



# A novel variant of the $^{13}\text{C}$ -methacetin liver function breath test that eliminates the confounding effect of individual differences in systemic $\text{CO}_2$ kinetics

Hermann-Georg Holzhütter<sup>1</sup> · Tilo Wuensch<sup>2</sup> · Robert Gajowski<sup>3,4</sup> · Nikolaus Berndt<sup>1,5</sup> · Sascha Bulik<sup>6</sup> · David Meierhofer<sup>3</sup> · Martin Stockmann<sup>2</sup>

Received: 18 June 2019 / Accepted: 27 January 2020 / Published online: 4 February 2020  
© The Author(s) 2020

## Abstract

The principle of dynamic liver function breath tests is founded on the administration of a  $^{13}\text{C}$ -labeled drug and subsequent monitoring of  $^{13}\text{CO}_2$  in the breath, quantified as time series delta over natural baseline  $^{13}\text{CO}_2$  (DOB) liberated from the drug during hepatic CYP-dependent detoxification. One confounding factor limiting the diagnostic value of such tests is that only a fraction of the liberated  $^{13}\text{CO}_2$  is immediately exhaled, while another fraction is taken up by body compartments from which it returns with delay to the plasma. The aims of this study were to establish a novel variant of the methacetin-based breath test LiMAX that allows to estimate and to eliminate the confounding effect of systemic  $^{13}\text{CO}_2$  distribution on the DOB curve and thus enables a more reliable assessment of the hepatic detoxification capacity compared with the conventional LiMAX test. We designed a new test variant (named "2DOB") consisting of two consecutive phases. Phase 1 is initiated by the intravenous administration of  $^{13}\text{C}$ -bicarbonate. Phase 2 starts about 30 min later with the intravenous administration of the  $^{13}\text{C}$ -labelled test drug. Using compartment modelling, the resulting 2-phasic DOB curve yields the rate constants for the irreversible elimination and the reversible exchange of plasma  $^{13}\text{CO}_2$  with body compartments (phase 1) and for the detoxification and exchange of the drug with body compartments (phase 2). We carried out the 2DOB test with the test drug  $^{13}\text{C}$ -methacetin in 16 subjects with chronic liver pathologies and 22 normal subjects, who also underwent the conventional LiMAX test. Individual differences in the systemic  $\text{CO}_2$  kinetics can lead to deviations up to a factor of 2 in the maximum of DOB curves (coefficient of variation  $\text{CV} \approx 0.2$ ) which, in particular, may hamper the discrimination between subjects with normal or mildly impaired detoxification capacities. The novel test revealed that a significant portion of the drug is not immediately metabolized, but transiently taken up into a storage compartment. Intriguingly, not only the hepatic detoxification rate but also the storage capacity of the drug, turned out to be indicative for a normal liver function. We thus used both parameters to define a scoring function which yielded an excellent disease classification ( $\text{AUC} = 0.95$ ) and a high correlation with the MELD score ( $R_{\text{Spearman}} = 0.92$ ). The novel test variant 2DOB promises a significant improvement in the assessment of impaired hepatic detoxification capacity. The suitability of the test for the reliable characterization of the natural history of chronic liver diseases (fatty liver—fibrosis—cirrhosis) has to be assessed in further studies.

**Keywords**  $^{13}\text{C}$ -methacetin · Liver function · Breath test · Compartment model ·  $^{13}\text{CO}_2$  kinetics

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s00204-020-02654-0>) contains supplementary material, which is available to authorized users.

✉ Hermann-Georg Holzhütter  
hergo@charite.de

Extended author information available on the last page of the article

## Introduction

The rationale of dynamic liver function tests consists in using the hepatic elimination rate of a properly chosen test compound as a measure for the overall 'hepatic functional mass'. Liver damage may include dysfunction of subcellular organelles (e.g. endoplasmic reticulum and mitochondria) as well as a reduction in the number of metabolically active hepatocytes, both processes affecting the CYP-450-dependent detoxification rate of a test drug. Breath

tests, in particular, share the principle that a subject is administered an exogenous test compound (e.g. aminopyrine, galactose, phenylalanine or methacetin) in which the common  $^{12}\text{C}$  atom of a functional group has been replaced by the stable  $^{13}\text{C}$  isotope. In the liver,  $^{13}\text{CO}_2$  is enzymatically cleaved from the functional group, typically by an enzyme of the CYP-450 family, enters the plasma and is then expired. As about 1% of the natural carbon atoms of our body compounds are present as the stable isotope  $^{13}\text{C}$ , the abundance of the excess  $^{13}\text{CO}_2$  in the breath is quantified as delta-over-baseline (DOB) defined as relative concentration difference ( $[\text{C}^{13}\text{CO}_2] - [\text{C}^{13}\text{CO}_2]_{\text{baseline}} / [\text{C}^{13}\text{CO}_2]_{\text{baseline}}$  given in per mille (‰)).

Breath tests are attractive for being less invasive, relatively simple and having a high patient acceptance (Bonfrate et al. 2015; Braden et al. 2007). Amongst the various substrates utilized to evaluate quantitative liver function, the  $^{13}\text{C}$ -methacetin breath test (MBT) has shown to be most promising (Buechter et al. 2018; Klatt et al. 1997). In its original design, the MBT is performed by oral administration of the test drug and subsequent measurement of the cumulative percentage of the administered dose of  $^{13}\text{CO}_2$  recovered in the breath over time. Stravitz et al. (2015) reported that the oral MBT was superior to the MELD score in predicting the risk of cirrhotic complications and mortality in patients listed for liver transplantation. The oral  $^{13}\text{C}$ -MBT was also successfully evaluated in the assessment of acute liver injury in a rat model (Zhu et al. 2014).

As individual differences in the intestinal uptake of the test drug may distort the test results, an improved variant of the MBT test (LiMAX test) has been developed at the Charite Berlin based on the intravenous administration of the drug and quasi-continuous monitoring of  $^{13}\text{CO}_2$  in the breath by means of an ultra-sensitive laser system (Stockmann et al. 2009). The LiMAX test has been successfully validated as a valuable non-invasive test in liver surgery for the pre- and postoperative assessment of organ function (Stockmann et al. 2009; Rubin et al. 2016) and for classifying the risk of cirrhotic complications and mortality in patients evaluated for liver transplantation (Jara et al. 2015). Moreover, in a retrospective study encompassing 102 patients with chronic liver disease who underwent a liver biopsy, the LiMAX test performed better than transient elastography (TE) and several serum biomarkers in detecting different stages of liver fibrosis and cirrhosis (Buechter et al. 2018). Nevertheless, the capability of the LiMAX test to reliably discriminate between a non-fibrotic (F0) and mildly-fibrotic (F1) liver was equally poor than observed with the oral version of the MBT (Dinesen et al. 2008; Razlan et al. 2011). Therefore, to become accepted as a reliable clinical tool for the early detection of the onset loss of the liver's functional parenchyma, it requires further improvement of the LiMAX test.

