdonor differences are substantial and are manifested in multiple levels (intrinsic metabolic clearance, formation of liver-specific molecules, gene expression). Moreover, this work supports the use of modeling and simulation as an indispensable tool to analyze and translate the in vitro results emerging from such microphysiological systems to the in vivo context. Finally, the current work provides further confidence regarding the use of liver MPS technologies as an alternative for drug metabolism–related investigations.

Visual abstract. Schematic overview of the framework proposed in this work with regard to the analysis and the subsequent in vivo translation (at the population level) of the in vitro liver MPS data. See *Discussion* for detailed explanation.

Acknowledgments

The authors thank the members of the MIT PhysioMimetics team for the comments and discussions.

Authorship Contributions

Participated in research design: Tsamandouras, Kostrzewski, Stokes, Hughes, Cirit.

Conducted experiments: Kostrzewski.

Performed data analysis: Tsamandouras, Kostrzewski.

Wrote or contributed to the writing of the manuscript: Tsamandouras, Kostrzewski, Stokes, Griffith, Hughes, Cirit.

References

Benet LZ, Broccatelli F, and Oprea TI (2011) BDDCS applied to over 900 drugs. *AAPS J* **13**:519–547.

Bowman CM and Benet LZ (2016) Hepatic clearance predictions from in vitro-in vivo extrapolation and BDDCS. *Drug Metab Dispos* DOI: 10.1124/dmd.116.071514.

Brown HS, Griffin M, and Houston JB (2007) Evaluation of cryopreserved human hepatocytes as an alternative in vitro system to microsomes for the prediction of metabolic clearance. *Drug Metab Dispos* **35**:293–301.

Dash A, Inman W, Hoffmaster K, Sevidal S, Kelly J, Obach RS, Griffith LG, and Tannenbaum SR (2009) Liver tissue engineering in the evaluation of drug safety. *Expert Opin Drug Metab Toxicol* **5**:1159–1174.

Di L, Keefer C, Scott DO, Strelevitz TJ, Chang G, Bi Y-A, Lai Y, Duckworth J, Fenner K, Troutman MD, et al. (2012) Mechanistic insights from comparing intrinsic clearance values between human liver microsomes and hepatocytes to guide drug design. *Eur J Med Chem* **57**:441–448.

Di L and Obach RS (2015) Addressing the challenges of low clearance in drug research. *AAPS J* **17**:352–357.

Domansky K, Inman W, Serdy J, Dash A, Lim MHM, and Griffith LG (2010) Perfused multiwell plate for 3D liver tissue engineering. *Lab Chip* **10**:51–58.

Ebrahimkhani MR, Neiman JAS, Raredon MSB, Hughes DJ, and Griffith LG (2014) Bioreactor technologies to support liver function in vitro. *Adv Drug Deliv Rev* **69–70**:132–157.

Fabre KM, Livingston C, and Tagle DA (2014) Organs-on-chips (microphysiological systems): tools to expedite efficacy and toxicity testing in human tissue. *Exp Biol Med (Maywood)* **239**:1073–1077.

Galetin A (2014) Rationalizing underprediction of drug clearance from enzyme and transporter kinetic data: from in vitro tools to mechanistic modeling, in *Enzyme Kinetics in Drug Metabolism: Fundamentals and Applications* (Nagar S, Argikar AU, and Tweedie JD eds) pp 255–288, Humana Press, Totowa, NJ.

Hallifax D, Foster JA, and Houston JB (2010) Prediction of human metabolic clearance from in vitro systems: retrospective analysis and prospective view. *Pharm Res* **27**:2150–2161.

Hutzler JM, Ring BJ, and Anderson SR (2015) Low-turnover drug molecules: A current challenge for drug metabolism scientists. *Drug Metab Dispos* **43**: 1917–1928.

Jamei M, Dickinson GL, and Rostami-Hodjegan A (2009) A framework for assessing inter-individual variability in pharmacokinetics using virtual human populations and integrating general knowledge of physical chemistry, biology, anatomy, physiology and genetics: A tale of 'bottom-up' vs 'top-down' recognition of covariates. *Drug Metab Pharmacokinet* **24**:53–75.

Jones H and Rowland-Yeo K (2013) Basic concepts in physiologically based pharmacokinetic modeling in drug discovery and development. *CPT Pharmacometrics Syst Pharmacol* **2**:e63.

Karlsson MO and Savic RM (2007) Diagnosing model diagnostics. *Clin Pharmacol Ther* **82**:17–20.

Long TJ, Cosgrove PA, Dunn RT, Stolz DB, Hamadeh HK, Afshari C, McBride H, and Griffith LG (2016) Modeling therapeutic antibody-small molecule drug-drug interactions using a 3d perfusable human liver co-culture platform. *Drug Metab Dispos* DOI: dmd.116.071456 [published ahead of print].

Mould DR and Upton RN (2013) Basic concepts in population modeling, simulation, and model-based drug development-part 2: introduction to pharmacokinetic modeling methods. *CPT Pharmacometrics Syst Pharmacol* **2**:e38.

Obach RS (1999) Prediction of human clearance of twenty-nine drugs from hepatic microsomal intrinsic clearance data: An examination of in vitro half-life

approach and nonspecific binding to microsomes. *Drug Metab Dispos* **27**: 1350–1359.
Rostami-Hodjegan A (2012) Physiologically based pharmacokinetics joined with in vitro-in vivo extrapolation of ADME: a marriage under the arch of systems pharmacology. *Clin Pharmacol Ther* **92**:50–61.
Sarkar U, Rivera-Burgos D, Large EM, Hughes DJ, Ravindra KC, Dyer RL, Ebrahimkhani MR, Wishnok JS, Griffith LG, and Tannenbaum SR (2015) Metabolite profiling and pharmacokinetic evaluation of hydrocortisone in a perfused three-dimensional human liver bioreactor. *Drug Metab Dispos* **43**: 1091–1099.
Sheiner LB and Beal SL (1981) Evaluation of methods for estimating population pharmacokinetic parameters. II. Biexponential model and experimental pharmacokinetic data. *J Pharmacokinet Biopharm* **9**:635–651.
Sheiner LB and Beal SL (1983) Evaluation of methods for estimating population pharmacokinetic parameters. III. Monoexponential model: routine clinical pharmacokinetic data. *J Pharmacokinet Biopharm* **11**:303–319.
Shibata Y, Takahashi H, Chiba M, and Ishii Y (2002) Prediction of hepatic clearance and availability by cryopreserved human hepatocytes: an application of serum incubation method. *Drug Metab Dispos* **30**:892–896.
Stokes CL, Cirit M, and Lauffenburger DA (2015) Physiome-on-a-chip: The challenge of "scaling" in design, operation, and translation of microphysiological systems. *CPT Pharmacometrics Syst Pharmacol* **4**:559–562.
Trame MN, Biliouris K, Lesko LJ, and Mettetal JT (2016) Systems pharmacology to predict drug safety in drug development. *Eur J Pharm Sci* **94**:93–95.
Tsamandouras N, Dickinson G, Guo Y, Hall S, Rostami-Hodjegan A, Galetin A, and Aarons L (2015a) Development and application of a mechanistic pharmacokinetic model for simvastatin and its active metabolite simvastatin acid using an integrated population PBPK approach. *Pharm Res* **32**:1864–1883.
Tsamandouras N, Rostami-Hodjegan A, and Aarons L (2015b) Combining the 'bottom up' and 'top down' approaches in pharmacokinetic modelling: fitting PBPK models to observed clinical data. *Br J Clin Pharmacol* **79**:48–55.
Tsamandouras N, Wendling T, Rostami-Hodjegan A, Galetin A, and Aarons L (2015c) Incorporation of stochastic variability in mechanistic population pharmacokinetic models: handling the physiological constraints using normal transformations. *J Pharmacokinet Pharmacodyn* **42**:349–373.
Tucker GT and Boas RA (1971) Pharmacokinetic aspects of intravenous regional anesthesia. *Anesthesiology* **34**:538–549.
Vivares A, Salle-Lefort S, Arabeyre-Fabre C, Ngo R, Penarier G, Bremond M, Moliner P, Gallas JF, Fabre G, and Klieber S (2015) Morphological behaviour and metabolic capacity of cryopreserved human primary hepatocytes cultivated in a perfused multiwell device. *Xenobiotica* **45**:29–44.
Wheeler SE, Clark AM, Taylor DP, Young CL, Pillai VC, Stolz DB, Venkataramanan R, Lauffenburger D, Griffith L, and Wells A (2014) Spontaneous dormancy of metastatic breast cancer cells in an all human liver microphysiologic system. *Br J Cancer* **111**: 2342–2350.
Yu J, Cilfone NA, Large EM, Sarkar U, Wishnok JS, Tannenbaum SR, Hughes DJ, Lauffenburger DA, Griffith LG, Stokes CL, et al. (2015) Quantitative systems pharmacology approaches applied to microphysiological systems (mps): Data interpretation and multi-mps integration. *CPT Pharmacometrics Syst Pharmacol* **4**: 585–594.

Address correspondence to: Dr. Murat Cirit, Room 16-469, Building 16, Department of Biological Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139. E-mail: mcirit@mit.edu

Supplementary Material for the manuscript: "Quantitative assessment of population variability in hepatic drug metabolism using a perfused 3D human liver microphysiological system."

N. Tsamandouras[1], T. Kostrzewski[2], C. Stokes[3], L. Griffith[1], D. Hughes[2], M. Cirit[1]

1. Department of Biological Engineering, Massachusetts Institute of Technology, USA.

2. CN Bio Innovations, UK.

3. Stokes Consulting, USA.

Table of Contents

1. Supplementary Methods

1.1. Chemicals and reagents

HPLC grade methanol, acetonitrile, isopropanol, dimethyl sulfoxide (DMSO) and formic acid were purchased from Fisher Scientific (Loughborough, UK). Chloroform and ammonium acetate were purchased from Sigma, (Gillingham, UK). Water for HPLC was purified on a Milli Q system (Millipore, Watford, UK). Rapid Equilibrium Dialysis (RED) devices were purchased from Fisher Scientific (Loughborough, UK). Chemicals used for the drug metabolism studies and standard curves were purchased from the suppliers specified in Table S5. All test compound stock solutions were prepared in DMSO.

1.2. RNA isolation and gene expression analysis

Total RNA was extracted from freshly thawed hepatocytes or from LiverChip scaffolds cultured for 6 days (run in parallel to the drug metabolism study without the addition of any drug), using TRIzol® Reagent (Ambion) and a chloroform phase separation. RNA was precipitated from aqueous phase samples using 100% isopropranolol and RNA pellets were re-suspended in dH_2O. RNA was subsequently treated with TURBO DNA-free™ Kit (Ambion) to remove genomic DNA, before cDNA was synthesized from each sample using the High-Capacity RNA-to-cDNA™ Kit (Applied Biosystems, Foster City, CA). QPCR was performed using SYBR® Green PCR Master Mix and primers designed against transcripts related to hepatic genes of specific interest (CYP1A2, CYP7A1, BSEP, MRP2, NTCP and HNF4a, see also Table S1). Samples were analyzed using a

Quantstudio 6 real time PCR system (Applied Biosystems, UK). Ct values from samples were compared and normalized to GAPDH expression. Additionally, a super-array (see below) analysis was performed in order to investigate the expression of 84 metabolism-related genes. For this super-array analysis, reverse transcription and PCR were performed using RT^2 First Strand Kit and RT^2 Profiler PCR Arrays (Qiagen, UK). Human Drug Metabolism (PAHS-002ZC-12) RT^2 Profiler™ PCR Array was used to analyze each sample. Ct values from samples were compared and normalized to the average expression across five different housekeeping genes (ACTB, B2M, GAPDH, HPRT1 and RPLP0). The super-array and QPCR data (referring to 84 and 6 genes respectively) were merged to generate a dataset of 90 genes the expression of which was investigated. A complete list of all these genes is provided in Table S2.

The analysis of gene expression data was focused on addressing three research questions. The first was to assess the expression of metabolism-related genes in hepatocytes cultured in the liver MPS for 6 days in comparison to that in freshly thawed hepatocytes. For each of the 90 genes for which mRNA levels where available, we calculated the associated average fold-change by averaging (geometric mean) across all the different donors and pool of donors the ratio of relative mRNA expression in the liver MPS compared to freshly thawed hepatocytes. Thus, an average fold-change value higher than 1 for a specific gene indicates up-regulation in the liver MPS compared to the freshly thawed hepatocytes while a value lower than 1 indicates down-regulation. Also, to assess the statistical significance of the observed fold-change in gene expression, a paired t-test was performed for each of the 90 genes, comparing the mRNA levels obtained across the different donors and pool of donors in the liver MPS to the respective levels obtained in

freshly thawed hepatocytes. Multiple testing was taken into account by controling the false discovery rate at the 0.05 level with the Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995).

The second question was to assess the magnitude of inter-donor variability in mRNA expression levels and specifically investigate whether this variability is higher or lower in the liver MPS compared to freshly thawed hepatocytes. Therefore, the coefficient of variation (CV%) associated with the mRNA expression levels across the five hepatocyte donors was calculated for both the liver MPS and freshly thawed hepatocytes.

Finally, the third question was to assess whether the mRNA expression levels obtained from pooled hepatocytes is representative of the average mRNA expression across the different donors. For each of the 90 genes for which mRNA levels were available and for both the liver MPS and the freshly thawed hepatocytes separately, we calculated an average fold-difference metric by averaging across all the different donors, the ratio of mRNA expression in a specific donor compared to the respective expression in the pooled hepatocytes. Thus, an average fold-difference metric value higher than 1 for a specific gene indicates that pooled hepatocytes under-predict the average mRNA expression obtained across the different donors for this gene, while a value lower than 1 indicates over-prediction. In addition, a one sample t-test was performed for each gene, testing the null hypothesis that the mRNA expression levels across the different donors come from a distribution with mean equal to the expression levels determined in the pooled hepatocytes for this specific gene. Multiple testing was taken into account by controling the false discovery rate at the 0.05 level with the Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995).

1.3. Sample preparation and LC-MS/MS analysis

60 µL supernatant samples (i.e. corresponding to the extracellular compartment) were taken at pre-determined sampling times and added to 40 µl water, 150 µl methanol and 150 µl methanol containing 1 µM Tolbutamide as internal standard (IS). Standard curves ranging from 1.5 nM - 12,000 nM were prepared for all the compounds in the same matrixes as the media samples. All analytical samples were vortex mixed and kept at -20°C for a minimum of 2 hours to allow complete protein precipitation. Samples were centrifuged at 2,500 x g, 4°C for 20 minutes and supernatants were transferred to fresh polypropylene microtiter plates (Greiner, Stonehouse, UK), sealed with pre-slit silicone cap mats (Chromatography Direct, Runcorn, UK) and analyzed by liquid chromatography tandem mass spectrometry (LC-MS/MS). The individual multiple reaction monitoring (MRM) methods are summarized in Table S5. The instruments used and the LC-MS/MS parameters are summarized in Table S6. Mass spectrometry was performed by an independent contract research organization (Xenogesis Ltd, Nottingham, UK).

1.4. Rapid equilibrium dialysis to investigate drug binding to tissue culture media

Binding to tissue media components (e.g. BSA) was assessed by rapid equilibrium dialysis (RED). 100 µM compound stock solutions were prepared in DMSO and subsequently diluted 100-fold in WEM maintenance culture media. This solution was transferred to the sample side of a RED plate (Fisher Scientific, Loughborough, UK) and PBS was added to the buffer side. RED plates were sealed and incubated at 37°C for 4 hours with orbital shaking at 100 rpm. Remaining culture medium was mixed 1:1 with

PBS and quenched with methanol/IS. This served as the t=0 sample for recovery determinations. At the end of the incubation, samples from the medium side were mixed 1:1 with PBS and from the PBS side were mixed with 50 µl blank culture media and all samples were quenched with methanol/IS. Compound DMSO stock solutions were diluted in methanol/IS and serially diluted to cover a standard concentration range 0.16 nM - 1000 nM. Each methanolic working solution was spiked into a 1:1 mix of blank media and PBS and quenched with methanol/IS. All standards and samples were mixed, and kept at -20°C for a minimum of 2 hours to allow complete protein precipitation. Samples were centrifuged at 2,500 x g, 4°C for 20 minutes and supernatants were transferred to polypropylene microtiter plates for analysis by LC-MS/MS. The individual MRM methods are summarized in Table S7 and the instruments used and the LC-MS/MS parameters are summarized in Table S6.

1.5. Identification of *in vitro* clearance predictors

It was of particular interest in this study to identify variables that can be used as predictors of *in vitro* clearance. Therefore, the values (empirical Bayes estimates) of intrinsic clearance obtained across different donors/wells ($CL^{ij}_{\mathrm{int}(u)}$) from the modeling of the individual-donor data were investigated in relation to the values of the respective available pre-dose phenotypic metrics (e.g. albumin production). It was desirable to preform such an investigation not independently/separately for each compound but simultaneously across all the studied compounds in order to detect common/universal predictors. Therefore, for each compound, the intrinsic clearance across the different

donors/wells ($CL^{ij}_{\mathrm{int}(u)}$) was normalized by the typical (median) value of intrinsic clearance ($CL_{\mathrm{int}(u)}$) for this specific compound (see Eq.1).

$$normalized\ CL^{ij}_{\mathrm{int}(u)} = \frac{CL^{ij}_{\mathrm{int}(u)}}{CL_{\mathrm{int}(u)}} \qquad \textbf{(Eq.1)}$$

Consequently, the normalized intrinsic clearance will be less than one for donors/wells that we observed a clearance smaller than the typical for this compound clearance and higher than 1 in the opposite case. The normalized intrinsic clearance values across all compounds were subsequently regressed against the following four variables: 1) pre-dose albumin production in the respective donors/wells; 2) pre-dose urea production in the respective donors/wells; 3) pre-dose LDH release levels in the respective donors/wells; and 4) donor-specific CYP mRNA levels determined in the liver MPS wells that were sacrificed at day 6 for RNA analysis (see above). For each compound the mRNA levels of the primarily responsible for its metabolism CYP were used (CYP2D6 for propranolol (Masubuchi et al., 1994); CYP3A4 for predinisolone (Zhang et al., 2009); CYP1A2 for phenacetin (Venkatakrishnan et al., 1998) and lidocaine (Wang et al., 2000); and CYP2C9 for ibuprofen (Davies, 1998) and diclofenac (Bort et al., 1999)). Similarly to the normalization applied in intrinsic clearance values, the mRNA levels of each of the above CYPs were normalized by the average mRNA levels for this specific CYP across all donors. Thus, the normalized CYP mRNA expression (being less than one for donors with mRNA levels lower than the donor average and higher than 1 in the opposite case) was used as predictor in the regression. The Lasso, a regularized least-squares regression method (Tibshirani, 1996), was employed as it was desirable to avoid overfitting / improve prediction accuracy while simultaneously selecting only the subset of predictors

which are important to the dependent variable (*in vitro* intrinsic clearance) by filtering out any less relevant/redundant predictors. The Lasso regression was performed in Matlab R2015b (The MathWorks, Inc., Natick, Massachusetts, USA) using the *lasso* function of the statistics and machine learning toolbox. Optimal tuning of the regularization parameter lambda (λ), and thus model (covariate) selection, was performed by 5-fold cross validation (Hastie et al., 2009) and identification of the λ value that minimizes the cross-validated mean square error.

