## Online Methods

The presented liver model is a multi-scale model comprising the metabolism of individual hepatocytes on cellular scale (Figure 1A), the individual sinusoidal unit on tissue scale (Figure 1B), the representation of lobulus via integration of multiple sinusoidal units (lFigure 1C) and the liver in individual subjects and variability in the population (Figure 1D).

### *Availability of data and models*

The complete source code of modelling and analysis is open source and available under GPL from<https://github.com/matthiaskoenig/multiscale-galactose>. All literature data sets are made accessible on request. The mathematical model on cellular and tissue-scale are published as SBML under creative commons (CC BY-SA 4.0) in the Supplementary Information as well as on Biomodels.org and JWS Online. A HTML representation of the information is available in the Supplement.

### *Numerical integration*

The hepatocyte and sinusoidal unit models are kinetic model based on ordinary differential equation (ODE) systems. All numerical integrations were performed directly with the SBML with libRoadRunner v1.3 {Somogyi2014, Somogyi2015} with absolute and relative tolerances of 1E-6 on a computer cluster. libRoadRunner was extended to efficiently handle the large models via ...

All results were stored in a database.

### *Cellular scale - galactose metabolism*

The kinetic model of galactose metabolism for individual hepatocytes consists of three main enzymatic steps i) the phosphorylation of galactose (gal) to galactose 1-phosphate (gal1p) catalysed by galactokinase (GALK, EC 2.7.1.6); ii) the conversion of gal1p to UDP-galactose (udpgal) by galactose-1-phosphate uridyl transferase (GALT, EC 2.7.7.10) and iii) the interconversion of udpgal and UDP-glucose (udpglc) by UDP-galactose 4´-epimerase (GALE, EC 5.1.3.2) {Novelli2000, Petry1998}. Galactose can enter glycolysis as glucose-1 phosphate (glc1p), one of the GALT reaction products, or can be incorporated as udpgal, the substrate donor of all galactosylation reactions, in glycoproteins and glycolipids {Novelli2000}. The alternative processes important in galactosemias and ATP synthesis (ATPS) and NADP reduction (NADPR) for cofactor regeneration were added to the model. Detailed information on metabolites, initial concentrations, rate equations and enzymatic parameters is provided in Supplementary Table 1 and Supplementary Table 2. All literature based kinetic parameters were included in SABIO-RK {Wittig2012} (see Supplementary Tables and SBML annotations). Maximal enzyme activities (Vmax) were chosen to achieve good correspondence of model simulations with reported galactose elimination rates in young subjects (20 years).

### *Tissue scale - sinusoidal unit*

The tissue-scale model of the sinusoidal unit (Figure 1B) consists of a central blood vessel (sinusoid) surrounded by the space of Disse and adjacent hepatocytes in cylindrical geometry with parameters in Supplementary Table 3 and Supplementary Table 4. The periportal (pp) and perivenous (pv) blood compartments are located adjacent to the first and last sinusoidal volume, respectively. A single sinusoidal unit consists of Nc hepatocytes with each cell having a single associated sinusoid and Disse volume. In the sinusoid substances are transported by blood flow and diffusion, in the space of Disse solely by diffusion. Red blood cells (RBC) are constricted to the sinusoid, whereas all other model substances smaller than the fenestrae (rsubstance≤rfen) pass in the space of Disse owing to the fenestration of the endothelial cells {Wisse1985}, i.e. galactose, albumin, sucrose and water. Galactose and water are exchanged between the space of Disse and the hepatocytes, whereas sucrose and albumin are restricted to the space of Disse.

Diffusion and blood flow are modelled via discretized one-dimensional diffusion and convection equations (analogue to {Konig2013}). The diffusion through the sinusoidal fenestration, small cylindrical channels in the endothelial cells is described via pore theory {Pappenheimer1953, Renkin1954}. The total restriction to diffusion due to the combined effects of steric hindrance at the entrance of the pores and frictional resistance within the pores for substance a with radius ra is given as actual diffusion Da relative to unhindered Diffusion Da,0 with radius of the substance ra

and pore radius rfen as DaDa,0=(1-rarfen)21-2.104 (rarfen)+2.09 (rarfen)3-0.95 (rarfen)5{Renkin1954}.