A detailed discussion of various confounding factors limiting the accuracy of the MBT is given in Gorowska-Kowolik et al. (2017). In brief, there are two different categories of confounding factors diminishing the reliability of functional breath tests. First, even if the breath test was able to provide a precise information of the detoxification rate of the test drug, the utility of this parameter for the distinction between normal and diseased livers is restricted by several confounding factors such as the intake of medical drugs, smoking, aging or genetic variability which all may influence the expression level of the targeted metabolic pathway in the liver. Second, the kinetics of  $^{13}\text{CO}_2$  in the exhaled breath does not truly reflect the kinetics of enzymatic  $^{13}\text{CO}_2$  formation in the liver. This is a consequence of the systemic distribution of  $^{13}\text{CO}_2$ : A certain fraction of newly formed  $^{13}\text{CO}_2$  is taken up into several body compartments from which it returns with delay to the plasma. Earlier studies on the systemic  $\text{CO}_2$  dynamics in humans (Barstow et al. 1989, 1990; Irving et al. 1983) revealed large variations in the rate of irreversible  $\text{CO}_2$  elimination and  $\text{CO}_2$  exchange with body compartments. Hence, depending on the size of the transiently trapped  $^{13}\text{CO}_2$  fraction in a specific subject, the DOB values may either over- or underestimate the actual hepatic formation rate of  $^{13}\text{CO}_2$ . This fact is usually disregarded when equating the DOB value with the formation rate of  $^{13}\text{CO}_2$  in the liver.

To correct for the 'CO<sub>2</sub> bias', we recently proposed a novel test variant (named "2DOB") which is initialized with the injection of a standard dose of  $^{13}\text{C}$ -labeled bicarbonate (yielding information on the actual systemic  $\text{CO}_2$  kinetics in the body of the individual) followed by injection of the  $^{13}\text{C}$ -labeled test drug just as in the conventional breath test (Holzhutter et al. 2013). Computer simulations suggested that the predictive power of the proposed 2DOB breath test to reliably assess the CYP-specific hepatic detoxification activity should be significantly higher compared to the conventional breath test. Here we report on a first preliminary study with 38 subjects on the practicability and utility of the proposed 2DOB test.

## Methods

### Subjects

16 patients with different types of liver pathologies (10 males, 6 females) and 22 normal subjects (14 males, 8 females) with no history or biochemical signs of liver disease, were enrolled in this study (for detailed information see Supplementary Table 1). In all except one patient, the diagnosis of the liver disease was histologically proven by histologic evaluation of either needle biopsy or surgical specimen. From healthy controls, a panel of laboratory

parameters (aminotransferases,  $\gamma$ -glutamyltransferase, coagulation factors, bilirubin, alkaline phosphatase and albumin) was determined to exclude presence of a liver disease.

## Experimental protocols

The study was carried out at the Department of Surgery, Charité Campus Mitte and Campus Virchow-Klinikum, Charité-Universitätsmedizin Berlin from July 2016 to February 2018. The study protocol was approved by the ethics committee of the Charité-Universitätsmedizin Berlin and adhered to the latest version of the Declaration of Helsinki, and all the study participants provided written informed consent before enrollment. Every subject underwent two examination sessions, held on separate days.

### 2DOB test

To avoid isotopic effects from the food intake (Jonderko et al. 2008), the test was started after at least 3 h of fasting. The subject was placed in a prone position. To analyze the whole breath continuously, a specifically designed face mask was used for easy collection of exhaled breath. The gas analyses were performed in real time at the bedside using the commercially available FLIP<sup>®</sup> detection device (Humedics GmbH, Berlin, Germany) with a newly developed device for spectral analysis of the  $^{13}\text{CO}_2$  and  $^{12}\text{CO}_2$  amounts in the exhaled breath (Rubin et al. 2011, 2016). The intravenous injection of  $^{13}\text{C}$ -labeled bicarbonate (named  $^{13}\text{C-B}$  in the following) was performed after about 10 min when the DOB baseline has reached stable values. This injection resulted in a rapid increase of DOB values reaching a maximum after a few seconds. For the model-based computational analyses of the experimentally determined DOB data, the time point at which the DOB value of an individual patient has reached the maximum after injection of  $^{13}\text{C-B}$  was designated "0". Then, at time  $T = T_m \approx 30$  min after the injection of  $^{13}\text{C-B}$ , a dose of 2 mg/kg body weight of  $^{13}\text{C}$ -labeled methacetin ( $^{13}\text{C-M}$ ) was administered intravenously followed by a flush of 20 ml of isotonic saline solution. To keep the duration of the full test as short as possible, a time span of 30 min for phase 1 turned out to be sufficient for the unequivocal estimation of the parameters  $k_{-C}$ ,  $k_{+C}$  and  $k_R$ . Administration of  $^{13}\text{C-M}$  gave rise to a second increase of the DOB values followed by a more or less steep decline depending on the methacetin metabolizing capacity of the liver (see Fig. 2). Online sampling and analysis of the  $^{13}\text{CO}_2/^{12}\text{CO}_2$  ratio in the breath were performed with a time resolution of 25 s. To maintain a constant hemodynamic state, test subjects had to rest in a supine position throughout the test.

A compartment model (see below) was used to estimate from the experimentally determined time-dependent DOB

data numerical values for the kinetic parameters characterizing the chemical conversion rate of  $^{13}\text{C-M}$ , the exchange of  $^{13}\text{C-B}$  and  $^{13}\text{C-M}$  with body compartments and the irreversible elimination of  $^{13}\text{C-B}$ . These parameters were then tested for their ability to serve as classifiers for the discrimination between normal and diseased livers (see below).

### LiMAx test

The LiMAx test is basically identical with the second phase of the 2DOB test. After calibration of the DOB baseline curve, the test is initiated with the intravenous injection of a dose of 2 mg/kg body weight of  $^{13}\text{C-M}$ . Blood samples were taken from eight subjects at six time points (= 60, 120, 300, 600, 1200, 1800 s) after  $^{13}\text{C-M}$  injection for the determination of the plasma values of  $^{13}\text{C-M}$ . The classification score of the test, LiMAx, was introduced in (Stockmann et al. 2009) as

$$\text{LiMAx} = \frac{\text{DOB}_{\max} \times R(\text{PDB}) \times P \times \text{BSA} \times \text{MM}}{\text{BW}}, \quad (1)$$