1.6. Prediction of *in vivo* hepatic clearance

The typical value of unbound intrinsic clearance ($CL_{\mathrm{int}(u)}$) determined for each drug from the pharmacokinetic analysis of the individual-donor *in vitro* data, was subsequently scaled up to a human liver equivalent unbound intrinsic clearance ($CL_{\mathrm{int}(u),H}$) using Eq.2,

$$CL_{\mathrm{int}(u),H} = CL_{\mathrm{int}(u)} \cdot HC \cdot LW \qquad \textbf{(Eq.2)}$$

, where *HC* is the human hepatocellularity of 120 million cells / g of liver (Hakooz et al., 2006) and *LW* is the average human liver weight of 25.7g / kg of body weight (Brown et al., 1997). The hepatic clearance (referring to whole blood concentrations) was then predicted ($CL_{H(pred)}$) using either the Well-Stirred (WS) (Eq.3) or the Parallel Tube (PT) (Eq.4) liver model (Pang and Rowland, 1977).

$$CL_{H(pred)} = \frac{Q_H \cdot fu_b \cdot CL_{\mathrm{int}(u),H}}{Q_H + fu_b \cdot CL_{\mathrm{int}(u),H}} \qquad \textbf{(Eq.3)}$$

$$CL_{H(pred)} = Q_H \cdot \left(1 - e^{-fu_b \cdot CL_{\mathrm{int}(u),H}/Q_H}\right) \qquad \textbf{(Eq.4)}$$

, where Q_H is the average hepatic blood flow of 20.7 mL/min/kg of body weight (Davies and Morris, 1993) and fu_b is the fraction of drug which is unbound in blood (see Table S4). The predicted hepatic clearance ($CL_{H(pred)}$) values were then compared to observed hepatic clearance ($CL_{H(obs)}$) values (referring to whole blood concentrations), which were calculated with Eq.5 under the assumption of no extra-hepatic metabolism.

$$CL_{H(obs)} = CL_{total} \cdot \left(1 - f_{ren}\right) \qquad \textbf{(Eq.5)}$$

, where CL_{total} is the total clearance (referring to whole blood) that has been observed in humans (see values in Table S4) and f_{ren} is the fraction of drug excreted unchanged in urine (see values in Table S4). The overall agreement between the observed and the predicted hepatic clearance values was determined by the calculation (see Eq.6) of the average fold error (AFE) across all the evaluated compounds (N=6).

$$AFE = 10^{\frac{1}{N}\sum_{q=1}^{N}\log\left(\frac{CL_{H(obs),q}}{CL_{H(pred),q}}\right)} \qquad \textbf{(Eq.6)}$$

In order to determine an empirical scaling factor (ESF) that could be used prospectively for PBPK modeling of a new compound, the average fold error was also calculated in the scale of intrinsic clearances (fold-error between the *in vitro* determined intrinsic clearance derived from Eq.2 and the *in vivo* observed intrinsic clearance deriving from the rearrangement of Eqs.3, 4 for a given value of observed hepatic clearance (Hallifax et al., 2010)).

1.7. Lidocaine as a case-example of population *in vitro – in vivo* translation

The development of a population PBPK model for lidocaine is described in this work, in order to clearly illustrate with the aid of a case-example, the framework under which liver MPS data can be translated to predictions of *in vivo* concentration-time profiles at the population level. The main reasons why lidocaine was selected among the other *in vitro* evaluated compounds for such an illustration are reported below. Firstly, high-quality *in vivo* pharmacokinetic data after IV administration are available across different individuals (Tucker and Boas, 1971), offering the ability to validate our predictions. In order to do so, model simulations were performed in accordance to the design of this published study. Secondly, lidocaine was the compound for which we obtained the most precise estimate of inter-donor variability through the model-based analysis of the *in vitro* data (see Table 1, main manuscript). Finally the hepatic extraction ratio of lidocaine (0.63 (Tucker and Mather, 1975)) is such that the observed *in vivo* variability in clearance should depend jointly on both the variability in intrinsic hepatic clearance and the variability in hepatic blood flow. This gives us the chance to illustrate the importance of considering population variability on both drug-related and system-related parameters.

1.8. Mass balance differential equations of the lidocaine PBPK model

The lidocaine PBPK model was mathematically described with a system of 14 mass balance ordinary differential equations (see Eqs.7-11).

Arterial blood concentrations (Eq.7):

$$V_{ar} \cdot \frac{dC_{ar}}{dt} = Q_{lu} \cdot \left(\frac{C_{lu}}{KP_{lu:B}} - C_{art} \right)$$

Lung concentrations (Eq.8):

$$V_{lu} \cdot \frac{dC_{lu}}{dt} = Q_{lu} \cdot \left(C_{ve} - \frac{C_{lu}}{KP_{lu:B}} \right)$$

Venous blood concentrations (Eq.9):

$$V_{ve} \cdot \frac{dC_{ve}}{dt} = \sum_{T \neq sp,gu} Q_T \cdot \frac{C_T}{KP_{T:B}} + INF \cdot IR - Q_{lu} \cdot C_{ve}$$

Liver concentrations (Eq.10):

$$V_{li} \cdot \frac{dC_{li}}{dt} = Q_{sp} \cdot \frac{C_{sp}}{KP_{sp:B}} + Q_{gu} \cdot \frac{C_{gu}}{KP_{gu:B}} + Q_{ha} \cdot C_{ar} - Q_{li} \cdot \frac{C_{li}}{KP_{li:B}} - CL_{\text{int}(u),H} \cdot fu_b \cdot \frac{C_{li}}{KP_{li:B}}$$

Other non-eliminating tissues concentrations (Eq.11):

$$V_T \cdot \frac{dC_T}{dt} = Q_T \cdot \left(C_{ar} - \frac{C_T}{KP_{T:B}} \right) \qquad , T \neq li$$

, where V , Q , C and $KP_{T:B}$ correspond respectively to volumes, blood flows, concentrations and tissue-to-blood partition coefficients associated with the model tissues/compartments. The subscripts *ar*, *lu* and *ve* correspond to arterial blood, lung and venous blood compartments respectively; the subscript *ha* refers to the hepatic artery; while the subscript *T* can refer to any of the following compartments: muscle (*mu*), brain (*br*), kidney (*ki*), heart (*he*), spleen (*sp*), gut (*gu*), liver (*li*), adipose (*ad*), skin (*sk*), bone (*bo*) and rest of the body (*ro*). *IR* corresponds to the lidocaine infusion rate of 3 mg/kg given intravenously over a 3-minute period; *INF* is a dummy variable which takes the

value of 1 when t ≤ 3 mins and 0 when t > 3 mins; and fu_b is the fraction of drug which is unbound in blood.

1.9. System-related parameters and the associated population variability

The generation of system-related parameters (organ volumes, blood flows) that were used as an input in the lidocaine PBPK model simulations is described below. The volumes of the different model compartments were allowed to vary across the simulated population through dependence on the total body weight of each individual (see Eq.12).

$$V_{ji} = f_{WT_{ji}} \cdot WT_i \cdot \left(\frac{1}{d_j} \right) \qquad \textbf{(Eq.12)}$$

, where V_{ji} is the volume that corresponds to tissue/compartment j in the individual i , WT_i is the total body weight of individual i; $f_{WT_{ji}}$ is the fraction of total body weight corresponding to the tissue/compartment j in the individual i (see Table S8 for the average population values); and d_j is the density relating to the tissue/compartment j (used for mass-to-volume conversion, see Table S8). The total body weight of each simulated individual i was generated by random sampling from a normal distribution with mean 69 kg and standard deviation 6.2 (average weight and standard deviation in the published lidocaine study (Tucker and Boas, 1971)). In order to take into account additional stochastic variability in organ volumes that cannot be explained just by total body weight differences, the fractional weight parameters ($f_{WT_{ji}}$) in Eq.12 were also allowed to vary between individuals (for example the fraction of total body weight corresponding to muscle tissue is not exactly 0.4 for all individuals). More specifically

these parameters were randomly sampled from a 14-dimensional (equal to the number of compartments in the model) logistic-normal distribution in order to constrain their sum across all the different model compartments to be always equal to 1 (Tsamandouras et al., 2015). The characteristics of this logistic-normal distribution were tuned accordingly to a previously proposed methodology (Tsamandouras et al., 2015), in order to generate population distributions of fractional weight parameters ($f_{WT_{ji}}$) that have means matching exactly the values known from physiology for each organ (see Table S8) and a coefficient of variation of 10%.

The blood flows associated with the different model compartments were allowed to vary across the simulated population through dependence on the cardiac output of each individual (see Eq.13)

$$Q_{ji} = f_{CO_{ji}} \cdot CO_i \qquad \textbf{(Eq.13)}$$

, where Q_{ji} is the blood flow that corresponds to tissue/compartment *j* in the individual *i* ; CO_i is the cardiac output of individual *i*; and $f_{CO_{ji}}$ is the fraction of cardiac output corresponding to the tissue/compartment *j* in the individual *i* (see Table S8 for the average population values). The cardiac output of each simulated individual *i* was generated from an allometric relationship that uses the sampled weight of each individual as an input (Eq.14) (Nestorov, 2001)

$$CO_i = 187 \cdot WT_i^{0.81} \qquad \textbf{(Eq.14)}$$

, where CO_i is expressed in mL/min and WT_i in kg. In order to take into account additional stochastic variability in blood flows that cannot be explained just by differences in cardiac output, the respective fractional parameters ($f_{CO_{ji}}$) in Eq.13 were also allowed to vary between individuals (for example the fraction of cardiac output that goes to the muscle is not exactly 0.17 for all individuals). These parameters were sampled from an 11-dimensional (equal to the number of compartments in the model (14) minus the two blood compartments and the lung compartment which is perfused by the total cardiac output) logistic normal distribution. Similarly to above, the characteristics of this logistic normal distribution were tuned accordingly to a previously proposed methodology (Tsamandouras et al., 2015), in order to generate population distributions of fractional cardiac output parameters ($f_{CO_{ji}}$) that have means matching exactly the values known from physiology for each organ (see Table S8) and a coefficient of variation of 10%.

1.10. Drug-related parameters and the associated population variability

All the drug-related parameters that were used as an input in the lidocaine PBPK model simulations are summarized in Table S9. In a previous study with 24 individuals (Routledge et al., 1980), the lidocaine fraction unbound in plasma (fu_p) was reported to be on average 0.302 with a standard deviation of 0.055 and thus this information was used to construct a normal distribution from which the fu_p parameter was randomly sampled. The lidocaine blood-to-plasma ratio (BP) has been also previously reported in humans (Tucker and Mather, 1975) to be on average 0.84 with a standard deviation of

0.08 (n=5 subjects) and thus this parameter was also randomly sampled from the respective normal distribution. The fraction unbound in blood parameter (fu_b) was calculated for each individual with Eq.15.

$$fu_b = \frac{fu_p}{BP} \qquad \textbf{(Eq.15)}$$

Tissue-to-plasma unbound partition coefficients ($KPu_{T:P}$) for all tissues (apart form the rest of body compartment) were *in silico* predicted with the mechanistic equations developed by Rodgers and Rowland (Rodgers et al., 2005). These equations require input parameters referring to human tissue composition (e.g. fraction of intracellular / extracellular water, volume of neutral lipids / phospholipids etc.), which were extracted across different organs from (Poulin et al., 2011). Additional drug-related physicochemical parameters are needed for these equations, thus a logP of 2.26 and a pKa of 8.01 were used for lidocaine (Poulin and Theil, 2009). The predicted tissue-to-plasma unbound partition coefficients ($KPu_{T:P}$) were subsequently converted with Eq.16 for each individual to the tissue-to-blood partition coefficients ($KP_{T:B}$), which were used in the model's differential equations (see Eqs.7-11).

$$KP_{T:B} = \frac{KPu_{T:P} \cdot fu_p}{BP} \qquad \textbf{(Eq.16)}$$

Note that population variability on the tissue composition parameters was not considered due to the lack of any relevant information and thus any variability on the calculated tissue-to-blood partition coefficients arise solely from the variability in the fu_p and BP

parameters. Finally, the predicted partition coefficients were averaged for each individual across all the different tissues and this average was assigned as the partition coefficient of the complementary rest of body compartment.

The value of unbound intrinsic clearance ($CL_{\text{int}(u)}$) along with the associated inter-donor variability determined in the liver MPS for lidocaine (see Table 1, main manuscript) was used to construct the respective log-normal distribution from which this parameter was randomly sampled for model simulations. The sampled unbound intrinsic clearance for each individual was subsequently scaled up to a human liver unbound intrinsic clearance ($CL_{\text{int}(u),H}$) using Eq.17.

$$CL_{\text{int}(u),H} = CL_{\text{int}(u)} \cdot HC \cdot f_{WT_{li}} \cdot WT \cdot ESF \qquad \textbf{(Eq.17)}$$

, where HC is the standard human hepatocellularity of 120 million cells / g of liver (Hakooz et al., 2006); $f_{WT_{li}}$ is the fraction of total body weight that corresponds to liver in each individual; WT is the total body weight of each individual (both $f_{WT_{li}}$ and WT are randomly generated for each simulated individual with a procedure described above); and finally ESF is an empirical scaling factor that was determined across all the evaluated in this work compounds and aims to correct for the intrinsic clearance under-prediction of the *in vitro* system (see section 1.6). The value for ESF was 8.7 (see "Prediction of *in vivo* hepatic clearance" in Results, main manuscript) when the well-stirred model (which is more relevant to our PBPK model assumptions compared to the parallel-tube model) was considered. The idea behind the utilization of this empirical scaling factor stems from the fact that we wanted to mimic the situation where it is desirable to perform

predictions for a new chemical entity, when clinical data for this compound are not yet available (as it is usually the case in the early stages of drug development) and the only available information is how the *in vitro* system performs across a range of different compounds. Since the experimental data for only 6 compounds were available in this work, the employed ESF is only for illustration of the approach purposes. A larger and more diverse set of compounds needs to be evaluated in the studied liver MPS, in order to determine a robust system-specific empirical scaling factor.

2. Supplementary Results

2.1. Quantitative analysis of the cell culture phenotypic metrics

Out of the total of 108 wells ((5 donors + 1 pool) x 6 compounds x 3 replicate wells) that were seeded and intended for the drug metabolism study, the vast majority (103) were associated with robust formation of 3D micro-tissues that were consistent/comparable across the different donors and were maintained throughout the culture period (Figure S6). Poor tissue formation was observed for only 5 wells, justified based on microscope pictures in conjunction with an albumin production rate being less than 2 μg / 10^6 cells / day. Results corresponding to these 5 wells were discarded from the subsequent analysis.

Pre-dose (measured at day 4) albumin, urea and LDH levels are illustrated in Figure 1 (main manuscript), stratified across different donors. Substantial and statistically significant differences in albumin production ($p = 4 \cdot 10^{-15}$, one-way ANOVA) were observed across hepatocytes from the 5 different donors (see Figure 1, main manuscript).

Similarly, significant inter-donor differences (see Figure 1, main manuscript) were detected in both urea production ($p = 6 \cdot 10^{-7}$, one-way ANOVA) and LDH release ($p = 5 \cdot 10^{-34}$, one-way ANOVA). These three metrics (albumin, urea and LDH) were correlated at the individual level, as the donors that were associated with high albumin production were also jointly associated with high urea production and low LDH release levels (e.g. donor Hu1601, see Figure 1, main manuscript). On the contrary, donors that were associated with low albumin production were also jointly associated with low urea production and high LDH release levels (e.g. donor Hu8150, see Figure 1, main manuscript). The pre-dose quality control (phenotypic) metrics (albumin, urea and LDH) were also merged across the 5 different donors and compared with the equivalent metrics from the pooled hepatocytes (see Figure 1, main manuscript). Although no difference was observed in albumin production ($p = 0.79$, two sample t-test), the pooled hepatocytes were associated with higher urea production ($p = 0.001$, two sample t-test) and higher LDH levels ($p = 3 \cdot 10^{-4}$, Welch's unequal variances t-test).

A comparison between the pre-dose (measured at day 4) and the post-dose (measured at day 6 for diclofenac, propranol, lidocaine and ibuprofen and day 5 and 7 for phenacetin and prednisolone respectively) quality control (phenotypic) metrics, stratified across different treatments (compounds) is illustrated in Figure 2 (main manuscript). It was evident that across all treatments, post-dose albumin production for a given donor/well was significantly increased ($p < 0.001$, paired sample t-test) compared to the equivalent pre-dose levels (see Figure 2, main manuscript). In addition, across all treatments significantly lower levels of LDH release ($p < 0.001$, paired sample t-test) were detected at post-dose measurements (see Figure 2, main manuscript). With regard to urea

production an overall trend of significant increase at post-dose ($p < 0.05$, paired sample t-test) was observed for all treatments except ibuprofen (no significant change between pre- and post-dose) and prednisolone (significant decrease post- compared to pre-dose, $p<0.001$, paired sample t-test). A one-way ANOVA between the post-dose metrics across all the treatments with post-dose metrics evaluated at the same day (day 6 for diclofenac, ibuprofen, lidocaine and propranolol) indicated no significant inter-treatment differences at the post-dose albumin ($p = 0.53$) urea ($p = 0.24$) and LDH ($p = 0.09$) levels (see Figure 2, main manuscript). This further indicates that the observed differences between pre-dose and post-dose metrics (e.g. albumin) are due to the increased period in culture rather than due to a treatment effect.