### *Liver region of interest*

The heterogeneity of sinusoidal units within a lobulus was modeled via Monte Carlo simulation of varying sinusoidal units based on experimental distributions for parameters of ultrastructure (sinusoidal length Lsin, sinusoidal radius ysin, width space of Disse ydis, hepatocyte sheet thickness ycell) and microcirculation (sinusoidal blood flow velocity vblood). The output of a region of interest was calculated via integration of the results from N=1000 sinusoidal units. All parameter distributions were assumed log-normal and statistically independent of each other. Distributions of ysin, vblood and ycell were fitted based on maximum-likelihood method for uni-variate distributions . For Lsin and ydis the log-normal parameters were calculated from reported mean m and standard deviation std via stdlog=log(1+std²m²)and mlog=log(m²m²+std²). All parameters and references are given in Supplementary Table 4.

### *Variation in Perfusion*

For the variation of perfusion, i.e. blood flow per tissue volume, the distribution of blood flows was scaled via pscaled(vblood)= p(fflow\*vblood) mit fflow=1 corresponding to the experimental microcirculation.

### *Multiple indicator dilution curves*

The multiple indicator dilution curves under varying unlabeled galactose concentration were modeled via: i) running simulation to steady state under given unlabeled galactose concentration; ii)

For comparison of the simulated diluat

The catheter and nonexchangeable vessel transit time t0 were estimated from the time of first appearance of radioactivity above background levels in the experimental and simulated dilution curves.

Integration of the single a region of interest for the periportal output was done via the volume flow weighted average of the individual sinusoidal units

cvenuoss(t)=i=1Nwicpps(t)=i=1NQsinunit,ii=1NQsinunit,icpps(t)

The model is a distributed model of flow based on parallel, non-interacting sinusoids joined at the venous terminus. The dispersion characteristics is due to the *a priori*  incorporated experimental velocity and path length variations within the ensemble of sinusoids {Weiss1995}

The dilution curves were simulated with reported GEC values for dogs of ~ 0.5 \* GEC of humans (see supplement).

### *Galactose Elimination, Extraction Ratio and Clearance*

The galactose elimination rate (GE[k]), extraction ratio (ER[k]) and clearance (CL[k]) for a single sinusoidal unit k are calculated from the sinusoidal bloodflow Q[k] and periportal and perivenous concentrations of galactose cppgal[k]andcpvgal[k]in steady state {Schirmer1986}

GE[k]=Q[k] (cppgal[k]-cpvgal[k])

ER[k]=(cppgal[k]-cpvgal[k])cppgal[k]

CL[k]=Q[k] (cppgal[k]-cppgal[k])cppgal[k]

The apparent GE, ER and CL in a tissue region of interest, i.e. the integrated response over Nsin sinusoidal units, are calculated with the volume of the individual sinusoidal unitsV[k]=and the short notations xtot=k=1Nsinx[k]and<x>=1Nsink=1Nsinx[k]=1Nsinxtotas

GEROI=QtotVtot(<cppgal-cpvgal>)

ERROI=<cppgal-cpvgalcppgal>

CLROI=QtotVtot<cppgal-cpvgalcppgal>

PROI=QtotVtot

The integrated clearance and results in values pedepend on the total blood

### *Liver*

To scale the output from region of interests to main things have to be taken into account. 1) The region of interests describe the parenchymal liver tissue. To account for whole liver function the non parenchymal volume of the liver, mainly consisting of large vessel volume, has to be taken into account. The parenchymal fraction of the liver is arround 80% (ftissue=0.8).

The relatiionship between tissue perfusion Ptissue=QtissueVtissue=i=1NQsinuiti=1NVsinunit and liver perfusion Pliver=QliverVliveris given via Pliver=QtissueVtissue+Vnotissue=(12-ftissue)QtissueVtissue=(12-ftissue)Ptissue

Via the relationship for normal perfusion of 1.2ml/min/ml an necessary adaption of the microcirculation of fflow=0.3 results corresponding to a mean sinusoidal flow velocity of 81µm/s. This is still in the range of the experimentally obtained values. OPS values and microcirculation is taken on the surface of the liver, with partly larger arterial components and properly not representative of the whole liver.

### *Liver in aging*

The age dependent change in total liver volume, blood flow and perfusion are taken into account in the age-dependent GAMLSS curves. The additional changes in ultrastructure of the liver (pseudocapillarization) were modeled using the age-dependent change in fenestration, i.e. change in fenestration number per area (Nfen), and endothelial thickness (yend) based on experimental data as input for the model predictions (supplementary information).

Practically, for different combinations of (Nfen, yend) corresponding to certain ages the GEC per tissue and perfusion were calculated.