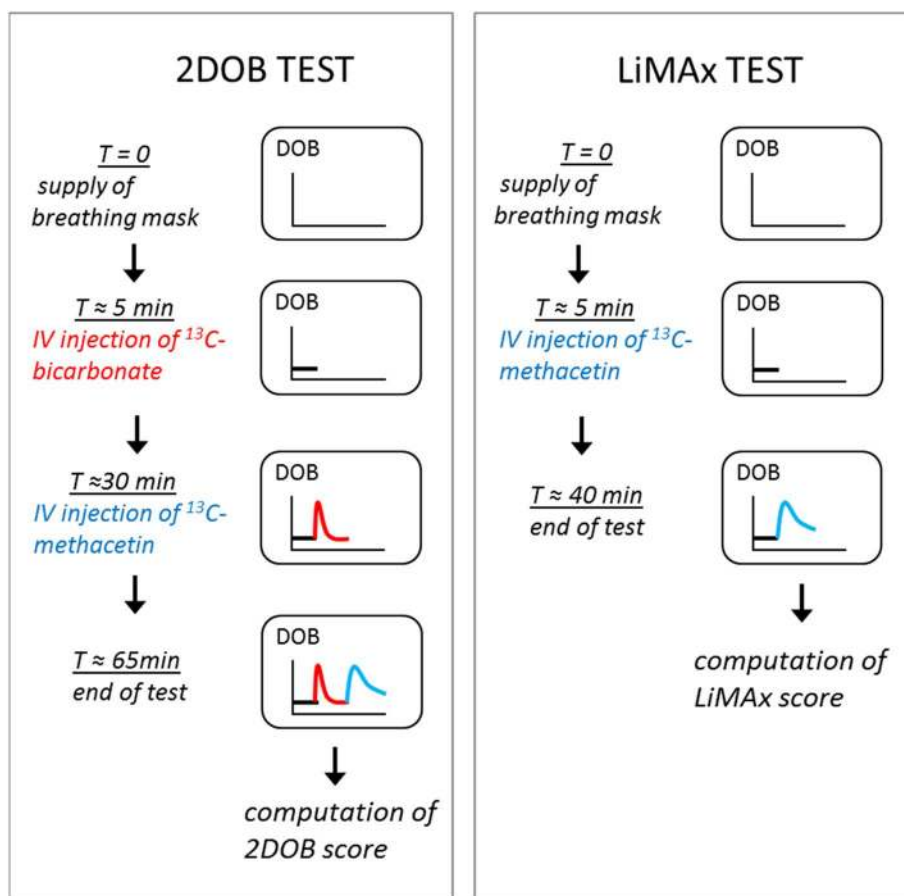
where  $\text{DOB}_{\max}$  is the maximum value of DOB,  $R(\text{PDB}) = 0.011237$  is the baseline ratio  $^{13}\text{CO}_2/^{12}\text{CO}_2$  according to the Pee Dee Belemnite standard (Craig 1957),  $P = 300$  mmol/h/m<sup>2</sup> is the  $\text{CO}_2$  production rate per m<sup>2</sup> body surface, BSA is the body surface area (in m<sup>2</sup>),  $\text{MM} = 166$  is the molecular mass of the test drug  $^{13}\text{C}$ -methacetin and BW denotes the body weight. BSA was computed by the formula  $\text{BSA} = 0.024256 \times \text{BW}^{0.5378} \times H^{0.3964}$  [m<sup>2</sup>], where  $H$  is the height [in m] of the subject (Haycock et al. 1978).

The data of the LiMAx test were also analyzed by means of the reduced compartment model (represented by the non-shaded part in the reaction scheme shown in Fig. 2) in which the kinetic parameters for the exchange of  $\text{CO}_2$ /bicarbonate with body compartments were not included (see Fig. 1).

### Targeted multiple reaction monitoring (MRM) of methacetine and acetaminophene in the plasma

Pure metabolites were dissolved to a final concentration of 1  $\mu\text{M}$  in MeOH and 0.1% formic acid, and injected by a syringe (7  $\mu\text{L}/\text{min}$ ) into the triple quadrupole hybrid ion trap mass spectrometer (QTrap 6500, Sciex, Framingham, MA). Precursor ions were fragmented in positive and negative electrospray ionization (ESI) modes, and the most intense fragment peaks were chosen and optimized for the following parameters: declustering potential (DP) for precursor ions, collision energy (CE), and collision cell exit potential (CXP) for fragment ions. Transitions were monitored and acquired at unit resolution (peak width at 50% was  $0.7 \pm 0.1$  Da tolerance) in quadrupole Q1 and Q3. Metabolites were separated on a Reprosil-PUR C18-AQ (1.9  $\mu\text{m}$ , 120  $\text{\AA}$ , 150  $\times$  2 mm ID;

**Fig. 1** Schematic representation of the procedures of the 2DOB test and the LiMAx test



Dr. Maisch; Ammerbuch, Germany) column at a controlled temperature of 30 °C, the four best transitions were chosen and all MRM instrument settings are given in Table 1.

For the standard curve, 50  $\mu\text{L}$  of plasma was mixed with 20  $\mu\text{L}$  of a dilution series of acetaminophen (paracetamol) and methacetin of following concentrations (800 nM, 200 nM, 50 nM, 12.5 nM, 3.125 nM, 781.25 pM, 195.31 pM, 48.83 pM, 12.21 pM, 3.05 pM) and 20  $\mu\text{L}$  of internal standard. 180  $\mu\text{L}$  of acetonitrile was added to precipitate plasma proteins, vortexed for 20 s and incubated

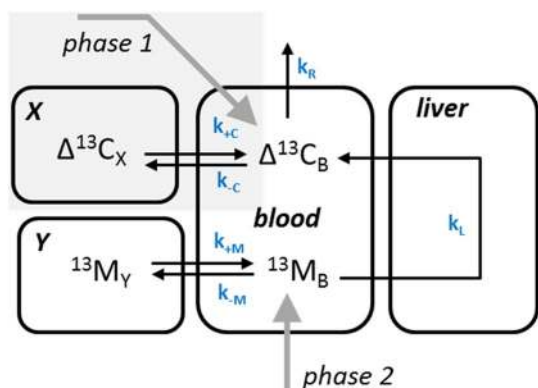
at room temperature for 10 min. Thereafter, the mixture was centrifuged for at 16,000 rcf for 10 min. 100  $\mu\text{L}$  of the supernatant was taken and mixed with 44.4  $\mu\text{L}$  of HPLC grade water. 5  $\mu\text{L}$  of the mixture was finally injected into the LC–MS system. The standard curve was re-measured after the acquisition of two complete time series. The standard curve was acquired interlaced with the measurement of samples.

Blood samples were handled as described above with the following modification: 50  $\mu\text{L}$  of plasma was mixed with

**Table 1** Statistical measures of the predictive power of the LiMAx test and parameters of the 2DOB test

	LiMAx	$k_L$	$M_L$	2DOB score	
				$k_L$	$M_L$
Classifier threshold value	363	2.93E–02	61	5.16E–02	61
True positive rate (TP)	0.75	0.75	0.94	0.94	
False positive rate (FP)	0.18	0.14	0.32	0.09	
True negative rate (TN)	0.82	0.86	0.68	0.91	
False negative rate (FN)	0.25	0.25	0.06	0.06	
Positive predictive value (PPV)	0.75	0.80	0.68	0.88	
Negative predictive value (NPV)	0.82	0.83	0.94	0.95	
Area under the ROC curve (AUC)	0.82	0.85	0.81	0.92	

The classification of the individual subjects is given in Supplementary Table 1



**Fig. 2** Scheme of the compartment model. The variable  $\Delta^{13}C$  denotes the sum of additionally (above baseline!) formed  $^{13}CO_2$  and  $H^{13}CO_3$ ,  $^{13}M$  denotes  $^{13}C$ -methacetin. Both metabolites can be exchanged between the plasma and other body compartments lumped together into single compartments (X and Y, respectively). Uptake of  $^{13}C$ -methacetin into the liver, conversion to the reaction products acetaminophen and  $^{13}CO_2$  by CYP1A2, and release of the enzymatically formed  $^{13}CO_2$  into the blood plasma are described by an overall reaction with rate constant  $k_L$ . Phase 1 starts with the intravenous administration von  $^{13}C$ -labeled bicarbonate. Phase 2 starts about 30 min later with the intravenous administration von  $^{13}C$ -labeled methacetin (see test protocols)

20  $\mu$ L of internal standard and 20  $\mu$ L of HPLC grade water. Each sample was injected in triplicates.