2.2. Numerical summary of the drug depletion data available to the pharmacokinetic analysis

From the designed 90 concentration measurements for the metabolic depletion of each compound ((5 donors + 1 pool) x 3 wells x 5 sampling times), the pharmacokinetc analysis included 79, 90, 84, 84, 85 and 85 measurements for propranolol, prednisolone, phenacetin, lidocaine, ibuprofen and diclofenac respectively. The rest of the measurements were discarded either because they were below the limit of quantification (6, 0, 1, 1, 0 and 0 measurements for propranolol, prednisolone, phenacetin, lidocaine, ibuprofen and diclofenac respectively) or because they corresponded to one of the five wells (see section 2.1) with poor tissue formation (5, 0, 5, 5, 5 and 5 measurements for propranolol, prednisolone, phenacetin, lidocaine, ibuprofen and diclofenac respectively).

2.3. Investigation of bias introduced by the assumption of mono-exponential decline for diclofenac

The one-compartment pharmacokinetic model that was assumed by default for drug depletion, provided an adequate description of the data, as the majority of the compounds exhibited mono-exponential declines in their concentration-time profiles. A special note should be made only with regard to diclofenac, as the observed concentration-time profiles indicate that a model of bi-exponential decline might be more appropriate. Indeed a two-compartment pharmacokinetic model provided a slightly better description of the data in terms both of objective function value and model diagnostics (data not shown). However, the instance of such a second (peripheral) compartment is difficult to be explained physiologically, as the results of the non-specific binding to the platform analysis (see main manuscript) indicated no partition of diclofenac to the actual platform. Several other speculations can be made regarding the origin of this biphasic behavior (e.g. nonlinear binding to BSA in the media, distribution to intracellular components, time-dependent change in metabolic activity etc.), however none of these can be robustly supported with the available data in order to be taken into account in our modeling. In order to assure that the assumption of mono-exponential decline used here did not substantially bias the obtained diclofenac intrinsic clearance and the associated inter-donor variability, the intrinsic clearance was also calculated in each donor/well using a standard non-compartmental analysis approach (ratio of the amount of drug in media at time 0 over the calculated 0 to infinity area under the concentration-time curve ($AUC_{0-\infty}$)), as has been previously suggested for such cases (Di and Obach, 2015). Using the latter approach the estimate of the typical unbound intrinsic clearance ($CL_{\mathrm{int}(u)}$) for

diclofenac was 18.04 μL/min/10^6 cells (calculated as the geometric mean across all donors/wells) and the associated inter-donor variability CV was 35.3% (calculated using the averaged intrinsic clearances across different wells for each donor). These values were in very close agreement with the values obtained from the modeling of the data (17.8 μL/min/10^6 cells and 36.2% respectively, see Table 1 in main manuscript), indicating that the mono-exponential decline assumption for diclofenac is not introducing any substantial bias for the purpose of this work.

3. Supplementary References

Benet LZ, Broccatelli F and Oprea TI (2011) BDDCS applied to over 900 drugs. *AAPS J* 13:519-547.

Benjamini Y and Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)* 57:289-300.

Bort R, Macé K, Boobis A, Gómez-Lechón Ma-J, Pfeifer A and Castell J (1999) Hepatic metabolism of diclofenac: Role of human CYP in the minor oxidative pathways. *Biochem Pharmacol* 58:787-796.

Brown RP, Delp MD, Lindstedt SL, Rhomberg LR and Beliles RP (1997) Physiological parameter values for physiologically based pharmacokinetic models. *Toxicol Ind Health* 13:407-484.

Davies B and Morris T (1993) Physiological parameters in laboratory animals and humans. *Pharm Res* 10:1093-1095.

Davies NM (1998) Clinical pharmacokinetics of ibuprofen. *Clin Pharmacokinet* 34:101-154.

Di L and Obach RS (2015) Addressing the challenges of low clearance in drug research. *AAPS J* 17:352-357.

Hakooz N, Ito K, Rawden H, Gill H, Lemmers L, Boobis AR, Edwards RJ, Carlile DJ, Lake BG and Houston JB (2006) Determination of a human hepatic microsomal scaling factor for predicting in vivo drug clearance. *Pharm Res* 23:533-539.

Hallifax D, Foster JA and Houston JB (2010) Prediction of human metabolic clearance from in vitro systems: Retrospective analysis and prospective view. *Pharm Res* 27:2150-2161.

Hastie T, Tibshirani R and Friedman J (2009) *The elements of statistical learning. Data mining, inference and prediction (second edition)*. Springer.

Masubuchi Y, Hosokawa S, Horie T, Suzuki T, Ohmori S, Kitada M and Narimatsu S (1994) Cytochrome P450 isozymes involved in propranolol metabolism in human liver microsomes. The role of CYP2D6 as ring-hydroxylase and CYP1A2 as n-desisopropylase. *Drug Metab Disposition* 22:909-915.

Nestorov I (2001) Modelling and simulation of variability and uncertainty in toxicokinetics and pharmacokinetics. *Toxicol Lett* 120:411-420.

Pang KS and Rowland M (1977) Hepatic clearance of drugs. I. Theoretical considerations of a "well-stirred" model and a "parallel tube" model. Influence of hepatic blood flow, plasma and blood cell binding, and the hepatocellular enzymatic activity on hepatic drug clearance. *J Pharmacokinet Biopharm* 5:625-653.

Poulin P, Jones RDO, Jones HM, Gibson CR, Rowland M, Chien JY, Ring BJ, Adkison KK, Ku MS, He H, Vuppugalla R, Marathe P, Fischer V, Dutta S, Sinha VK, Björnsson T, Lavé T and Yates JWT (2011) PHRMA CPCDC initiative on predictive models of human pharmacokinetics, part 5: Prediction of plasma concentration–time profiles in human by using the physiologically-based pharmacokinetic modeling approach. *J Pharm Sci* 100:4127-4157.

Poulin P and Theil F-P (2009) Development of a novel method for predicting human volume of distribution at steady-state of basic drugs and comparative assessment with existing methods. *J Pharm Sci* 98:4941-4961.

Prescott LF (1980) Kinetics and metabolism of paracetamol and phenacetin. *Br J Clin Pharmacol* 10:291S-298S.

Rodgers T, Leahy D and Rowland M (2005) Physiologically based pharmacokinetic modeling 1: Predicting the tissue distribution of moderate-to-strong bases. *J Pharm Sci* 94:1259-1276.

Routledge PA, Barchowsky A, Bjornsson TD, Kitchell BB and Shand DG (1980) Lidocaine plasma protein binding. *Clin Pharmacol Ther* 27:347-351.

Tibshirani R (1996) Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B (Methodological)* 58:267-288.

Tsamandouras N, Wendling T, Rostami-Hodjegan A, Galetin A and Aarons L (2015) Incorporation of stochastic variability in mechanistic population pharmacokinetic models: Handling the physiological constraints using normal transformations. *J Pharmacokinet Pharmacodyn* 42:349-373.

Tucker GT and Boas RA (1971) Pharmacokinetic aspects of intravenous regional anesthesia. *Anesthesiology* 34:538-549.

Tucker GT and Mather LE (1975) Pharmacology of local anaesthetic agents. Pharmacokinetics of local anaesthetic agents. *Br J Anaesth* 47 suppl:213-224.

Venkatakrishnan K, Moltke LLV and Greenblatt DJ (1998) Human cytochromes P450 mediating phenacetin o‐deethylation in vitro: Validation of the high affinity component as an index of CYP1A2 activity. *J Pharm Sci* 87:1502-1507.

Wang J-S, Backman JT, Taavitsainen P, Neuvonen PJ and Kivistö KT (2000) Involvement of CYP1A2 and CYP3A4 in lidocainen-deethylation and 3-hydroxylation in humans. *Drug Metab Disposition* 28:959-965.

Zhang Z-Y, Chen M, Chen J, Padval MV and Kansra VV (2009) Biotransformation and in vitro assessment of metabolism-associated drug–drug interaction for crx-102, a novel combination drug candidate. *J Pharm Biomed Anal* 50:200-209.

4. Supplementary Tables

Table S1: Primer sequences used to analyze hepatic gene expression by QPCR

Target transcript	Forward primer	Reverse primer
GAPDH	ACAGTTGCCATGTAGACC	TTTTTGGTTGAGCACAGG
ABCB11 / BSEP	CAGATTACAAATGAAGCCCTC	TCCATATCTGTAGGAAGCAG
ABCC2 / MRP2	AAATTGCTGATCTCCTTTGC	GATAGCTGTCCGTACTTTTAC
CYP1A2	CACTATCAGGACTTTGACAAG	AGGTTGACAATCTTCTCCTG
CYP7A1	AAATCTACCCAGACCCTTTG	TTCCAGGACATATTGTAGCTC
SLC10A1 / NTCP	CTTTCTGCTGGGTTATGTTC	CTGGAAAATCATGTAGAGGAG
HNF4α	AGTACATCCCAGCTTTCTG	AATGTAGTCATTGCCTAGGAG

Table S2: Gene expression differences between Liverchip-cultured and freshly isolated hepatocytes

Gene	fold-change [a]	p-value [b]	Gene description
ABCB1	1.54	0.0273	ATP-binding cassette, sub-family B (MDR/TAP), member 1
ABCC1	3.00	0.2114	ATP-binding cassette, sub-family C (CFTR/MRP), member 1
AOC1	0.28	0.0363	Amiloride binding protein 1 (amine oxidase (copper-containing))
ADH1B	0.51	0.1010	Alcohol dehydrogenase 1B (class I), beta polypeptide
ADH1C	0.58	0.1023	Alcohol dehydrogenase 1C (class I), gamma polypeptide
ADH4	0.06	0.0510	Alcohol dehydrogenase 4 (class II), pi polypeptide
ADH5	1.31	0.7147	Alcohol dehydrogenase 5 (class III), chi polypeptide
ADH6	0.25	0.0542	Alcohol dehydrogenase 6 (class V)
AHR	0.45	0.0182	Aryl hydrocarbon receptor
ALAD	0.52	0.0466	Aminolevulinate dehydratase
ALDH1A1	4.36	0.0127	Aldehyde dehydrogenase 1 family, member A1
ALOX12	2.97	0.0220	Arachidonate 12-lipoxygenase
ALOX15	0.82	0.5782	Arachidonate 15-lipoxygenase
ALOX5	1.12	0.8030	Arachidonate 5-lipoxygenase
APOE	6.60	0.5652	Apolipoprotein E
ARNT	1.10	0.5941	Aryl hydrocarbon receptor nuclear translocator
ASNA1	1.75	0.0002	ArsA arsenite transporter, ATP-binding, homolog 1 (bacterial)
BLVRA	14.10	0.0004	Biliverdin reductase A
BLVRB	0.64	0.0937	Biliverdin reductase B (flavin reductase (NADPH))
BSEP	1.31	0.7712	Bile salt export pump
CES1	1.40	0.1227	Carboxylesterase 1
CES2	0.82	0.3689	Carboxylesterase 2
CES3	0.51	0.1392	Carboxylesterase 3
CHST1	0.73	0.9417	Carbohydrate (keratan sulfate Gal-6) sulfotransferase 1
COMT	0.96	0.6338	Catechol-O-methyltransferase
CYB5R3	1.44	0.0572	Cytochrome b5 reductase 3
CYP11B2	0.30	0.1706	Cytochrome P450, family 11, subfamily B, polypeptide 2
CYP17A1	0.29	0.2176	Cytochrome P450, family 17, subfamily A, polypeptide 1
CYP19A1	1.12	0.5368	Cytochrome P450, family 19, subfamily A, polypeptide 1
CYP1A1	1.78	0.4547	Cytochrome P450, family 1, subfamily A, polypeptide 1
CYP1A2	0.37	0.0971	Cytochrome P450, family 1, subfamily A, polypeptide 2
CYP2B6	0.09	0.0232	Cytochrome P450, family 2, subfamily B, polypeptide 6
CYP2C19	0.37	0.1069	Cytochrome P450, family 2, subfamily C, polypeptide 19
CYP2C8	0.13	0.0188	Cytochrome P450, family 2, subfamily C, polypeptide 8
CYP2C9	0.26	0.0295	Cytochrome P450, family 2, subfamily C, polypeptide 9
CYP2D6	0.74	0.3833	Cytochrome P450, family 2, subfamily D, polypeptide 6
CYP2E1	0.01	0.0001	Cytochrome P450, family 2, subfamily E, polypeptide 1
CYP2F1	0.38	0.0337	Cytochrome P450, family 2, subfamily F, polypeptide 1
CYP2J2	0.41	0.0107	Cytochrome P450, family 2, subfamily J, polypeptide 2
CYP3A4	1.38	0.7600	Cytochrome P450, family 3, subfamily A, polypeptide 4
CYP3A5	0.49	0.0512	Cytochrome P450, family 3, subfamily A, polypeptide 5
CYP7A1	9.94	0.3922	Cytochrome P450, family 7, subfamily A, polypeptide 1
EPHX1	1.80	0.0443	Epoxide hydrolase 1, microsomal (xenobiotic)
FAAH	0.78	0.3217	Fatty acid amide hydrolase
FBP1	0.23	0.0370	Fructose-1,6-bisphosphatase 1
GAD1	3.19	0.0488	Glutamate decarboxylase 1 (brain, 67kDa)
GAD2	0.36	0.0734	Glutamate decarboxylase 2 (pancreatic islets and brain, 65kDa)
GCKR	0.27	0.0111	Glucokinase (hexokinase 4) regulator

GPI	1.95	0.0155	Glucose-6-phosphate isomerase
GPX1	2.42	0.0001	Glutathione peroxidase 1
GPX2	0.23	0.0345	Glutathione peroxidase 2 (gastrointestinal)
GPX3	0.94	0.4559	Glutathione peroxidase 3 (plasma)
GPX4	2.01	0.0357	Glutathione peroxidase 4 (phospholipid hydroperoxidase)
GPX5	0.38	0.0337	Glutathione peroxidase 5 (epididymal androgen-related protein)
GSR	3.18	0.0006	Glutathione reductase
GSTA3	0.25	0.0951	Glutathione S-transferase alpha 3
GSTA4	1.06	0.3173	Glutathione S-transferase alpha 4
GSTM2	0.61	0.2606	Glutathione S-transferase mu 2 (muscle)
GSTM3	3.14	0.0214	Glutathione S-transferase mu 3 (brain)
GSTM5	0.40	0.0419	Glutathione S-transferase mu 5
GSTP1	2.26	0.1520	Glutathione S-transferase pi 1
GSTT1	0.75	0.2724	Glutathione S-transferase theta 1
GSTZ1	0.36	0.0144	Glutathione transferase zeta 1
HK2	0.67	0.2588	Hexokinase 2
HNF4a	2.25	0.0069	Hepatocyte nuclear factor 4, alpha
HSD17B1	2.21	0.0631	Hydroxysteroid (17-beta) dehydrogenase 1
HSD17B2	0.44	0.0046	Hydroxysteroid (17-beta) dehydrogenase 2
HSD17B3	1.15	0.7343	Hydroxysteroid (17-beta) dehydrogenase 3
LPO	0.39	0.0338	Lactoperoxidase
MGST1	2.19	0.0502	Microsomal glutathione S-transferase 1
MGST2	0.36	0.0152	Microsomal glutathione S-transferase 2
MGST3	1.00	0.8597	Microsomal glutathione S-transferase 3
MPO	0.75	0.4944	Myeloperoxidase
MRP2	1.56	0.2334	Multidrug resistance-associated protein 2
MT2A	0.03	0.0054	Metallothionein 2A
MT3	0.38	0.0337	Metallothionein 3
MTHFR	0.72	0.3567	Methylenetetrahydrofolate reductase (NAD(P)H)
NAT1	0.51	0.0809	N-acetyltransferase 1 (arylamine N-acetyltransferase)
NAT2	0.80	0.1619	N-acetyltransferase 2 (arylamine N-acetyltransferase)
NOS3	0.58	0.4472	Nitric oxide synthase 3 (endothelial cell)
NQO1	25.24	0.0606	NAD(P)H dehydrogenase, quinone 1
NTCP	0.75	0.2870	Na(+)/Taurocholate transport protein
PKLR	2.42	0.1997	Pyruvate kinase, liver and RBC
PKM	19.41	0.0009	Pyruvate kinase, muscle
PON1	0.44	0.0528	Paraoxonase 1
PON2	0.91	0.4159	Paraoxonase 2
PON3	0.15	0.0008	Paraoxonase 3
SNN	4.23	0.0008	Stannin
SRD5A1	1.03	0.7832	Steroid-5-alpha-reductase, alpha polypeptide 1
SRD5A2	0.68	0.1688	Steroid-5-alpha-reductase, alpha polypeptide 2

[(a)] Average fold-change for each gene that has been calculated by averaging (geometric mean) across all the different donors and pool of donors the fold difference (ratio) of relative mRNA expression in the liver MPS (day 6) compared to freshly isolated hepatocytes. Therefore values higher than 1 indicate up-regulation in the liver MPS compared to the freshly isolated hepatocytes while values lower than 1 indicate down-regulation.

[(b)] p-values correspond to a paired t-test regarding the the null hypothesis of non-significant differences between the relative mRNA expression in the liver MPS compared to freshly isolated hepatocytes across all the different donors and pool of donors. The genes for which we detected significant differences after multiple testing correction (control false discovery rate at the 0.05 level with the Benjamini-Hochberg procedure) are highlighted in color (red for significant up-regulation in the liver MPS and blue for significant down-regulation in the liver MPS).

Table S3: Intercept and coefficients of the developed Lasso regression model

Term	Value
Intercept	1.0143
Albumin production (pre-dose)	0.0510
Urea production (pre-dose)	0.0102
LDH release (pre-dose)	-0.4254
Normalized CYP mRNA (day 6)	0.2514

The regression coefficients associated with albumin production, urea production and CYP mRNA levels indicate a positive relationship with intrinsic clearance, while the coefficient associated with LDH levels indicates a negative relationship of this variable with intrinsic clearance.

Table S4: Parameters related to the comparison of *in vitro* clearance to *in vivo* data

Compound	fu_b [(a)]	f_{ren} (%) [(b)]	CL_{total} [(c)]	$CL_{H(obs)}$ [(d)]	$CL_{H(pred,WS)}$ [(e)]	$CL_{H(pred,PT)}$ [(f)]
Propranolol	0.14	0.25	13.2	13.17	1.55	1.61
Prednisolone	0.10	16.0	2.4	2.02	0.25	0.25
Phenacetin	0.60	0.5	19.6	19.50	9.18	11.37
Lidocaine	0.33	8.0	11.5	10.58	3.67	4.01
Ibuprofen	0.015	0.5	0.81	0.806	0.23	0.23
Diclofenac	0.014	0.5	4.0	3.98	0.74	0.75

[(a)] Fraction of unbound drug in blood. Values obtained from (Hallifax et al., 2010).

[(b)] Fraction of drug excreted unchanged in urine. Values for all drugs apart from phenacetin were obtained from (Benet et al., 2011). Value for phenacetin was obtained from (Prescott, 1980).