### *Individualized predictions*

Based on the integration of detailed kinetic models of sinusoidal units over a region of interest the metabolic function for given liver structure, morphology, perfusion and metabolic function can be calculated. The total hepatic function of a person results from this regional function with the actual liver volume and blood flow of the person. This is achieved via prediction of the individual liver characteristics based on the anthropomorphic data of the person, i.e. based on sex, age, bodyweight, height and body surface area.

In a first step generalized additive models for location, scale and shape (GAMLSS) {Stasinopoulos2007} were fitted to describe the correlations between single liver features and single anthropomorphic features based on individual data from >3000 subjects from >30 studies (supplement GAMLSS). Individual models for male, female and all data were fitted. The resulting models enable the prediction of the probability distributions of liver features for single anthropomorphic features based on gender, for instance the distribution of liver volumes depending on age for femalespsex=F,volLiverage.Importantly, the observed population variability is part of the model.

In a second step, the information of the single feature models is combined under the assumption of statstical independence to generate a best estimate of liver volume based on the set of antropomorphic features observed in the person. The result is a personalized probability distribution of liver volumes pk(volLiver) for the subject k with sex=S, age=A, bodyweight=B, height=H and bsa=BS

pk(volLiver)= p(volLiver|sex=S, age=A, bodyweight=B, height=H, bsa=BS)=

psex=S,volLiverage(volLiver|age=A)

psex=S,volLiverkgage(volLiver|age=A, bodyweight=B)

psex=S,volLiverbodyweight(volLiver|bodyweight=B)

psex=S,volLiverkgbodyweight(volLiver|bodyweight=B)

psex=S,volLiverheight(volLiver|height=H)

psex=S,volLiverkgheight(volLiver|height=H, bodyweight=B)

psex=S,volLiverbsa(volLiver|bsa=BS)

psex=S,volLiverkgbsa(volLiver|bsa=BS, bodyweight=B)

Hepatic blood flows is calculated in a similar manner, but taking the additional correlation information between liver volume and blood flow into account. The resulting probability distribution of hepatic blood flow for person k with given liver volume V is

pk(flowLiver|volLiver=V)=p(flowLiver|sex=S, age=A, bodyweight=B, bsa=BS)psex=S,flowLivervolLiver(flowLiver|volLiver=V)

psex=S,flowLiverage(flowLiver|age=A)

psex=S,flowLiverkgage(flowLiver|age=A, bodyweight=B)

psex=S,flowLiverbodyweight(flowLiver|bodyweight=B)

psex=S,flowLiverkgbodyweight(flowLiver|bodyweight=B)

psex=S,flowLiverbsa(flowLiver|bsa=BS)

psex=S,flowLiverkgbsa(flowLiver|bsa=BS, bodyweight=B)

psex=S,flowLivervolLiver(flowLiver|volLiver=V)

Finally, the metabolic function of person k, in this case the galactose elimination capacity, is calculated by scaling the metabolic function per tissue volume for given perfusion and structure, to total blood flow and liver volume (flowLiverk and volLiverk)via the metabolic functions calculated for regions of the liver.

GECk = fGEC\_per\_volLiver(flowLiverk/volLiverk) \* volLiverk

Via Monte Carlo simulation, i.e. repeated sampling from the individualized probability distributions pk(volLiver) and pk(flowLiver|volLiver)the distribution of liver volumes, blood flows and metabolic function in people with the given anthropomorphic features can be calculated.

Body surface was (BSA) is calculated via DuBois formula from bodyweight and height BSA=0.007184bw0.725h0.425{Moesteller1987} .

### Population variability

To calculate the population variability in liver function the prediction of liver volume, blood flow and GEC was performed for a large cohort representative of the US population. The NHANES {NHANES} survey data between years 1999 - 2012 was used, with subjects filtered based on body mass index (18.5 ≤ BMI ≤ 24.9) and ethnicity (Non-Hispanic White). For all subjects with complete data sets of age, gender, height, and body weight the prediction was performed. Using the Monte Carlo approach, repeated computations based on inputs selected at random from statistical distributions for each input parameter are conducted to provide a statistical distribution of the output. Using high percentile (e.g. 95th) and 50th percentile, the intraspecies variability can be calculated. To derive this information, Monte Carlo simulations based on distributions of input parameters have frequently be used. (Lipscomb et al., 2003; Gentry et al., 2002; Haber et al., 2002; Lipscomb and Kedderis, 2002; Timchalk et al., 2002; Bogaards et al., 2001; El-Masri et al., 1999; Thomas et al., 1996a, b).

### *Classification & ROC curves*

ROC curves were calculated based on the following classificator

TODO

Performance of the classifier was evaluated against logistic regression on the same datasets.