The samples (5  $\mu$ L) were injected and compounds were separated on a LC instrument (1290 series UHPLC; Agilent, Santa Clara, CA), coupled online to a triple quadrupole hybrid ion trap mass spectrometer QTrap 6500 (Sciex). Data acquisition was performed with an ion spray voltage of 5.5 kV in the positive mode for the electro spray ionization source, nitrogen as the collision gas was set to medium, the curtain gas was at 30 psi, the ion source gas 1 and 2 was at 50 and 70 psi, respectively. The interface heater temperature was set to 350  $^{\circ}C$ .

The LC buffer compositions were as follows: (A) 10 mM ammonium acetate in LC–MS grade  $H_2O$  (adjusted with acetic acid to pH 3.5), (B) LC–MS grade acetonitrile with 0.1% formic acid. The 10 min gradient was set as follows: 2% B at minutes 0–1, linearly increasing to 45% until minute 5, 98% between minutes 6 and 7, and again equilibrated at 2% between minutes 7.1–10.

Metabolite identification was based on three levels: (i) the correct retention time, (ii) four transitions, (iii) and matching MRM ion ratios of tuned pure metabolites as described previously (Lau and Ahmad 2013). Peak integration was performed using MultiQuant<sup>TM</sup> software v.2.1.1 (Sciex). All peaks were reviewed manually and adjusted if necessary. The peak area of the first transition per metabolite was used for subsequent calculations. Peak integrations were reviewed manually. An internal standard ( $^{13}C_9^{15}N$  L-phenylalanine) was used to normalize all LC–MS runs for instrumental

variations. For the calculation of absolute concentrations, the averaged intensities of three replicates (samples and standard curve) and the two nearest values of the standard curve were taken and calculated as a linear interpolation of values on a log–log plot.

### Compartment modeling of systemic $^{13}CO_2$ and $^{13}C$ -methacetin kinetics

We used the compartment model shown in Fig. 2 to describe the kinetics of  $^{13}C$ -methacetin (abbreviated with  $^{13}M$ ) and the additionally formed  $H^{13}CO_3 + ^{13}CO_2$  (abbreviated with  $\Delta^{13}C$ ) resulting from the administration of  $^{13}C$ -labeled bicarbonate and enzymatic conversion of  $^{13}C$ -methacetin. Combining  $^{13}CO_2$  and  $H^{13}CO_3$  into one variable is backed up by the fact that both entities are in quasi-equilibrium due to the fast reaction catalyzed by the enzyme carboanhydrase. The model takes into account the exchange of  $^{13}CO_2 + H^{13}CO_3$  and  $^{13}C$ -methacetin between the blood and other body compartments (denoted with X and Y, respectively). Earlier studies of Barstow et al. (1985, 1990) and Irving et al. (1983) on the  $^{13}C$ -bicarbonate kinetics in humans have provided evidence for the existence of at least three  $CO_2$  exchange compartments differing in the characteristic time constants for the exchange kinetics. Over shorter time ranges of about 15–30 min as considered in our test, the error made by neglecting the slow exchange processes should remain sufficiently small.

The reaction scheme shown in Fig. 2 differs from the reaction scheme used in our previous work (Holzhutter et al. 2013) in the following items: (i) the reversible exchange of the drug with non-hepatic compartments is now taken into account, (ii) the kinetics of the reaction product acetaminophen is not considered in the model as the plasma concentration of this drug was not monitored, (iii) exchange of  $^{13}C$ -methacetin with the liver, conversion to the reaction products acetaminophen and  $^{13}CO_2$  and release of the formed  $^{13}CO_2$  into the blood plasma are lumped together into a single process with rate constant  $k_L$ . The latter simplification was necessary as the reversible exchange of  $^{13}C$ -methacetin with the liver and with extra-hepatic organs cannot be discerned in the shape and magnitude of the DOB curve. Note that the parameter  $k_R$  represents the total elimination rate of methacetin from the plasma, including the release of the drug from the plasma into the breath as well as other modes of irreversible loss as, for example, renal excretion and slow covalent fixation in organic molecules.

### Kinetic equations

The reaction scheme shown in Fig. 1 is governed by the following set of ordinary differential equations:

$$\begin{aligned}
 \frac{d\ ^{13}\text{M}_\text{B}}{dt} &= k_{+\text{M}}\ ^{13}\text{M}_\text{Y}^* - (k_{-\text{M}} + k_\text{L})\ ^{13}\text{M}_\text{B}, \\
 \frac{d\ ^{13}\text{M}_\text{Y}^*}{dt} &= -k_{+\text{M}}\ ^{13}\text{M}_\text{Y}^* + k_{-\text{M}}\ ^{13}\text{M}_\text{B}, \\
 \frac{d\ \Delta^{13}\text{C}_\text{B}}{dt} &= k_\text{L}\ ^{13}\text{M}_\text{B} + k_{+\text{C}}\ \Delta^{13}\text{C}_\text{X}^* - (k_{-\text{C}} + k_\text{R})\ \Delta^{13}\text{C}_\text{B}, \\
 \frac{d\ \Delta^{13}\text{C}_\text{X}^*}{dt} &= -k_{+\text{C}}\ \Delta^{13}\text{C}_\text{X}^* + k_{-\text{C}}\ \Delta^{13}\text{C}_\text{B}.
 \end{aligned}
 \tag{2}$$

For the meaning of the rate constants in equation system (2), see legend to Fig. 1. Note that the variables  $\Delta^{13}\text{C}_\text{X}^*$  and  $^{13}\text{M}_\text{Y}^*$  represent effective concentrations in the compartments X and Y that are related to the true concentrations  $C_\text{X}$  and  $M_\text{Y}$  by the relations  $C_\text{X} = \frac{\Omega_\text{B}}{\Omega_\text{X}} C_\text{X}^*$  and  $M_\text{Y} = \frac{\Omega_\text{B}}{\Omega_\text{Y}} M_\text{Y}^*$  with  $\Omega$  denoting the (unknown) volume of the respective compartment. For the numerical integration of equation system (2) and estimation of numerical parameter see the Supplementary Information (Computational Details).