[(c)] Total clearance (referring to whole blood) that has been observed in humans (expressed in mL/min/kg of body weight). Values for all drugs were obtained from (Hallifax et al., 2010).

[(d)] Hepatic clearance (referring to whole blood) observed in humans (expressed in mL/min/kg of body weight).

[(e)] Predicted hepatic clearance (referring to whole blood) in humans (expressed in mL/min/kg of body weight) using the *in vitro* data along with the well-stirred model.

[(e)] Predicted hepatic clearance (referring to whole blood) in humans (expressed in mL/min/kg of body weight) using the *in vitro* data along with the parallel tube model.

Table S5: Mass spectrometry parameters and suppliers for hepatocyte metabolism samples

Compound	Parent (m/z)	Daughter (m/z)	Cone (V)	Collision energy (eV)	Supplier	LC-MS/MS method
Ibuprofen	205.1	161.0	20	5	Sigma	A
Acetaminophen*	152.1	93.0	10	22	Sigma	B
Acetaminophen*	152.1	110.1	10	16	Sigma	B
Phenacetin*	180.1	93.0	50	22	Sigma	B
Phenacetin*	180.1	110.1	50	22	Sigma	B
2-Hydroxyibuprofen*	240.2	107.1	10	28	Santa Cruz Biotec.	B
2-Hydroxyibuprofen*	240.2	205.2	10	10	Santa Cruz Biotec.	B
Propranolol*	260.2	56.0	16	28	Sigma	B
Propranolol*	260.2	116.1	16	16	Sigma	B
Lidocaine*	235.2	58.1	22	34	Sigma	B
Lidocaine*	235.2	86.1	22	16	Sigma	B
Diclofenac	296.3	215.1	16	16	Sigma	B
4-Hydroxydiclofenac	312.0	230.1	34	28	Sigma	B
Prednisolone*	361.2	147.1	16	22	Sigma	B
Prednisolone*	361.2	343.2	16	10	Sigma	B
6β-Hydroxyprednisolone**	341.0	237.0	10	15	Santa Cruz Biotec.	B
6β-Hydroxyprednisolone**	341.1	323.0	10	10	Santa Cruz Biotec.	B
6β-Hydroxyprednisolone**	377.1	323.0	10	10	Santa Cruz Biotec.	B
6β-Hydroxyprednisolone**	377.1	341.0	10	5	Santa Cruz Biotec.	B
Acetaminophen	152.0	110.0	65	16	Sigma	C
Phenacetin	180.0	110.0	35	10	Sigma	C
2-Hydroxyibuprofen	240.1	204.8	35	10	SCBT	C
Propranolol	260.0	116.0	35	16	Sigma	C
Diclofenac	295.9	214.8	35	16	Sigma	C
4-Hydroxydiclofenac	312.0	230.9	41	16	Sigma	C
Prednisolone	361.2	343.1	53	10	Sigma	C
6β-Hydroxyprednisolone	377.1	341.1	47	10	Santa Cruz Biotec.	C

* Traces for these transitions were combined in order to maximize sensitivity.

** Parent *m/z* was reduced probably due to water loss in the mass spectrometer source, all 4 traces were combined to maximize sensitivity.

Table S6. Analytical instruments and methods used for hepatocyte metabolism samples and for tissue culture medium binding samples

	Method A		**Method B**		**Method C**		**Method D**		**Method E**	
LC system	Waters Acquity UPLC I-class		Waters Acquity UPLC I-class		Waters Acquity UPLC		Waters Acquity UPLC		Waters Acquity UPLC I-class	
Solvent A	Water + 0.1% formic acid		Water + 0.1% formic acid + 0.025% ammonium acetate		Water + 0.1% formic acid + 0.025% ammonium acetate		Water + 0.1% formic acid + 0.025% ammonium acetate		Water + 0.1% formic acid	
Solvent B	Methanol + 0.1% formic acid		Methanol + 0.1% formic acid + 0.025% ammonium acetate		Methanol + 0.1% formic acid + 0.025% ammonium acetate		Methanol + 0.1% Formic + 0.025% ammonium acetate		Methanol + 0.1% formic acid	
Column	Acquity UPLC CSH C18, 50 x 2.1 mm, 1.7 μm		ACE Excel C18-AR, 50 x 2.1 mm, 2 μm		ACE Excel C18-AR, 50 x 2.1mm, 2 μm		ACQUITY UPLC BEH C18, 50 x 2.1 mm, 1.7 μm		Acquity UPLC CSH C18, 50 x 2.1mm, 1.7μm	
Column temperature:	65°C		65°C		60°C		70°C		70°C	
Flow rate:	0.8 ml/min		0.8 ml/min		0.8 ml/min		0.8 ml/min		0.8 ml/min	
Gradient conditions:	**Time**	**% Solvent A**	**Time**	**% Solvent A**	**Time**	**% Solvent A**	**Time**	**% Solvent A**	**Time**	**% Solvent A**
	0.00	70	0.00	99	0.00	99	0.00	95	0.00	70
	0.40	1	0.50	1	0.50	1	2.00	95	2.00	70
	0.75	1	0.75	1	0.75	1	3.00	5	2.40	1
	0.80	70	0.80	99	0.80	99	3.05	1	3.80	1
	0.85	70	0.85	99	1.50	99	4.00	1	3.90	70
							4.05	95	4.00	70
							4.10	95		
Mass spectrometer:	Waters TQ-S Micro		Waters TQ-S Micro		Waters Quattro Platinum		Waters Premier XE		Waters TQ-S Micro	
Ionisation mode	ES-		ES+		ES+		ES+		ES-	
Capillary (kV)	4		4		4		0.6		3.5	
Source Temperature (°C)	150		150		120		140		150	
Desolvation Temperature (°C)	600		600		400		500		600	
Desolvation gas flow (L/h)	1000		1000		900		900		1000	
Cone gas flow (L/h)	20		20		20		20		20	

Table S7: Mass spectrometry parameters and suppliers for tissue culture medium binding samples

Compound	Parent (m/z)	Daughter (m/z)	Cone (V)	Collision energy (eV)	Supplier	LC-MS/MS method
Prednisolone	361.0	147.0	22	10	Sigma	D
Phenacetin	180.1	109.9	28	22	Sigma	D
Lidocaine	235.0	85.8	22	16	Sigma	D
Diclofenac*	295.9	214.0	16	28	Sigma	D
Diclofenac*	295.9	214.9	16	22	Sigma	D
Diclofenac*	295.9	250.0	16	10	Sigma	D
Propranolol*	260.0	116.0	22	16	Sigma	D
Propranolol*	260.0	183.0	22	16	Sigma	D
Ibuprofen	205.1	161.0	20	5	Sigma	E

* Traces for these transitions were combined in order to maximize sensitivity.

Table S8: System-related parameters related to the lidocaine PBPK model

Compartment	f_{WT} [(a)]	f_{CO} [(b)]	d_j [(c)]
Arterial blood	0.0257	1	1.05
Lungs	0.0076	1	1.05
Venous blood	0.0514	1	1.05
Spleen	0.0026	0.03	1.05
Gut	0.0171	0.16	1.04
Liver	0.0257	0.065	1.04
Brain	0.0200	0.12	1.04
Heart	0.0047	0.04	1.03
Kidney	0.0044	0.19	1.05
Skin	0.0371	0.05	1.12
Muscle	0.4000	0.17	1.04
Adipose	0.2142	0.05	0.92
Bone	0.1429	0.05	1.58
Rest of Body	0.0466	0.075	1.04

[(a)] Fraction of total body weight corresponding to each tissue/compartment. All values have been obtained from (Brown et al., 1997), apart from those corresponding to arterial and venous blood which have been obtained form (Jones et al., 2006) and the rest of body compartment which derives through the constraint that fractions across all tissue/compartments should add up to 1. Value reported for the gut compartment refers to gastrointestinal tract. Value reported for the bone compartment refers to both bone and marrow. A coefficient of variation of 10% has been assigned around the reported values through a logistic-normal distribution (see section 1.9).

[(b)] Fraction of cardiac output corresponding to each tissue/compartment. All values have been obtained from (ICRP, 2002), apart from the one corresponding to the rest of body compartment which derives through the constraint that fractions across all tissues (except of lung that receives the total cardiac output) should add up to 1. Value reported for the liver compartment refers only to hepatic artery. Value reported for the gut compartment has been calculated as the total hepatic blood flow (25.5% of cardiac output) minus the hepatic artery blood flow (6.5% of cardiac output) minus the spleen blood flow (3% of cardiac output). A coefficient of variation of 10% has been assigned around the reported values through a logistic-normal distribution (see section 1.9).

[(c)] Organ density used for mass-to-volume convertions. Values for lungs, spleen, gut, brain, heart, kidney, skin, muscle, adipose and bone have been obtained from (Brown et al., 1997). Specifically, gut density has been calculated as the weighted average of the reported values for stomach, small intestine and large intestine. Skin density refers to dermis. Bone density has been calculated as the weighted average of the values reported for cortical bone, trabecular bone, red marrow and wellow marrow. Blood density has been obtained from (Trudnowski and Rico, 1974). A value of 1.04 was assumed when density has not been reported (liver, rest of body) based on (Brown et al., 1997).

Table S9: Drug-related parameters related to the lidocaine PBPK model

Parameter	Value
fu_p	0.302
BP	0.84
$KP_{lu:B}$	5.51
$KP_{sp:B}$	4.91
$KP_{gu:B}$	4.24
$KP_{li:B}$	6.84
$KP_{br:B}$	1.96
$KP_{he:B}$	3.56
$KP_{ki:B}$	7.19
$KP_{sk:B}$	2.49
$KP_{mu:B}$	3.92
$KP_{ad:B}$	1.48
$KP_{bo:B}$	2.16
$KP_{ro:B}$	4.02
$CL_{\mathrm{int}(u)}$	4.38
ESF	8.70

Values above correspond to typical / average values. When population variability is incorporated in any of the above parameters in the model simulations, this is explicitly described in the text (see section 1.10). fu_p corresponds to fraction unbound in plasma and BP to the blood-to-plasma ratio; $KP_{T:B}$ corresponds to the tissue-to-blood partition coefficient and the indices lu, sp, gu, li, br, he, ki, sk, mu, ad, bo, ro, correspond to the lung, spleen, gut, liver, brain, heart, kidney, skin, muscle, adipose, bone and rest of body tissue/compartments respectively; $CL_{\mathrm{int}(u)}$ corresponds to the *in vitro* liver MPS-determined unbound intrinsic clearance and is reported in μL/min/10^6 cells; ESF corresponds to an empirical scaling factor determined across all the compounds evaluated in this work (see section 1.10).

5. Supplementary Figures

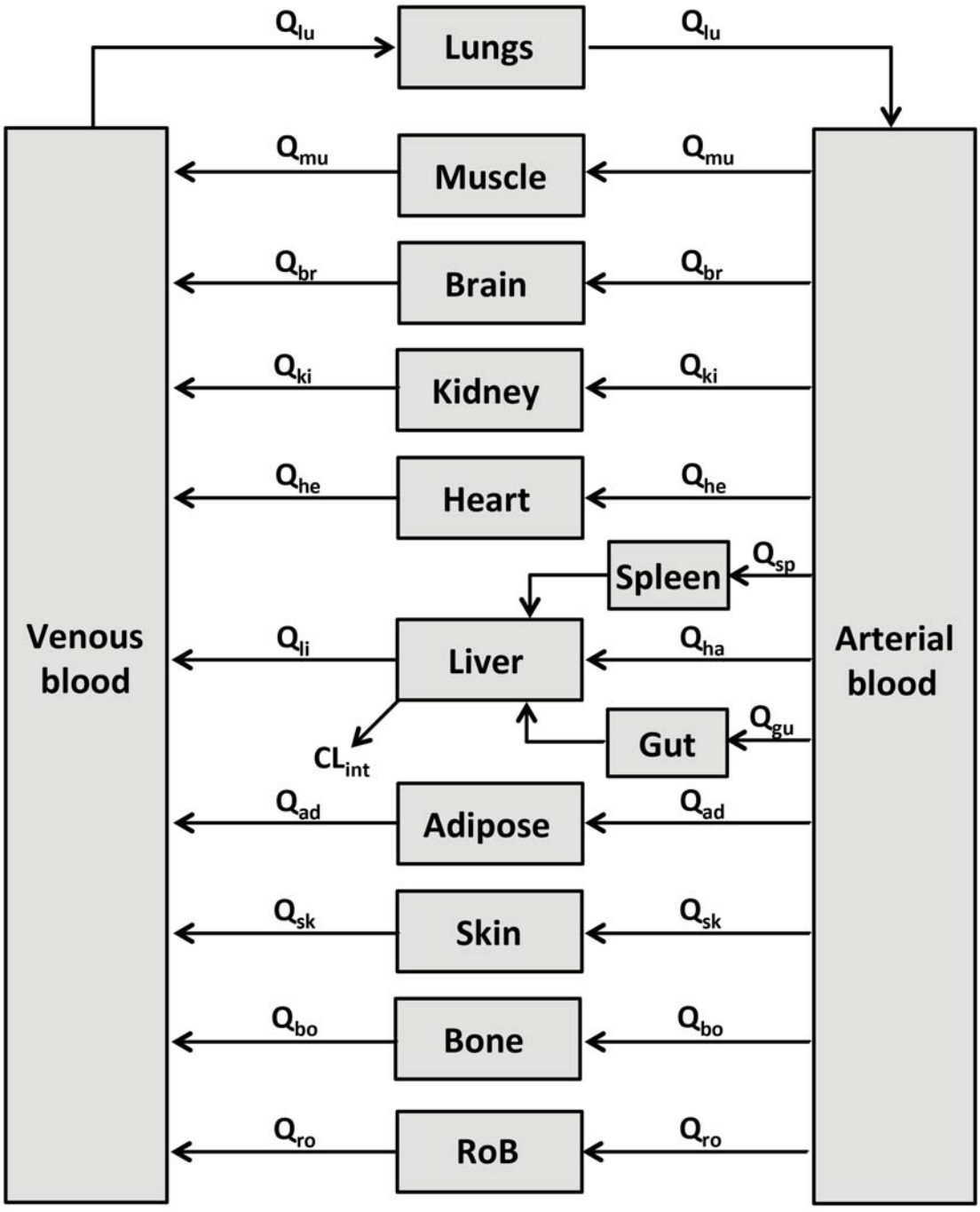

Figure S1. Schematic structure of the employed lidocaine PBPK model. Different organ/tissues are represented as different compartments connected by perfusing blood. RoB represents the rest of body compartment. CL_{int} represents the hepatic metabolic intrinsic clearance. Q_i refers to the blood flow perfusing the different organ/tissues and subscripts "i" are defined as: lu (lungs), mu (muscle), br (brain), ki (kidney), he (heart), sp (spleen), ha (hepatic artery), gu (gut), li (liver), ad (adipose), sk (skin), bo (none), ro (rest of body). The blood flow exiting the liver compartment is the sum of the blood flows assigned to the hepatic artery, gut and spleen ($Q_{li} = Q_{ha} + Q_{gu} + Q_{sp}$).

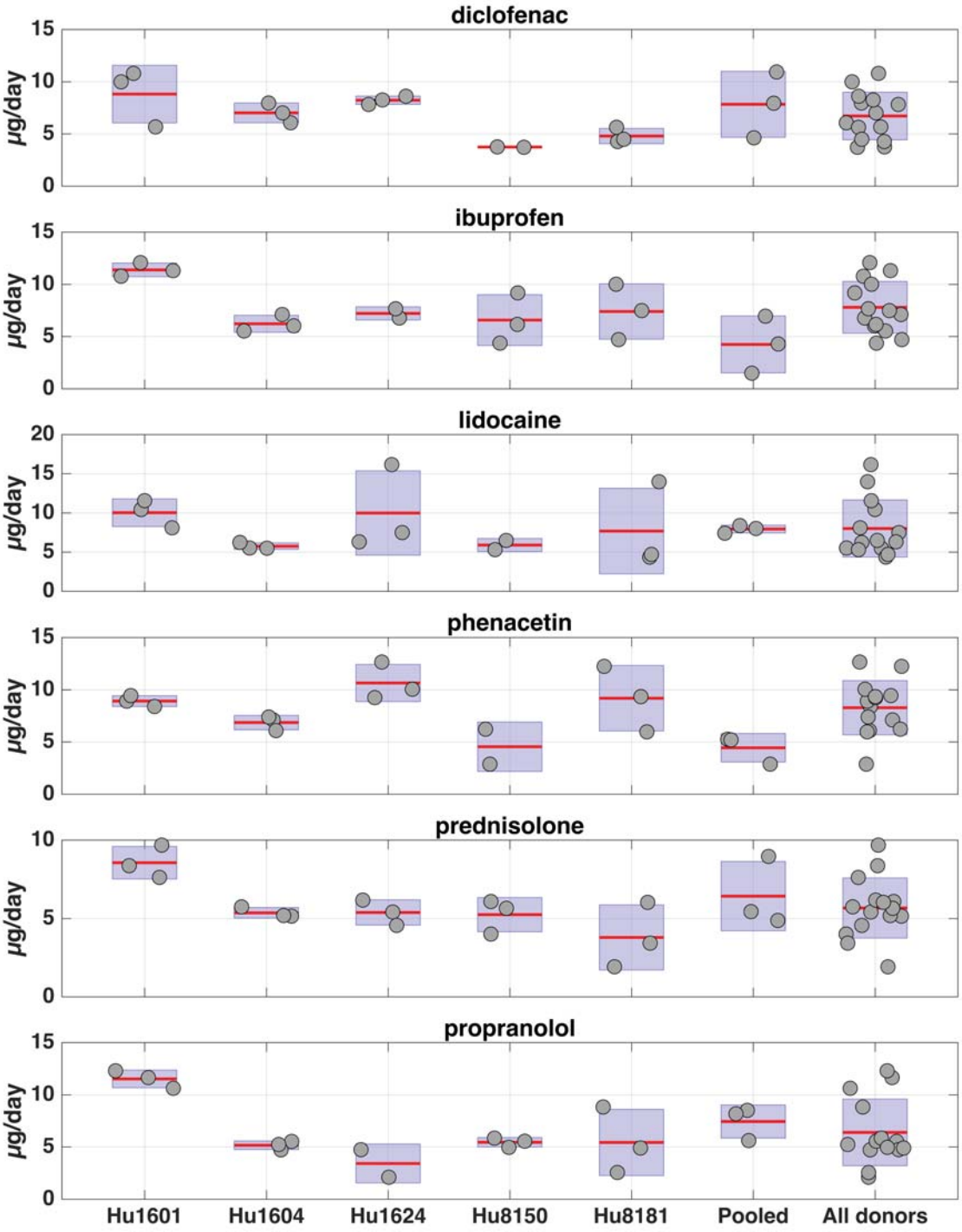

Figure S2. Post-dose albumin production (measured at day 6 for diclofenac, propranol, lidocaine and ibuprofen and day 5 and 7 for phenacetin and prednisolone respectively) stratified across different treatments and donors. "Hu1601", "Hu1604", "Hu1624", "Hu8150" and "Hu8181" are lot numbers corresponding to 5 different donors. "Pooled" refers to the pool of hepatocytes from the 5 donors and "All donors" refer to the data from all 5 donors merged together. Red lines correspond to the mean of the data and the purple boxes extend the mean by ± 1 SD (standard deviation).

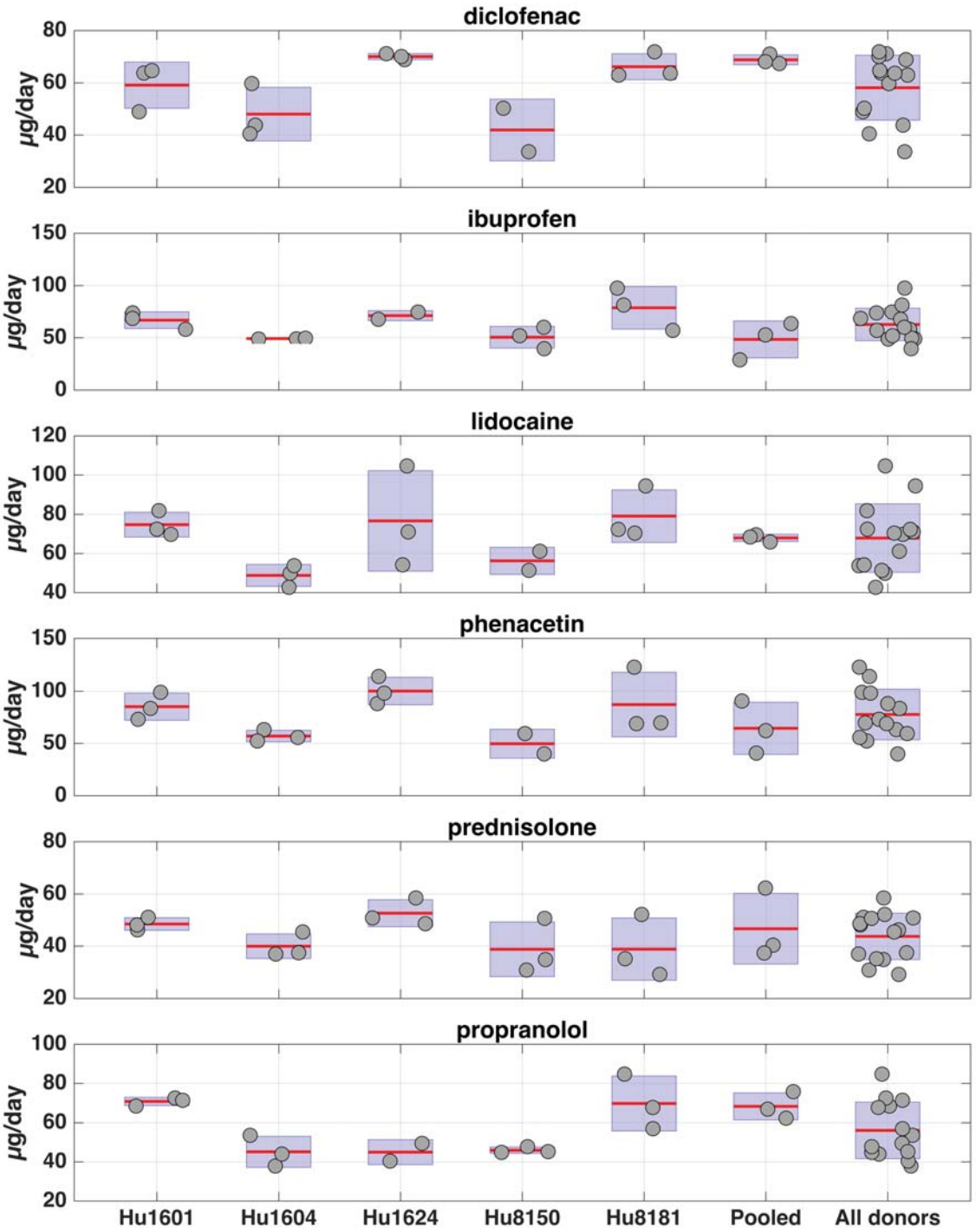

Figure S3. Post-dose urea production (measured at day 6 for diclofenac, propranol, lidocaine and ibuprofen and day 5 and 7 for phenacetin and prednisolone respectively) stratified across different treatments and donors. "Hu1601", "Hu1604", "Hu1624", "Hu8150" and "Hu8181" are lot numbers corresponding to 5 different donors. "Pooled" refers to the pool of hepatocytes from the 5 donors and "All donors" refer to the data from all 5 donors merged together. Red lines correspond to the mean of the data and the purple boxes extend the mean by ± 1 SD (standard deviation).

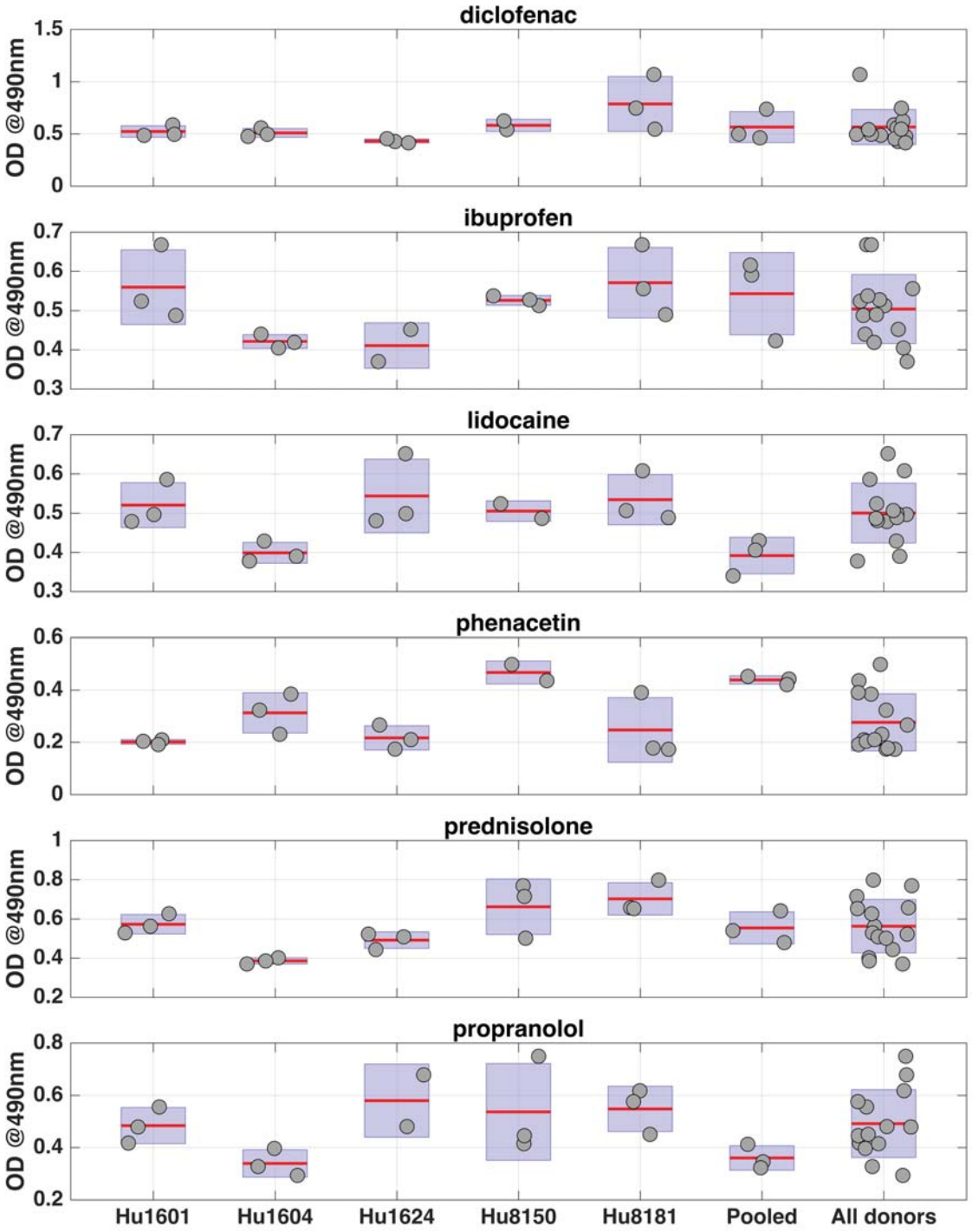

Figure S4. Post-dose LDH release (measured at day 6 for diclofenac, propranol, lidocaine and ibuprofen and day 5 and 7 for phenacetin and prednisolone respectively) stratified across different treatments and donors. "Hu1601", "Hu1604", "Hu1624", "Hu8150" and "Hu8181" are lot numbers corresponding to 5 different donors. "Pooled" refers to the pool of hepatocytes from the 5 donors and "All donors" refer to the data from all 5 donors merged together. Red lines correspond to the mean of the data and the purple boxes extend the mean by ± 1 SD (standard deviation). LDH levels are expressed in optical density (OD) units at 490nm.

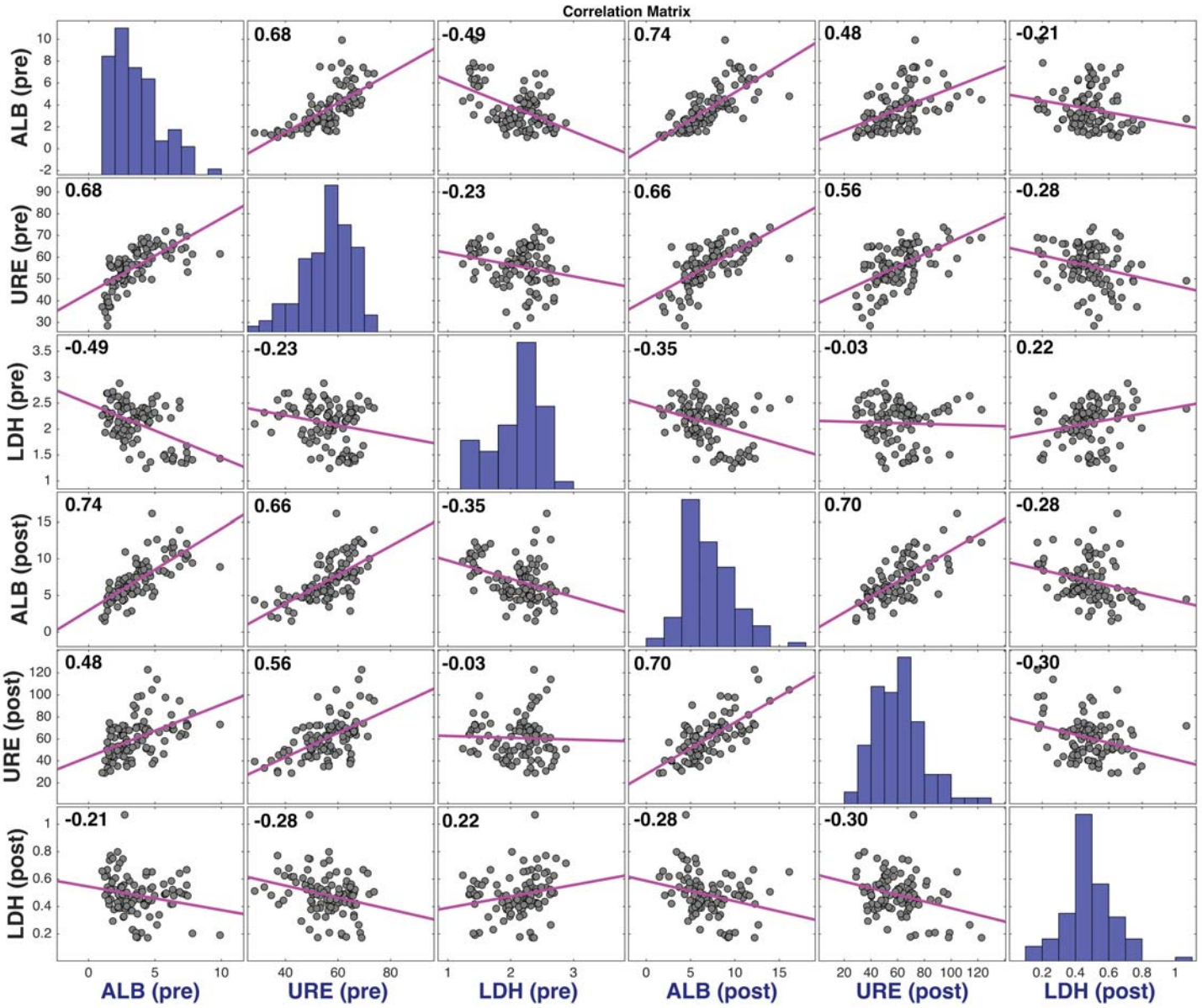

Figure S5. Correlation matrix plot of all the pre- and post-dose albumin (ALB), urea (URE) and LDH levels measured across different donors (or pool of donors) / wells. Each panel illustrates the pairwise correlation between the variables in the x and y axis, apart from the panels in the diagonal of the matrix which illustrate a histogram of the x-axis variable. The purple line inside each panel corresponds to a linear regression fit. The number in the top left corner of each panel refers to the pairwise Pearson's linear correlation coefficient (between -1 and 0 for negative correlations and between 0 and 1 for positive correlations). Indicatively, pre-dose and post-dose albumin productions exhibit a Pearson's correlation coefficient of 0.74, and post-dose albumin and post-dose urea production exhibit a coefficient of 0.70. Albumin and urea production is reported in units of μg/day and LDH release levels in units of optical density at 490nm.

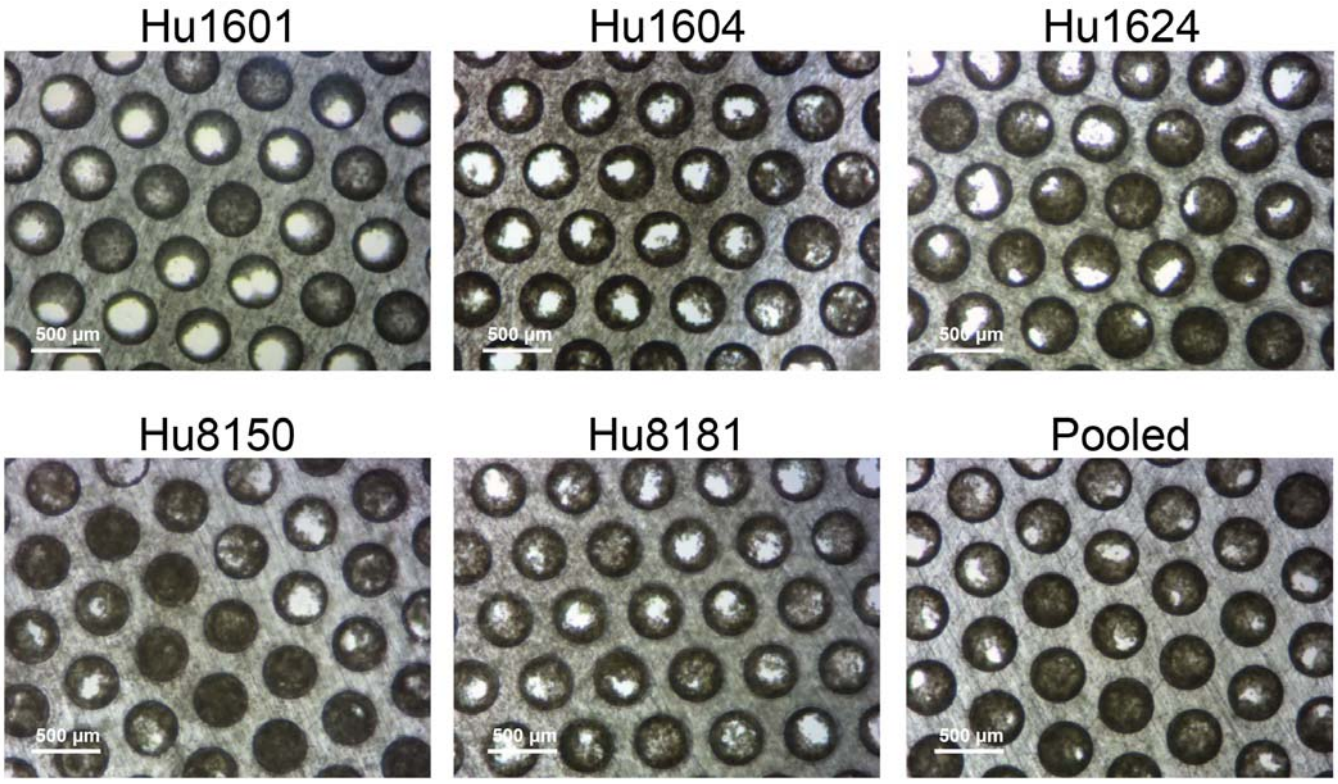

Figure S6. Micro-tissue formation across different donors. Scaffolds from the liver MPS were imaged after culture period (at the end of the drug metabolism study) by bright-field microscopy. One representative well is presented for each donor (or pool of donors). "Hu1601", "Hu1604", "Hu1624", "Hu8150" and "Hu8181" are lot numbers corresponding to 5 different donors. "Pooled" refers to the pool of hepatocytes from the 5 donors.