### Definition of a scoring function for the quantification of hepatic detoxification capacity

Different combinations of model parameters were tried to define an appropriate scoring function that reliably quantifies the hepatic detoxification capacity. The power of a scoring function to discriminate between the detoxification capacities of healthy and diseased livers was evaluated by means of a ROC (receiver operator characteristic) curve analysis. The area under the ROC curve (AUC) is an effective and widely used method for evaluating the discriminating power of a diagnostic test or statistical model. Threshold values of model parameters used in the scoring function were chosen such that the Youden index,  $Y = \text{TP} - (1 - \text{FP}) - 1$  (TP—true positive rate = fraction of ill subjects correctly classified, FP—false positive rate = fraction of healthy subjects falsely classified) becomes maximum. Assessment of the statistical significance of the difference between any two classifiers based on the AUC was performed by a non-parametric test that accounts for the correlation of the ROC curves (Vergara et al. 2008). Sensitivity of classification results against random variations of the training set of healthy and diseased subjects was checked by means of Bootstrap resampling (Mossman 1995) based synthetic data sets created by random selection (with permitted repetition) of original data.

## Results

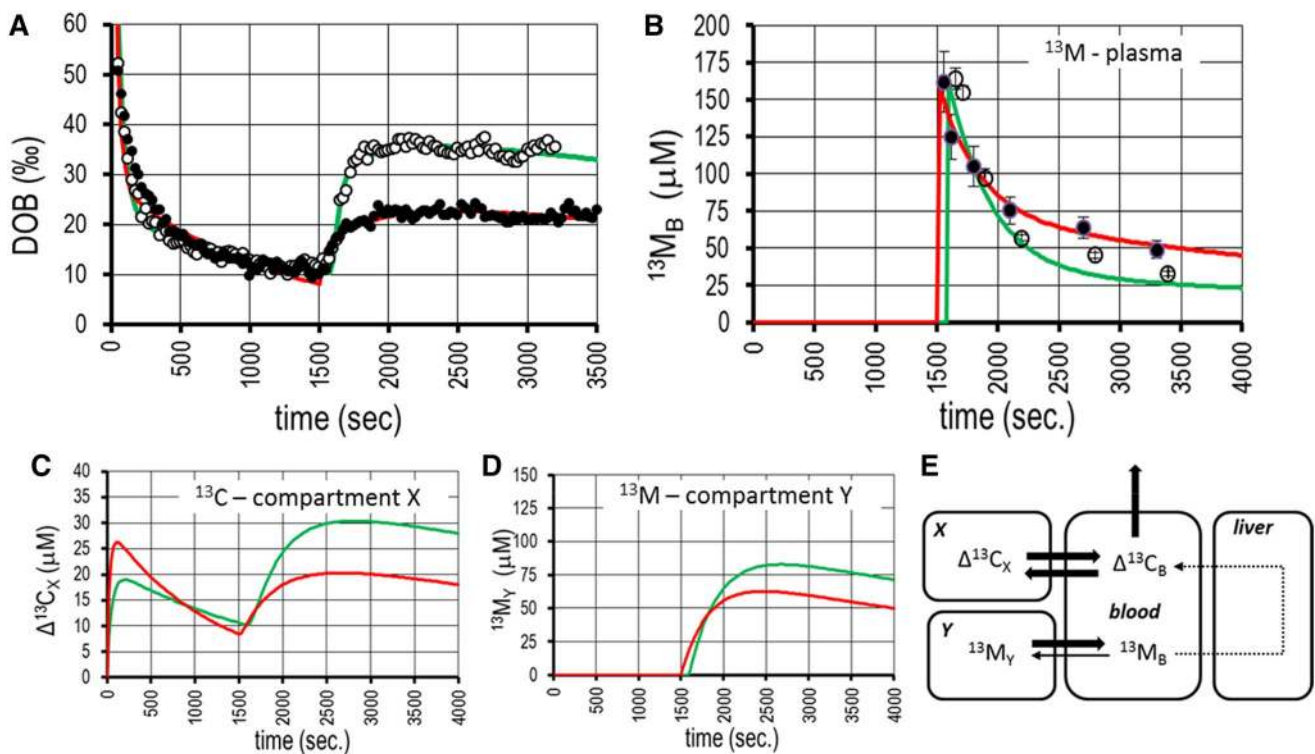
### Adjustment of the compartment model to measured DOB data

Both variants of the methacetin breath test, the conventional LiMAX test initiated by the injection of the test drug and subsequent monitoring of  $^{13}\text{CO}_2$  in the breath, and the proposed novel 2DOB test were carried out in 38 subjects (for test procedures see Fig. 1). Figure 3 shows typical DOB curves of 2DOB test obtained in a liver-healthy subject and a patient with liver cirrhosis. Experimental and computed DOB curves of all subjects are given in Fig. 1 of the Supplementary Information.

The DOB curve data were used to parameterize the compartmental model, i.e. to estimate numerical values for the model parameters (see Supplementary Information). Model fitting yielded numerical estimates for the rate constants  $k_{+\text{C}}$ ,  $k_{-\text{C}}$  of  $\Delta^{13}\text{C}$  exchange with body compartment X, the rate constant  $k_\text{R}$  of irreversible  $\Delta^{13}\text{C}$  elimination, the rate constants  $k_{+\text{M}}$ ,  $k_{-\text{M}}$  of  $^{13}\text{M}$  exchange with body compartment Y and the rate constant  $k_\text{L}$  for the hepatic detoxification of  $^{13}\text{M}$ . The two-step procedure applied for fitting subsets of model parameters separately to the first and second phase of the DOB curve was chosen as some parameters have a partially redundant influence on the height and shape of the DOB curve. For example, an increase of DOB values can be elicited by an increase of  $k_\text{L}$  or a decrease of the  $\text{CO}_2$  elimination rate  $k_\text{R}$ . The sensitivity analysis in the Supplement “Computational Details” illustrates that estimation of numerical parameter values from different parts of the 2DOB curve is necessary for unequivocal estimates of the parameters  $k_{+\text{M}}$  and  $k_{-\text{M}}$ .

From the DOB data of the LiMAX test, only for the four model parameters,  $k_\text{R}$ ,  $k_{+\text{M}}$ ,  $k_{-\text{M}}$  and  $k_\text{L}$ , numerical estimates could be derived from the data. Supplementary Table 1 depicts for each subject the numerical values of all model parameters.

Once parameterized, the model allows simulating the time course of  $\Delta^{13}\text{C}$  and  $^{13}\text{M}$  in the compartments X and Y, respectively. An example is shown in Fig. 3. Compared with the healthy subject (HS), the patient with cirrhosis (DS) displays a higher uptake rate of  $\Delta^{13}\text{C}$  into body compartment X after administration of an identical bolus of  $^{13}\text{C}$ -bicarbonate. Nevertheless, the rise of  $\Delta^{13}\text{C}$  in compartment X of DS is lower owing to the significantly lower detoxification rate of  $^{13}\text{M}$  and correspondingly lower liberation rate of  $\Delta^{13}\text{C}$  into the plasma (mirrored by the lower maximum of the DOB curve). Note that 40 min after administration of the test drug, a substantial portion of  $\Delta^{13}\text{C}$  is still retained in body compartment X. The model simulation predicts a faster decline of plasma  $^{13}\text{M}$  in HS