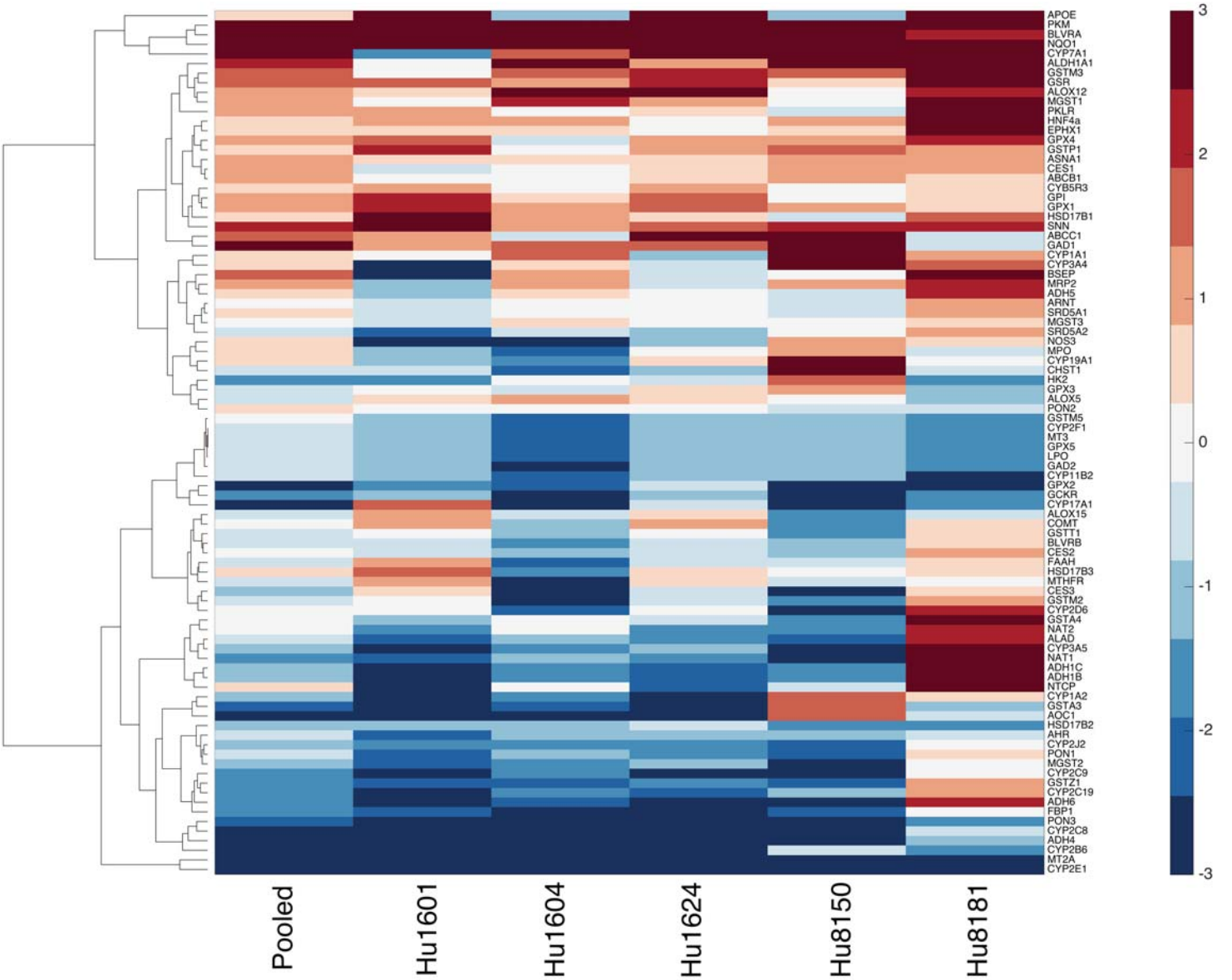

Figure S7. Heatmap and the associated dendrogram resulting from agglomerative hierarchical clustering with respect to the fold-change in gene expression between the liver MPS (day 6) and freshly thawed hepatocytes. The $\log_2$ of fold-change is being considered, thus positive values and red color (see colorbar) indicate up-regulation in the liver MPS compared to the freshly thawed hepatocytes and negative values and blue color (see colorbar) indicate down-regulation. The intensity of the color corresponds to the degree of up- or down-regulation (see colorbar) with darkest red corresponding to at least 8-fold up-regulation in the liver MPS and darkest blue to at least 8-fold down-regulation. "Hu1601", "Hu1604", "Hu1624", "Hu8150" and "Hu8181" are lot numbers corresponding to the 5 different donors. "Pooled" refers to the pool of hepatocytes from the 5 donors. Agglomerative hierarchical clustering was performed with the *clustergram* function in Matlab and more specifically using Ward's method with a Euclidean distance measure.

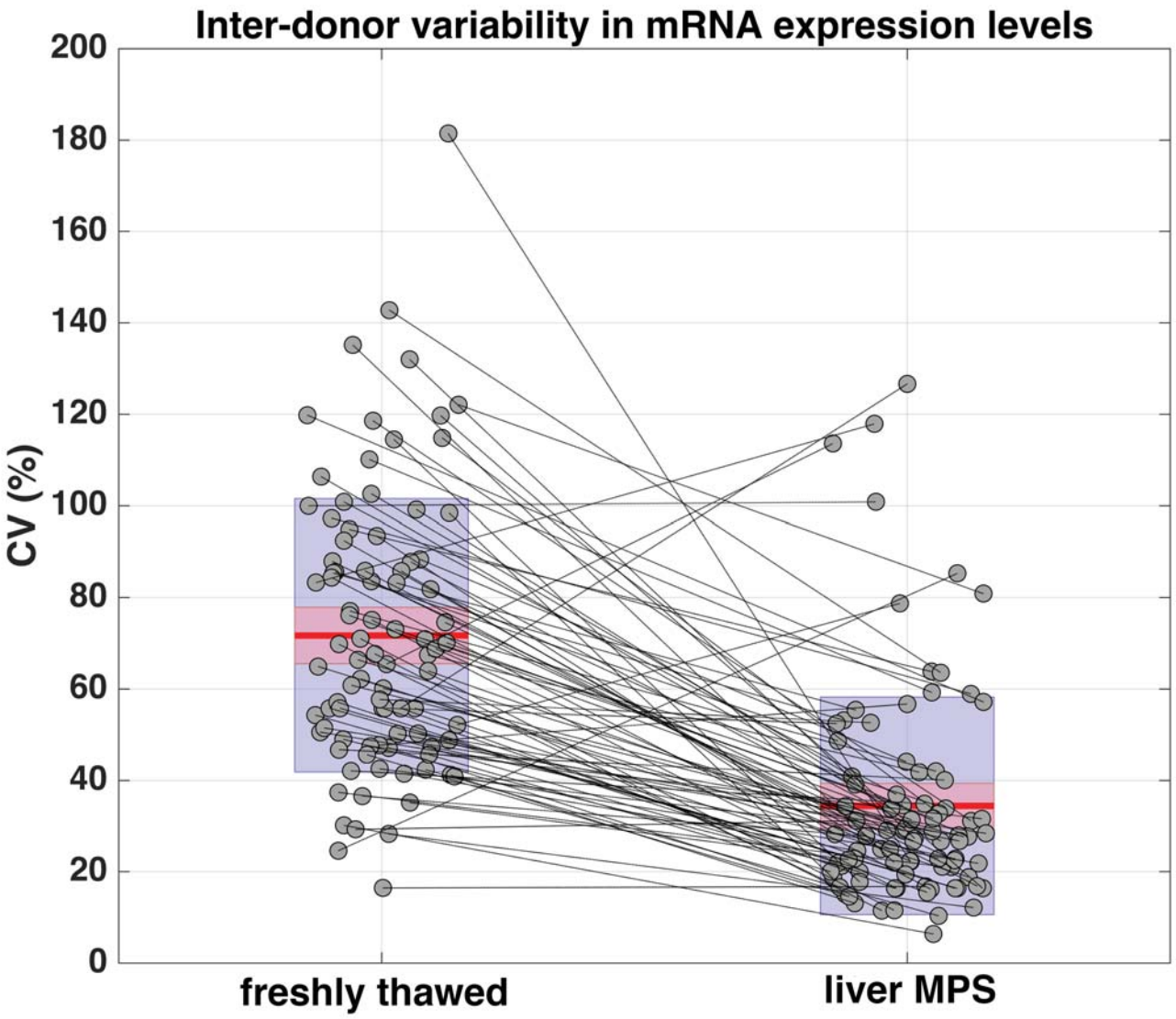

Figure S8. Comparison of inter-donor variability in mRNA expression levels between freshly thawed hepatocytes and liver MPS (day 6). CV on the y-axis refers to coefficient of variation in mRNA expression levels across the five donors. Closed gray circles refer to the 90 different genes and black lines connect the CV in freshly thawed hepatocytes with the CV in the liver MPS for the same gene. Red lines correspond to the mean CV (across all genes) in freshly thawed hepatocytes or the liver MPS, purple boxes extend the mean by ± 1 SD (standard deviation) and the pink boxes correspond to 95% confidence intervals around the mean.

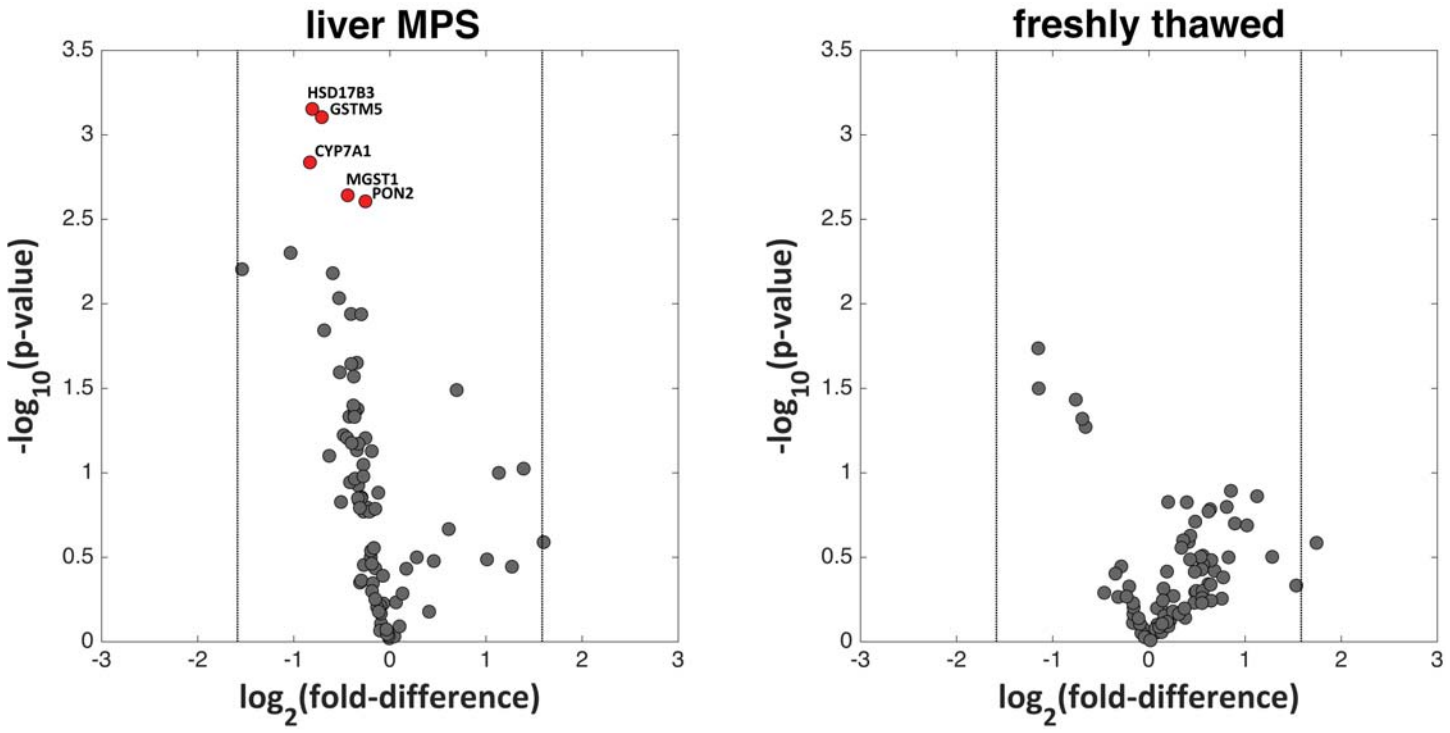

Figure S9. Volcano plots that illustrate the average fold-difference between mRNA expression levels determined across individual donors and mRNA expression levels determined in pooled hepatocytes, along with the associated statistical significance. The subplots on the left and right respectively refer to the liver MPS (day 6) and freshly thawed hepatocytes. The $\log_2$ of the fold-difference is plotted on the x-axis, thus positive values indicate that pooled hepatocytes under-predict the average mRNA expression obtained across the different donors, while negative values indicate over-prediction. Any genes outside the two black vertical lines are associated with a more than 3-fold difference. On the y-axis the $-\log_{10}$ of p-value is plotted, thus the higher the values, the stronger the statistical evidence that the mRNA expression levels across the different donors come from a distribution with mean different to the expression levels determined in the pooled hepatocytes. The genes that reached statistical significance after multiple testing correction are highlighted in red color and the respective gene names are reported.

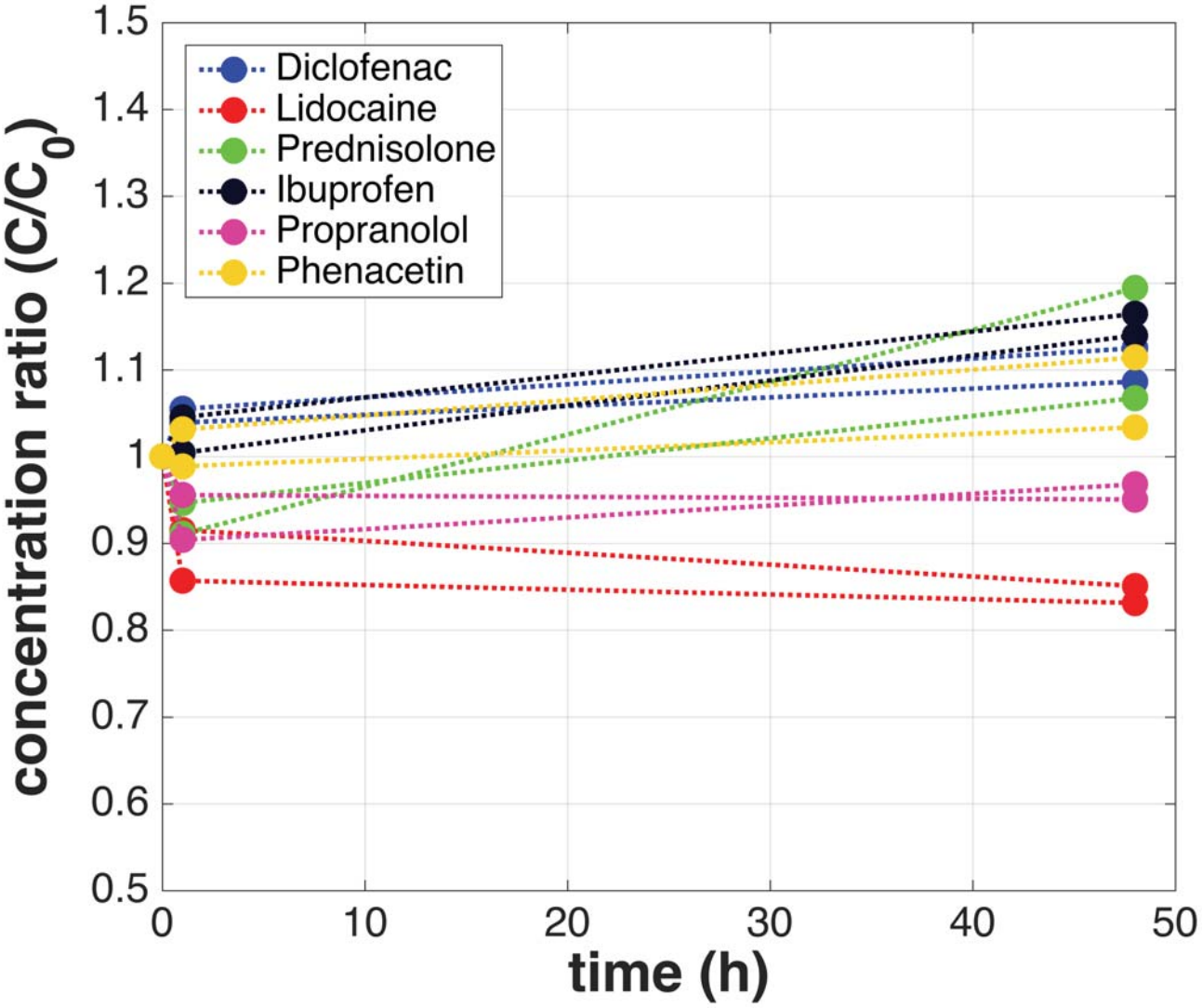

Figure S10. Non-specific binding to the platform materials of the 6 investigated compounds (experiments performed in hepatocyte-free LiverChip). The y-axis refers to the ratio of the measured concentrations at 1h and 48h to the respective initial concentration at time 0 (C_0). Experiments were performed in duplicate for each compound. The media concentrations in cell-free platforms were not perturbed more than 20% from the initial concentrations even after 48 hours of incubation.

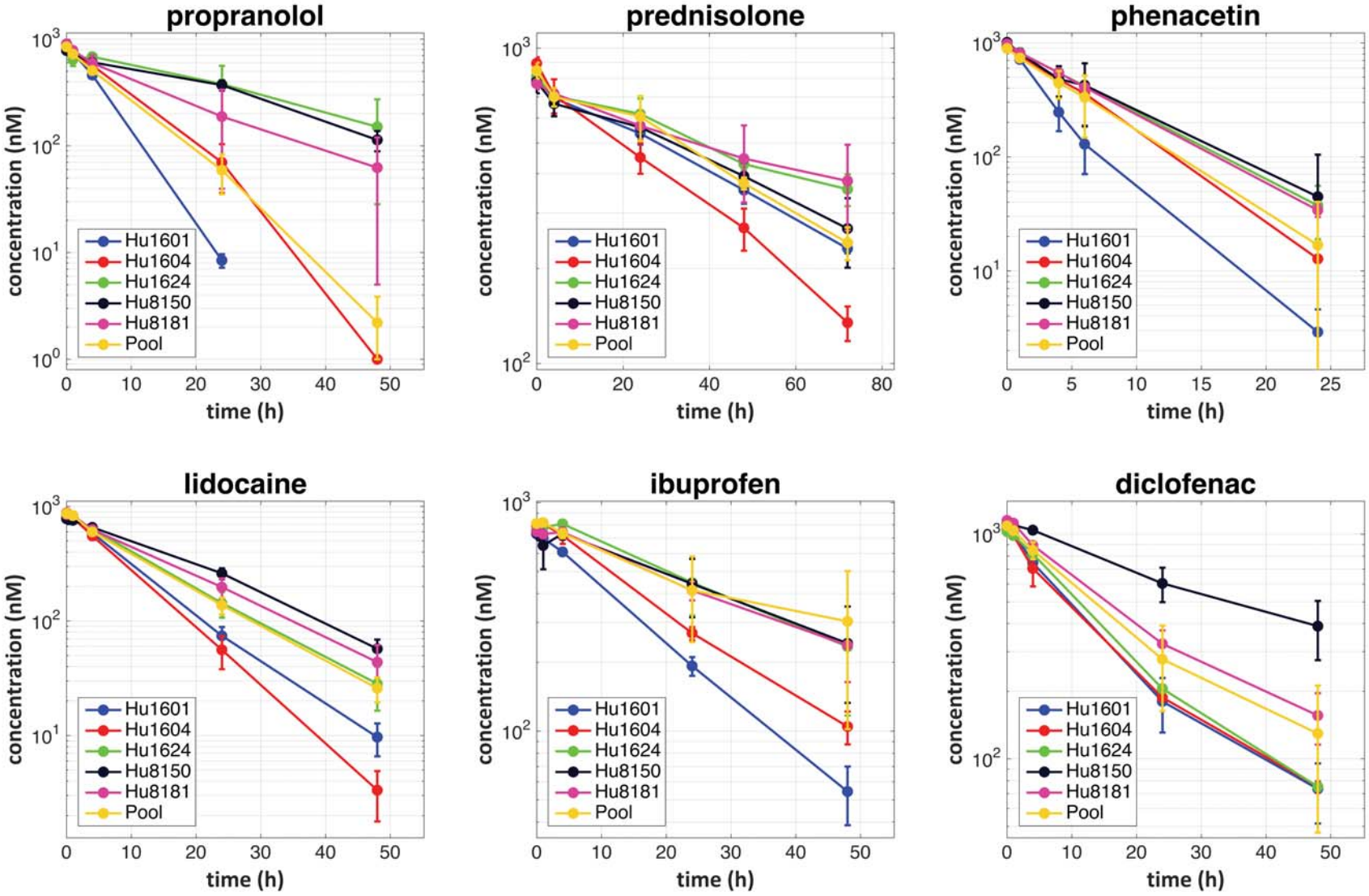

Figure S11. Drug depletion data available to the pharmacokinetic analysis. "Hu1601", "Hu1604", "Hu1624", "Hu8150" and "Hu8181" are lot numbers corresponding to 5 different donors. "Pool" refers to the pool of hepatocytes from the 5 donors. The average concentration profile across different wells is plotted for each donor (or pool of donors) and the bars are indicating inter-well variability by extending the mean by ± 1 standard deviation.

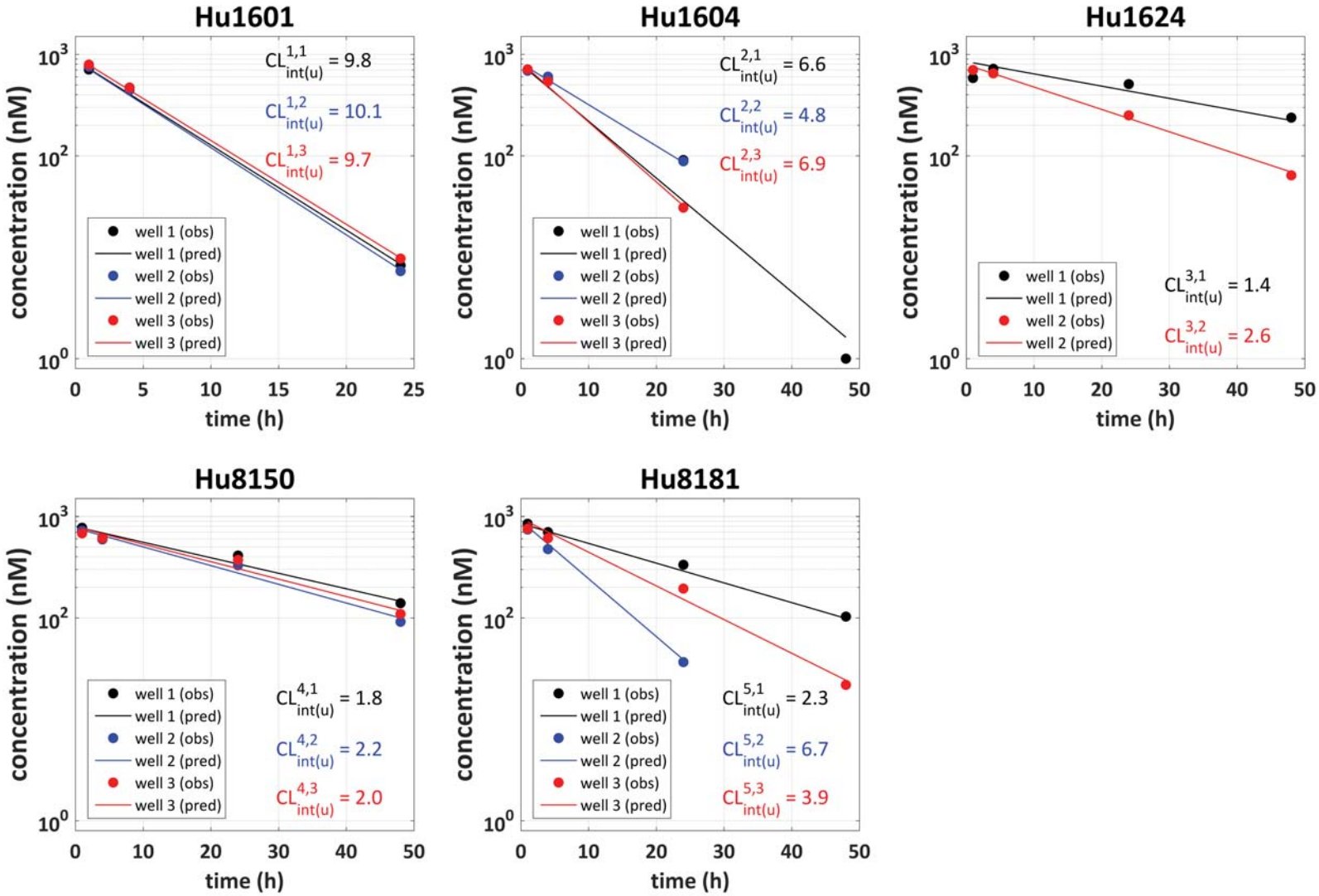

Figure S12. Model predictions regarding the *in vitro* metabolic depletion of propranolol at the level of each individual donor and well. "Hu1601", "Hu1604", "Hu1624", "Hu8150" and "Hu8181" are lot numbers corresponding to 5 different donors. Closed circles represent the observed data in each donor and well and solid lines represent the respective model predictions. The model (empirical Bayes) estimates for the unbound intrinsic clearance referring to the ith donor and the jth well ($CL^{ij}_{\mathrm{int}(u)}$) are also reported (in μL/min/10^6 cells).

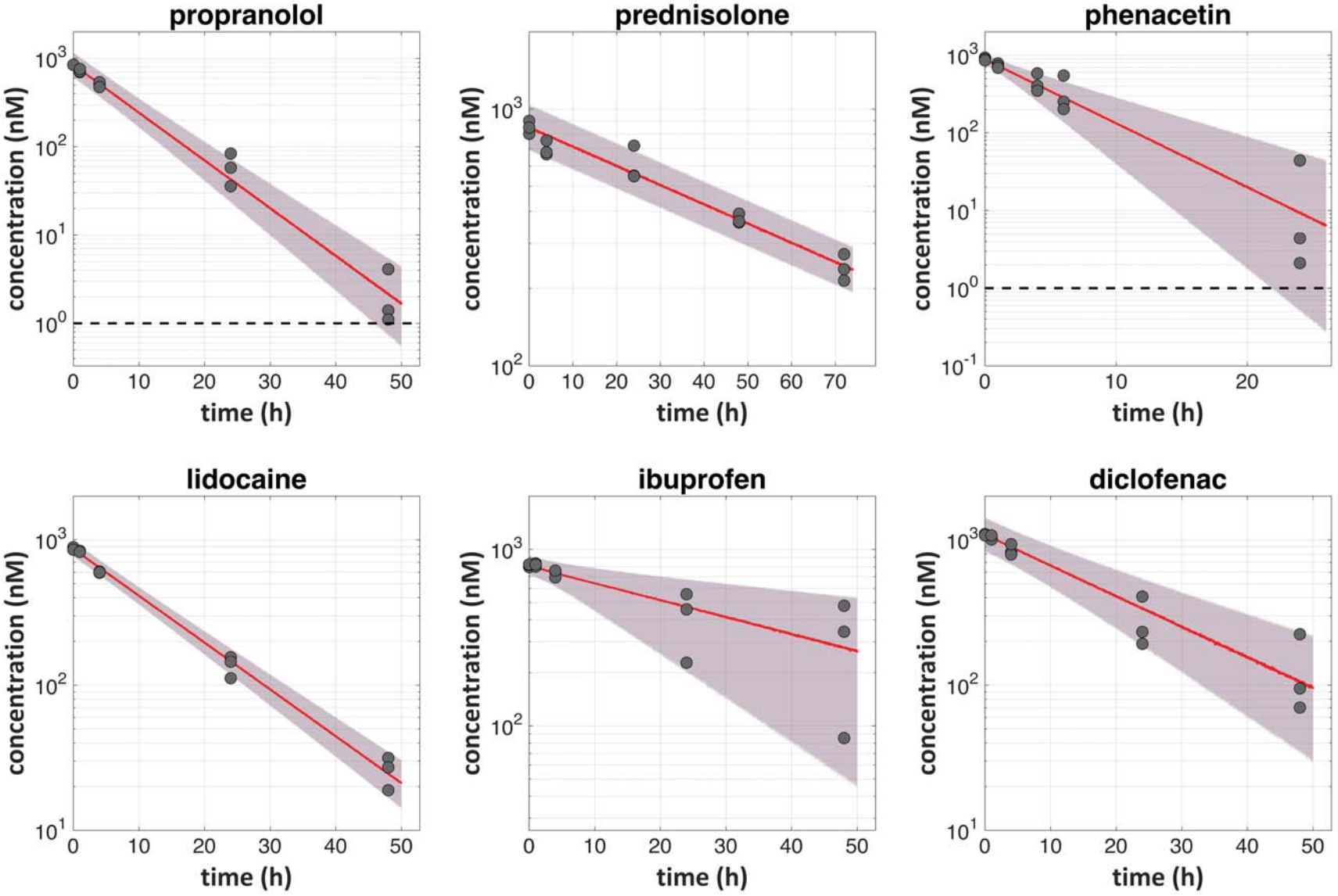

Figure S13. Model description of the drug depletion data determined in pooled hepatocytes. Closed gray circles represent the observed concentrations in media; highlighted with purple are the areas between the 5th and 95th percentiles of model simulations that take into account the different levels of variability (90% prediction intervals), whereas the red solid line represents their median (median prediction); the horizontal dashed black line represents the limit of quantification.

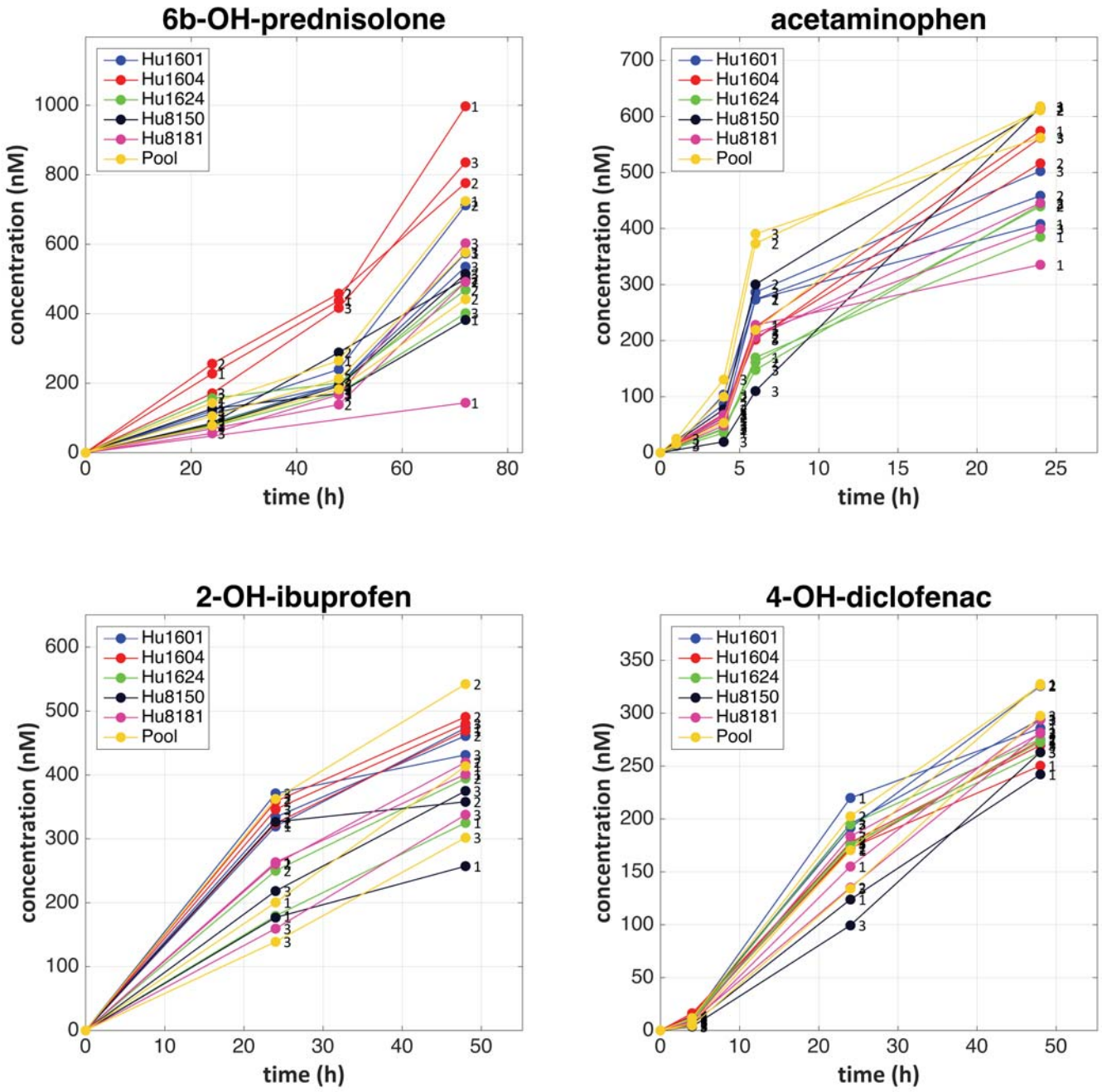

Figure S14. Available metabolite formation data across different donors/wells. "Hu1601", "Hu1604", "Hu1624", "Hu8150" and "Hu8181" are lot numbers corresponding to 5 different donors. "Pool" refers to the pool of hepatocytes from the 5 donors. The small numbers on the right of each concentration point (taking values of 1, 2 or 3) aim to distinguish different wells across the same donor (or pool of donors).

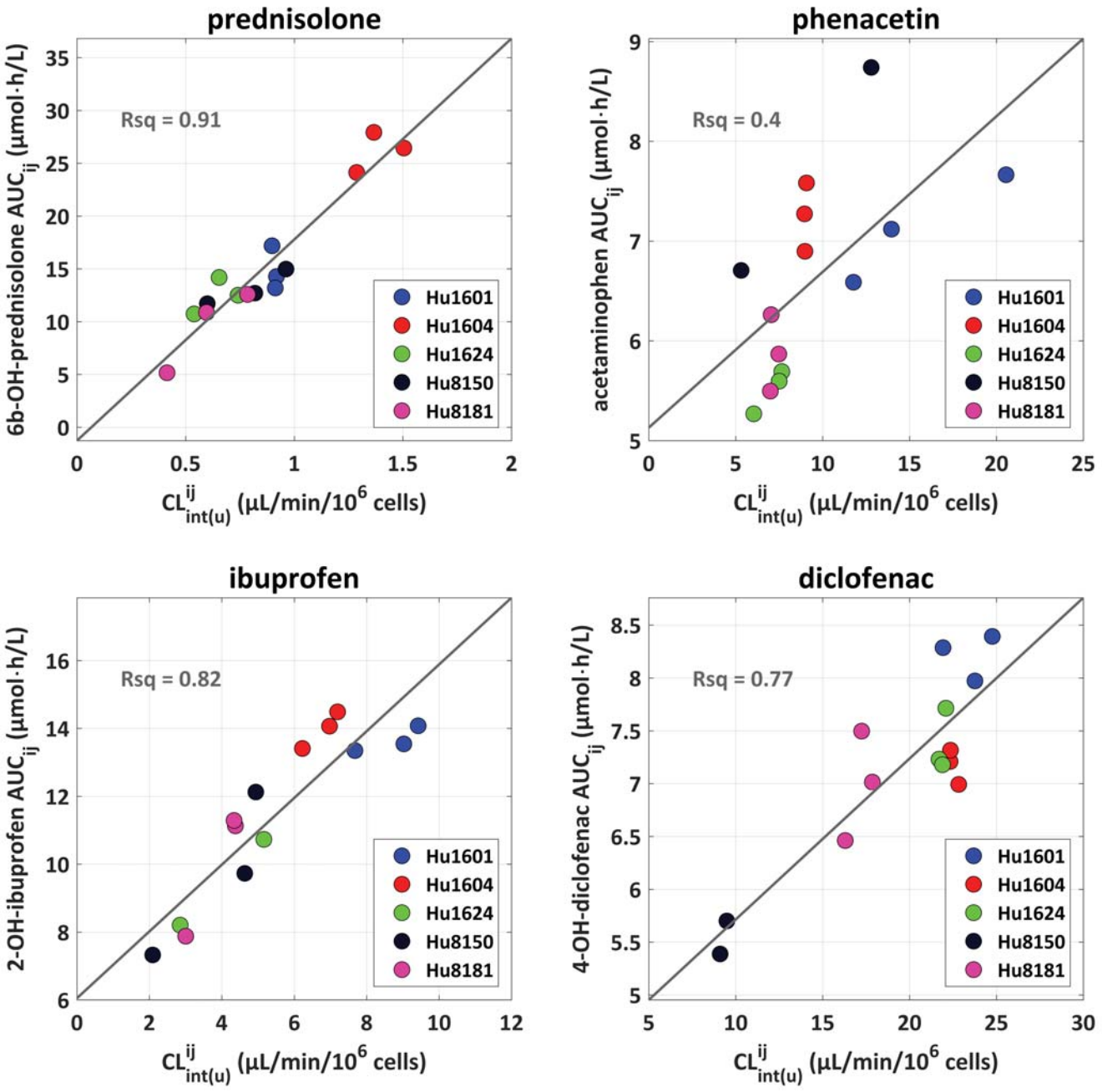

Figure S15. Correlation between the unbound intrinsic clearance for drug depletion regarding to the ith donor / jth well (x-axis) and the area under the metabolite concentration-time curve in the same donor / well (y-axis). The grey line represents the linear regression and "Rsq" the respective coefficient of determination (or R-squared) value. Data referring to different donors are highlighted with different colors (see legend). "Hu1601", "Hu1604", "Hu1624", "Hu8150" and "Hu8181" are lot numbers corresponding to the 5 different donors.