**Fig. 3** Patient-specific kinetics of  $^{13}\text{C}$ -labeled  $\text{CO}_2$  and methacetin in a healthy subject (HS) and a subject with liver cirrhosis (DS). **a** DOB curve data of HS (open circles, green line) and DS (filled circles, red line). The lines represent theoretical data obtained by fitting the compartment model to the measured DOB data. Over the full time course of the experiment, the DOB values were monitored every 25 s. After a calibration phase of about 600 s, at time "0" a dose of 2 mg/kg body weight of  $^{13}\text{C}$ -bicarbonate was injected intravenously. At time point  $T_M$  (indicated by the arrow), a dose of 2 mg/kg body weight of  $^{13}\text{C}$ -methacetin was injected intravenously. **b** Simulated and measured time course of plasma methacetin of the normal subject (green curve,

open circles) and the patient with liver cirrhosis (red curve, closed circles). **c** Simulated time course of additionally formed  $^{13}\text{C}$ - $\text{CO}_2$ /bicarbonate ( $\Delta^{13}\text{C}_X^*$ ) in body compartment X. Note that relative concentrations are shown which are linearly related to the true concentrations by a scaling factor that depends on the unknown volume of compartment X. **d** Simulated time course of  $^{13}\text{C}$ -methacetin ( $^{13}\text{M}_Y^*$ ) in body compartment Y. Note that relative concentrations are shown which are linearly related to the true concentrations by a scaling factor that depends on the unknown volume of compartment Y. **e** Flux changes in DS compared to HS. Bold arrow: increased flux, dotted arrow: decreased flux

compared with DS. This is a consequence of DS having both a lower detoxification rate of  $^{13}\text{M}$  and a lower uptake rate of  $^{13}\text{M}$  into compartment Y (Fig. 2d, e). We validated the model predictions by the good concordance between the theoretical curves and measurements of  $^{13}\text{M}$  at various time points after  $^{13}\text{M}$  administration (Fig. 2b). It has to be mentioned that for subjects with very low values of the detoxification rate constant  $k_L$ , there was a tendency of the computed time course of plasma methacetin to underestimate the true (= experimentally determined) clearance of  $^{13}\text{M}$  from the plasma (see Fig. 2, Supplementary Information). This discrepancy is due to the fact that with decreasing rise of the DOB curve after  $^{13}\text{M}$  administration, the two parameters  $k_{+M}$  and  $k_{-M}$  for the exchange of  $^{13}\text{C}$ -methacetin with compartment Y become less identifiable from the DOB curve (see Supplement "Computational details", Parameter Identifiability). At the extreme, if there is no rise of the DOB curve at all, indicating that detoxification of the drug is absent ( $k_L = 0$ ), the numerical values of  $k_{+M}$

and  $k_{-M}$  cannot be determined as the reversible exchange of  $^{13}\text{C}$ -methacetin with compartment Y is not connected with the  $^{13}\text{CO}_2$  kinetics.

### Influence of systemic $\text{CO}_2$ kinetics

The values of the model parameters  $k_{+C}$ ,  $k_{-C}$  and  $k_R$  derived from the first part of the DOB curve provide information on the systemic  $\text{CO}_2/\text{HCO}_3$  kinetics of the subject. The numerical estimates for the  $\Delta^{13}\text{C}$  elimination rate  $k_R$  varied between 0.085 and 0.40/min, the mean value across the 38 subjects amounts to 0.22/min. For a subject with a normal acid–base status of 25 mM plasma  $\text{CO}_2$  and a blood volume of 5 L, this translates into an average diurnal elimination rate of about 38 mol/day. This value is above the known resting value of about 20 mol/day for the respiratory  $\text{CO}_2$  elimination rate (Jonderko et al. 2008). Hence, on the average, about one-half of the eliminated  $\Delta^{13}\text{C}$  must be rapidly taken up by a "fast" body compartment in a quasi-irreversible manner such that

the slow return of this fraction to the plasma is not observable within the time span of the test. This is also reflected in an average  $^{13}\text{CO}_2$  recovery rate of 54% in the exhaled breath at  $T=15$  min after  $^{13}\text{C}$ -bicarbonate administration (for individual recovery rates see Fig. 4, Supplementary Information).

Notably, the distribution of values for the  $\text{CO}_2$  elimination rate  $k_R$  in the group of patients with a liver disease exhibits a general shift towards small values (see Fig. 3, Supplementary Information). Six liver patients had  $k_L$  values smaller than  $0.17 \text{ min}^{-1}$  ( $\approx 66\%$  of the mean). Nevertheless, there was no statistically significant difference between the mean  $k_L$  values in the group of healthy subjects ( $0.222/\text{min}$ ) and diseased subjects ( $0.220/\text{min}$ ) according to the two-sample  $t$  test ( $p=0.77$  at  $\alpha=0.05$ ).

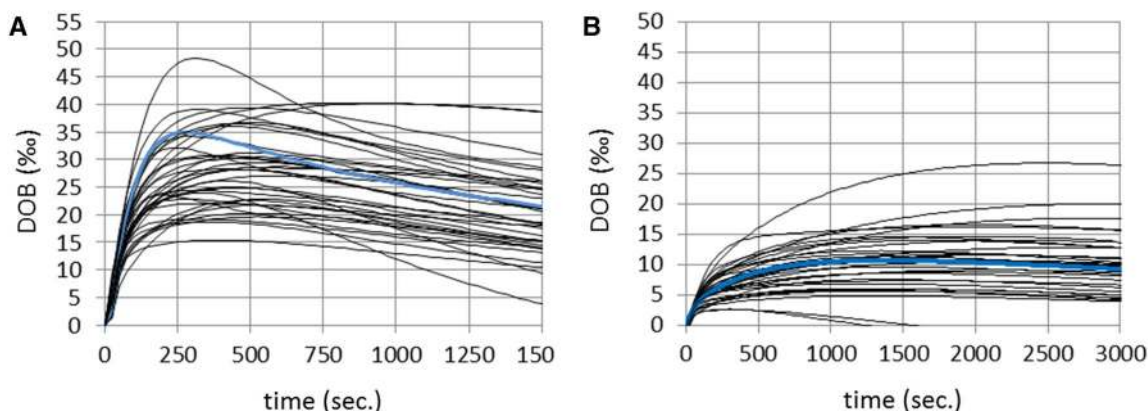
Figure 4 illustrates how individual variations of the systemic  $\text{CO}_2$  kinetics may influence the shape of the DOB curve at fixed  $^{13}\text{C}$ -methacetin kinetics. The curves were simulated with fixed values of the kinetic parameters  $k_L$ ,  $k_{+M}$ ,  $k_{-M}$  determining the methacetin kinetics, but varying values of the kinetic parameter  $k_{+C}$ ,  $k_{-C}$ ,  $k_R$  obtained for the 38 subjects of this study. The resulting curves differ substantially in characteristic curve parameters such as peak value, time-to-peak value, area under the curve and steepness of decline. For subject #33 with normal liver function, the peak values of the DOB curves (see Fig. 4a) may differ from the "true" peak value by more than a factor of 2. Extreme DOB curves would be produced if subject #33 had the  $\text{CO}_2$  kinetics of subject #23 (maximal positive deviation) or subject #28 (maximal negative deviation). The coefficient of variation

(CV = standard deviation/mean) of peak values is 0.25. For subject #37 with liver cirrhosis taken as reference, the spread of DOB curves is generally lower than for the normal subject (CV of peak values = 0.14). For the majority of DOB curves, the peak values at  $T=300$  s. differ less than a factor of 1.5 from the reference curve.

This suggests that the likelihood of overestimating the true drug-detoxifying capacity owing to variations of the  $\text{CO}_2$  kinetics is lower for a subject with a functionally impaired liver. The higher risk of over- or underestimating the 'true' hepatic detoxification capacity in subjects with normal or slightly impaired liver functionality compared with subjects with severe liver dysfunction is also reflected in the relationship between the detoxification rate constant  $k_L$  and the parameter LiMAX serving as predictor of the detoxification capacity in the conventional LiMAX test and representing up to a scaling factor the peak value of the DOB curve (see Fig. 5). Despite the highly significant overall correlation of these two parameters, the largest deviations from the hyperbolic regression line occur in the middle range of  $k_L$  values (grey-shaded area in Fig. 5) indicating either normal or mildly impaired detoxification capacity.

### The predictive capacity of model parameters as disease classifiers

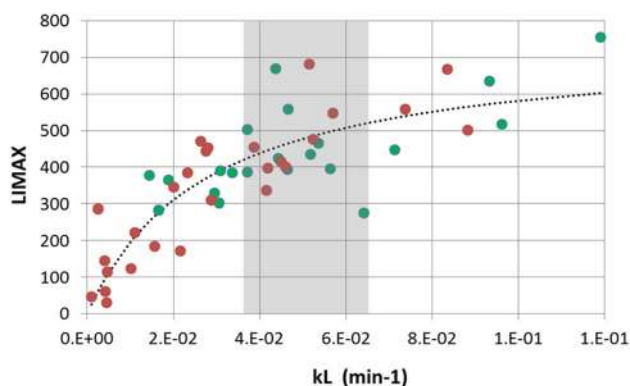
Next we studied the suitability of model parameters for the binary classification of the subjects liver into normal (disease class "1") or impaired (disease class "2"). Obviously, the model parameter  $k_L$  quantifying the hepatic detoxification



**Fig. 4** Simulated DOB curves at fixed methacetin kinetics, but variable  $\text{CO}_2$  kinetics. **a** The kinetic parameters determining the methacetin kinetics were fixed at those values obtained for subject #33 with normal liver function:  $k_L=0.039 \text{ min}^{-1}$ ,  $k_{+M}=0.036 \text{ min}^{-1}$ ,  $k_{-M}=0.082 \text{ min}^{-1}$ . **b** The kinetic parameters determining the methacetin kinetics were fixed at those values obtained for subject #37 with liver cirrhosis:  $k_L=0.010 \text{ min}^{-1}$ ,  $k_{+M}=0.131 \text{ min}^{-1}$ ,  $k_{-M}=0.06 \text{ min}^{-1}$ . The set of parameters,  $k_{+C}$ ,  $k_{-C}$ ,  $k_R$ , determining the  $\text{CO}_2$  kinetics were put to the values obtained for the 38 subjects, i.e. each curve represents the DOB curve of a subject comprising the

methacetin kinetics of subject #9 (**a**) or subject #37 (**b**) but the  $\text{CO}_2$  kinetics of one of the subjects 1–38. Note that the curves represent the second part of the 2DOB curves starting with the administration of  $^{13}\text{C}$ -methacetin, i.e. time "0" corresponds to time  $T_m$  of the full test curve. For a better comparison of curves, the residual DOB value at  $t=T_m$  was subtracted so that the initial DOB values at time=0 is zero for all curves. Blue curves: true DOB curve of the reference subject (subject #33 in **A**, subject #37 in **b**). Coefficient of variation (CV) of maximum DOB values: =0.25 (**a**), 0.14 (**b**)





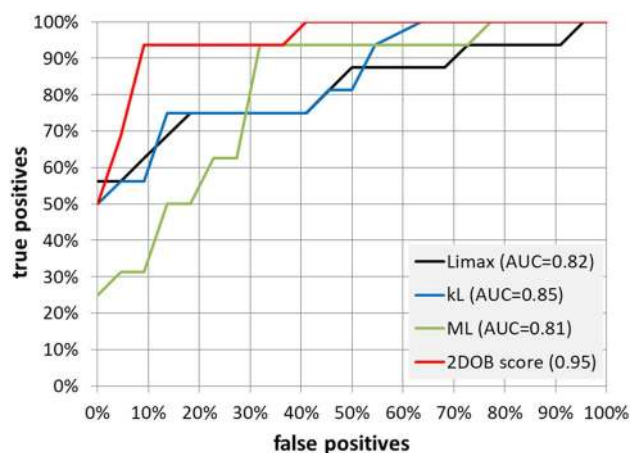
**Fig. 5** LiMAX versus model parameter  $k_L$  (detoxification rate). LiMAX values were determined from the DOB curves of the conventional LiMAX test according to Eq. (1) (data given in Supplementary Table 1). Green and red dots indicate data for controls and patients. The grey-shaded region highlights the range of largest deviations between  $k_L$  and LiMAX. The dotted curve represents the best-fit hyperbolic relationship between LiMAX and  $k_L$ :  $\text{LiMAX} = 744 \times k_L / (0.028 + k_L)$

rate of the test drug should serve as an appropriate classifier. Testing the predictive capacity of  $k_L$  by means of a receiver operating characteristics (ROC) yielded a value of  $\text{AUC} = 0.85$  for the area under the curve (AUC), a true positive rate (TP) of 0.75 and a true negative rate (TN) of 0.86. Thus, the predictive power achieved with the parameter  $k_L$  was not better than the predictive power of the LiMAX score ( $\text{AUC} = 0.82$ ,  $\text{TP} = 0.75$ ,  $\text{TN} = 0.86$ ). Notably, both  $k_L$  and the LiMAX yielded false negative classifications for the same group of subjects. This liver suggests that a chronic liver disease must not be necessarily paralleled by a significantly lowered chemical conversion rate of the test drug.

We then tested whether other model parameters may serve as possible disease predictors. Intriguingly, parameters  $k_{-M}$  describing the initial rate of methacetin uptake into the storage compartment Y yielded a surprisingly high quality of disease classification ( $\text{AUC} = 0.72$ ). Parameter  $k_{+M}$  describing the release rate of methacetin from compartment Y also yielded a statistically significant classification with  $\text{AUC} = 0.65$ . Hence the storage capacity of compartment Y for methacetin appears to be reduced in diseased livers. As the storage capacity of compartment Y depends on both  $k_{+M}$  and  $k_{-M}$ , we used the average concentration of methacetin stored in compartment Y over a time span  $T_S = 3000$  s. after administration of  $^{13}\text{M}$  at time  $t = T_M$ , (see Fig. 6)

$$M_L = \frac{1}{3000} \int_{T_M}^{T_M+T_S} {}^{13}M_Y^*(t) dt, \tag{3}$$

as a new disease classifier yielding an AUC value of 0.81. Remarkably, the ROC curve associated with the measure  $M_L$



**Fig. 6** Receiver operating characteristic to assess the predictive power of the 2DOD and LiMAX test. Black curve: LiMAX value used as classifier ( $\text{AUC} = 0.82$ ). Blue curve: parameter  $k_L$  used as classifier ( $\text{AUC} = 0.85$ ). Green curve: parameter  $M_L$  defined by equ (5) used as classifier ( $\text{AUC} = 0.81$ ). Red curve: 2DOB score defined by equ (6) used as classifier ( $\text{AUC} = 0.95$ )

reached already a true positive rate of more than 90% at a false positive rate of about 30%, i.e.  $M_L$  is predestined more than all other measures tested to reliably identify patients with liver disease.