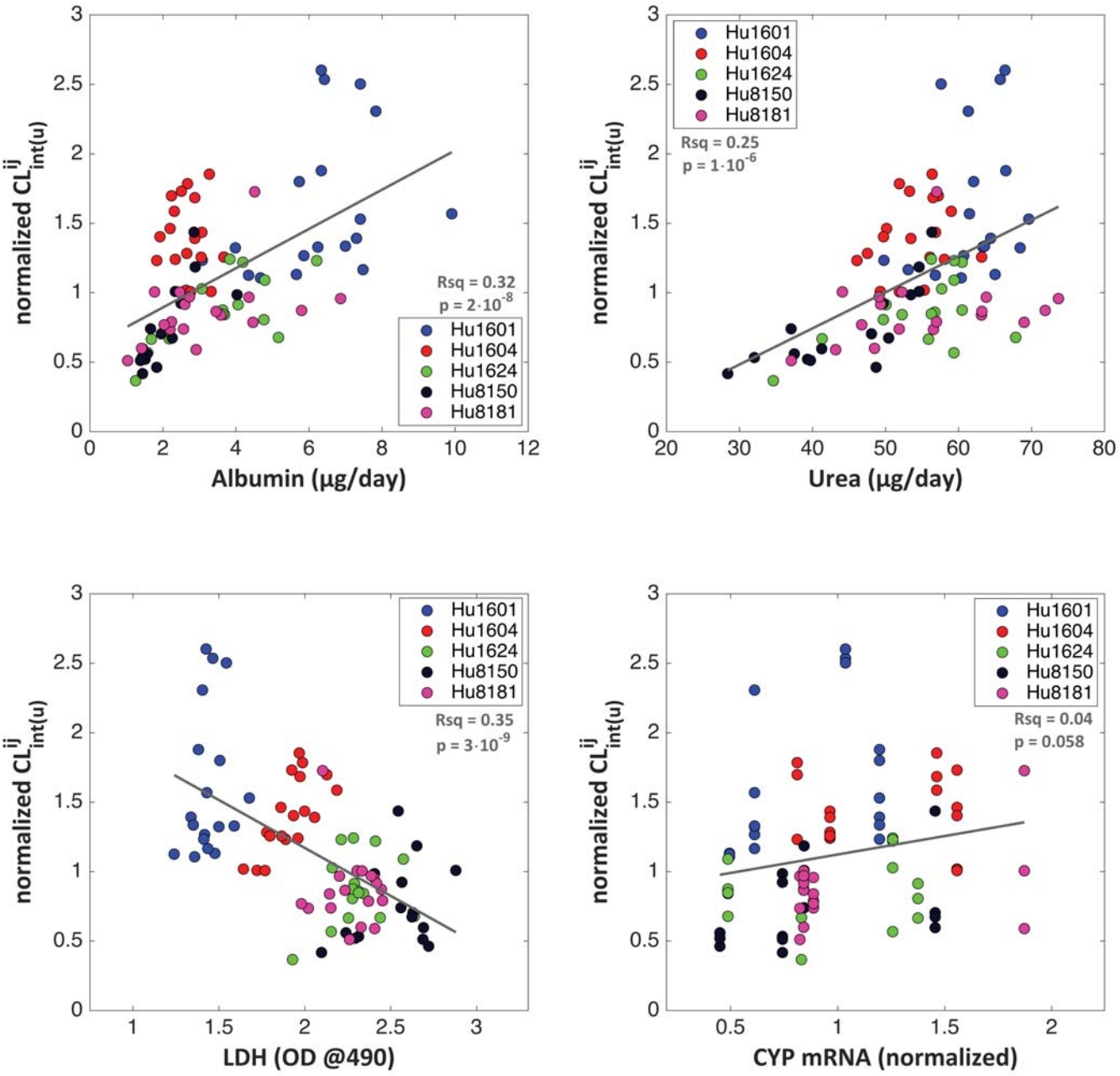

Figure S16. Normalized intrinsic clearance values obtained across different donors/wells plotted against the respective phenotypic metrics of interest (upper left: pre-dose albumin production, upper right: pre-dose urea production, bottom left: pre-dose LDH release, bottom right: normalized CYP mRNA levels). Data referring to different donors are highlighted with different colors (see legend). The grey line represents the linear regression. Inside each subplot we also report as "Rsq" the respective coefficient of determination (or R-squared) value; and as "p" the F-test p-value for a significant linear regression relationship between the response (clearance in y-axis) and the predictor variable (x-axis). LDH levels are expressed in optical density (OD) units at 490nm.

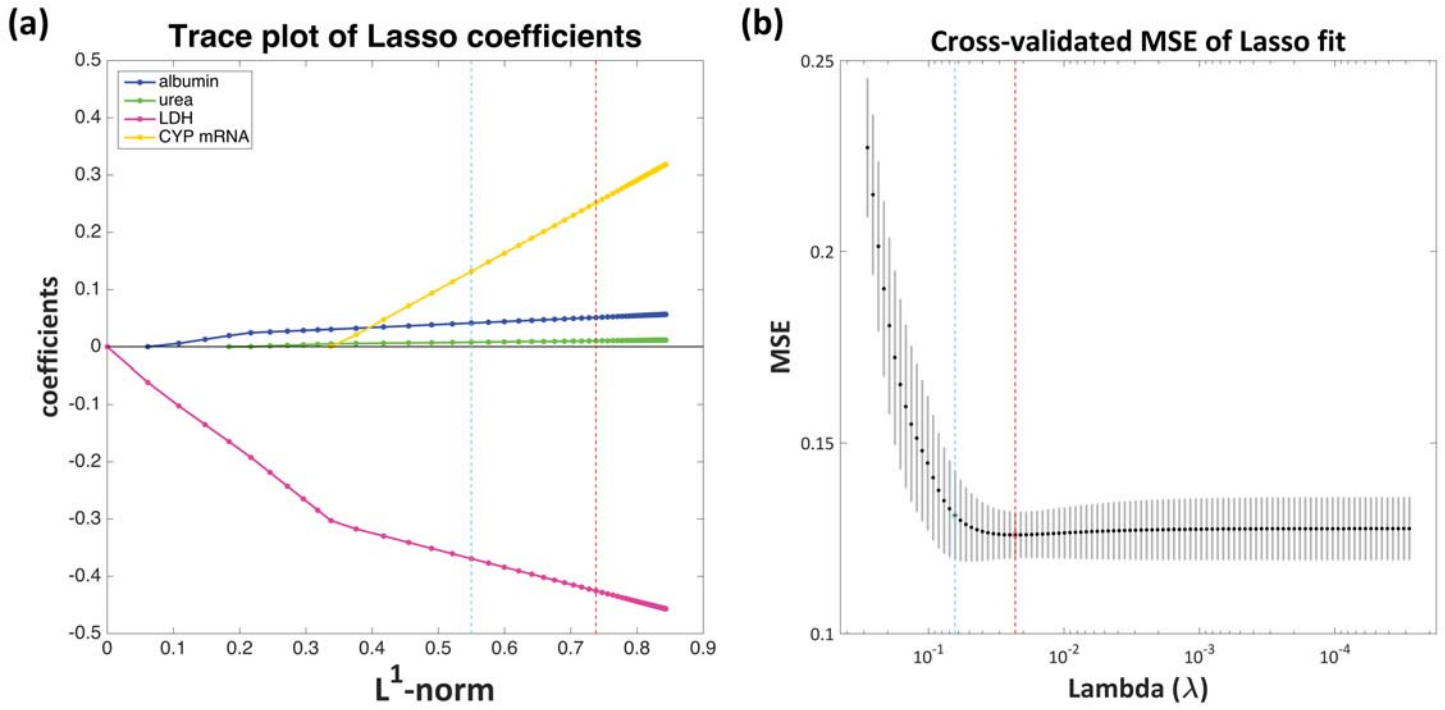

Figure S17. Trace plot referring to the regularization paths of the Lasso coefficients (a) and the cross-validated mean square error (MSE) of the Lasso fit for different values of the regularization parameter λ (b). In (a) the regularization paths of the 4 predictors are illustrated with different colors (see legend). In (b) each of the black dots represents the cross-validated MSE for different values of the regularization parameter λ, while the grey error-bars stretching out of each dot correspond to the associated standard error; the red vertical line highlights the value of λ that minimizes the cross-validated MSE (minimum MSE rule for tuning λ); the cyan vertical line highlights the highest value of λ (more regularization) that still gives a model with cross-validated MSE within 1 standard error of the minimum MSE (one standard error rule for tuning λ, which is a common alternative to the minimum MSE rule). The models obtained using either of these two rules are highlighted with equivalent color vertical lines in the trace plot of the Lasso coefficients (a). It can be observed in (a) that irrespectively of the rule used for tuning λ, all 4 variables have non-zero coefficients and thus are identified as significant predictors of intrinsic clearance. The results form the model obtained with the minimum MSE rule (red line) were finally reported. The trace plot of the Lasso coefficients (a) indicates that among all predictors, the CYP mRNA levels is the variable contributing the less to the predictive performance of the model, while LDH release levels is the variable contributing the most (the coefficients associated to CYP mRNA levels and LDH release are the first and the last respectively to shrink to zero when regularization/penalization is increased ($L^1 - norm$ is decreased)).

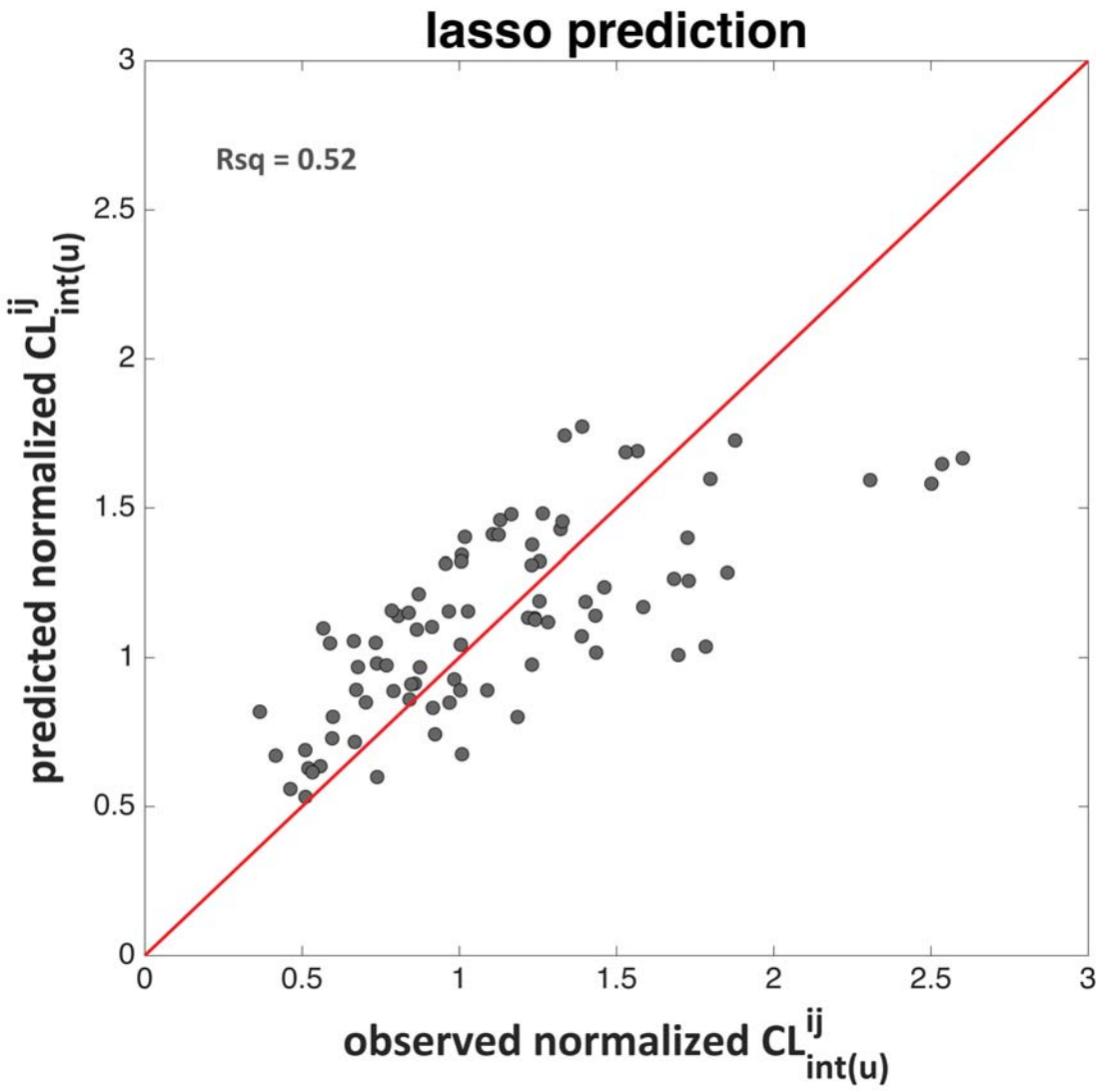

Figure S18. Performance of the Lasso regression model. The predicted from the Lasso model normalized intrinsic clearance values (using the model incorporated metrics/predictors measured across different donors/wells) are plotted against the respective observed normalized intrinsic clearance values (determined through modeling of the drug depletion data). The red diagonal line represents the line of unity. "Rsq" is the coefficient of determination (or R-squared) value.

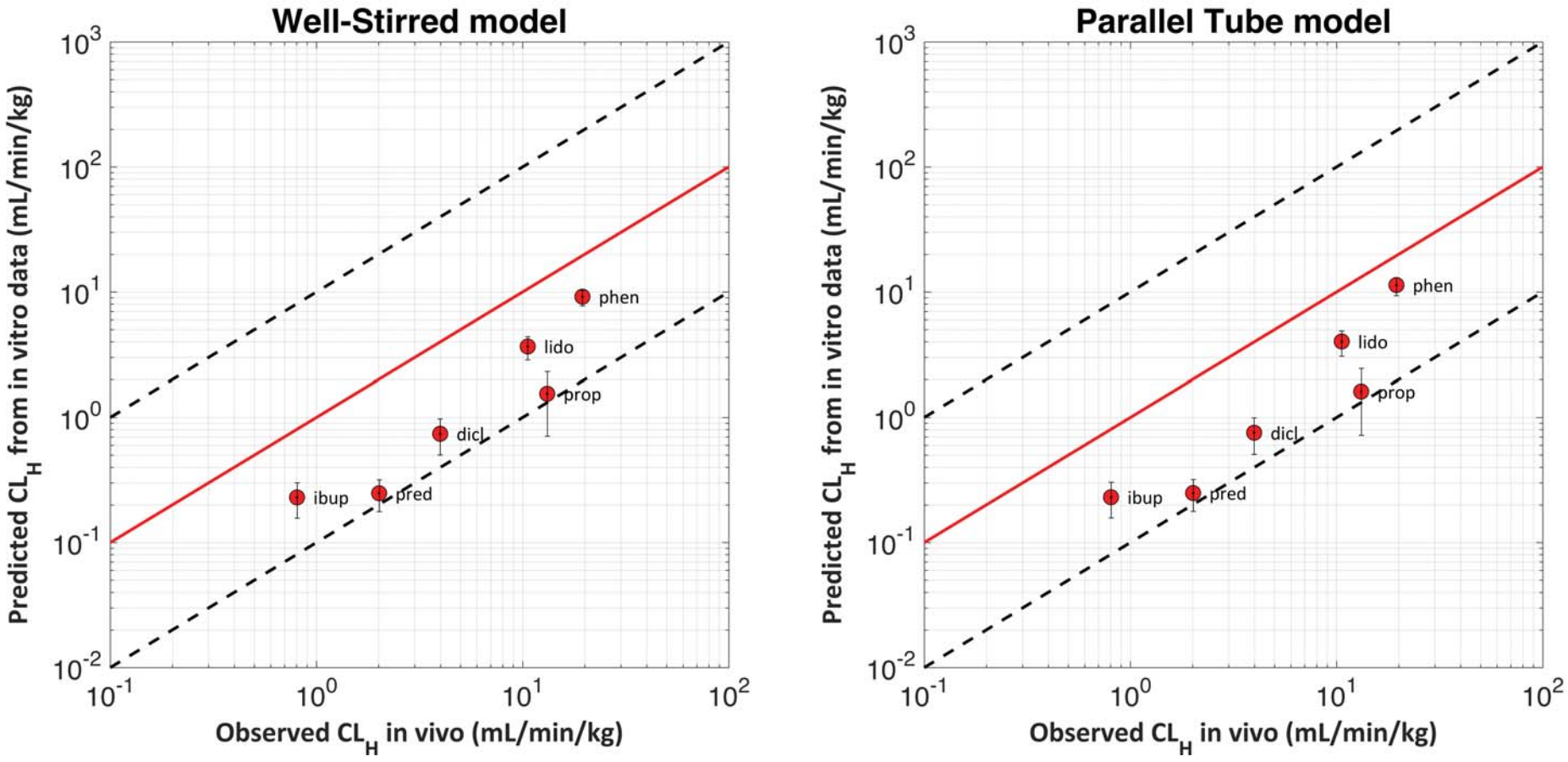

Figure S19. Comparison between the observed hepatic clearance *in vivo* (x-axis) and the predicted hepatic clearance from the *in vitro* data (y-axis) using the well-stirred (left) or the parallel tube (right) liver model. In red is the line of unity and the dashed black lines represent the 10-fold error. The vertical bars on each prediction represent 95% confidence intervals taking into account the uncertainty (standard error) associated with the *in vitro* estimate of intrinsic clearance. The labels on the right of each prediction are abbreviations of the corresponding compound's name: ibup for ibuprofen; pred for prednisolone; dicl for diclofenac; prop for propranolol; lido for lidocaine; and phen for phenacetin.