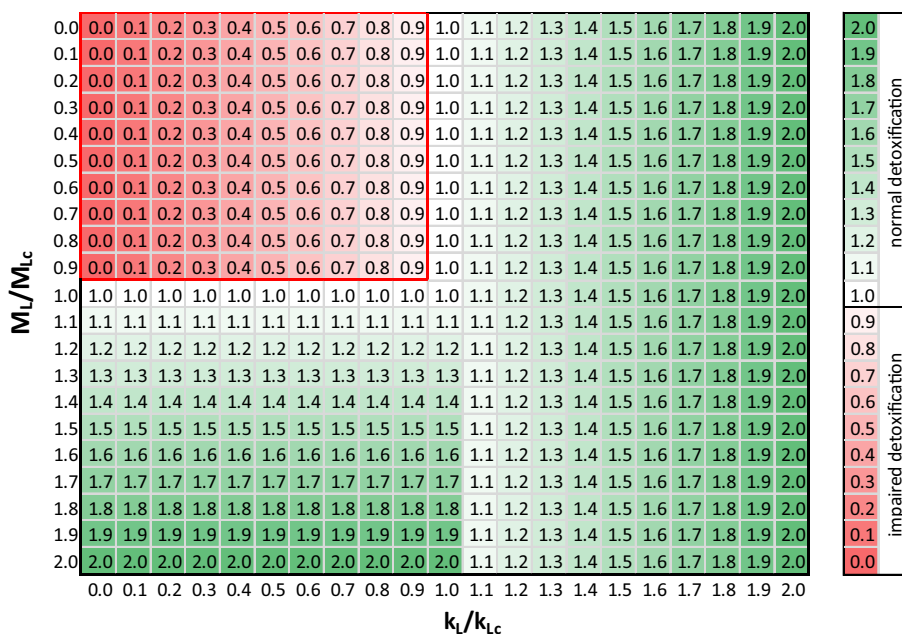
We then examined whether a combination of the two disease predictors,  $k_L$  and  $M_L$ , may provide an even better classification than either parameter alone. As there are endless possibilities to combine the model parameters into a scoring function, we followed Occam’s razor command “not to multiply entities without necessity” (Schaffer 2015) and used the most simple (linear) combination of the two model parameters  $k_L$  and  $M_L$  that individually provided the best classification results:

$$\text{2DOB score}(k_L, M_L; k_{Lc}, M_{Lc}) = \begin{cases} \frac{M_L}{M_{Lc}} & \text{if } k_L \leq k_{Lc} \text{ and } M_L \geq M_{Lc} \\ \frac{k_L}{k_{Lc}} & \text{else} \end{cases} \tag{4}$$

The 2DOB score defined in Eq. (4) is a continuous measure of the hepatic detoxifications capacity, whereby values smaller than unity indicate an impaired liver function. The 2DOB score is basically identical with the detoxification rate  $k_L$  except for those cases where  $k_L$  remains below the threshold value  $k_{Lc}$  (indicating an impaired detoxification rate) whereas the storage capacity  $M_L$  is larger than the threshold value  $M_{Lc}$ . This definition of the 2DOB score implicates that the liver is classified as impaired if and only if both parameters,  $k_L$  and  $M_L$ , are smaller than the cut-off values  $k_{Lc}$  and  $M_{Lc}$  (see Fig. 7).

With the 2DOB score used as classifier, indeed a significant improvement of the classification quality was achieved;

**Fig. 7** 2DOB score as function of the normalized (parameters  $k_L$  and  $M_L$ ). Parameters are normalized with respect to their cut-off values, i.e.  $k_L/k_{LC} \geq 1$  means  $k_L \geq k_{LC}$ . The red lines enclose the region of score values that are smaller than unity thus indicating an impaired hepatic detoxification capacity. Note that according to definition (6) of the 2DOB score, small values of the detoxification rate  $k_L$  below the threshold  $k_{LC}$  yield nevertheless a 2DOB score larger than 1 (=liver healthy) if the value of the storage capacity exceed the threshold value ( $M_L \geq M_{LC}$ )



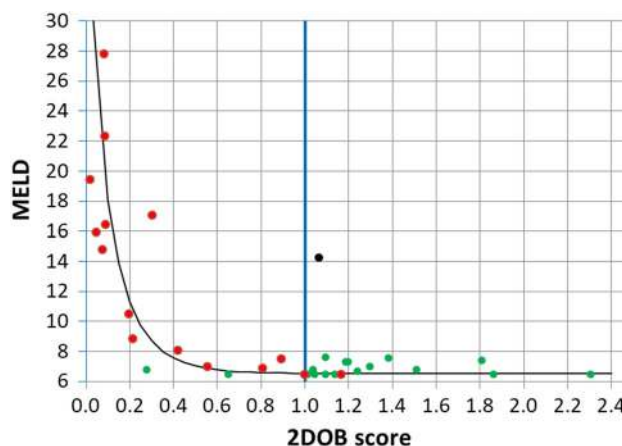
the  $AUC_{2DOBscore} = 0.95$  is significantly larger than the  $AUC_{LiMAX} = 0.82$  ( $z = 1.9, p = 0.028$ ).

The cut-off values  $k_{LC}$  and  $M_{LC}$  given in Table 1 were determined by maximizing the AUC value across the group of 38 subjects.

**Relationship between the 2DOB score and the MELD score**

As the 2DOB score is defined as a continuous measure of the hepatic detoxifications capacity, it was tempting to check whether the 2DOB score correlates with clinically applied parameters that are used in the clinics for assessing the severity of chronic liver diseases. Figure 8 reveals a high correlation between the 2DOB score and the MELD (Model for End-Stage Liver Diseases) core of the 38 test subjects ( $R_{Spearman} = 0.92, p < 0.00001$ ).

There exists a clear monotone and highly non-linear relationship between the 2DOB score and the MELD score. 2DOB scores smaller than 0.4 are consistently associated with MELD scores larger than 8. The only outlier occurred for subject #33 (black data point in Fig. 8) who has a clinically inconspicuous liver, correctly reflected by a 2DOB score  $> 1$ , but an ordinarily high MELD score of 14 due to the fact that this subject was under steady treatment with anticoagulants which are known to influence the MELD score. As shown in Fig. 8, the 2DOB score proves as a much more sensitive indicator of a beginning or moderate liver disease than the MELD score. That is no surprise as the MELD was developed and adapted as a



**Fig. 8** Relationship between the 2DOB score and the MELD score of the 38 test subjects. Green dots: normal subjects. Red dots: subjects with clinically diagnosed liver disease. The black line is intended to guide the eye and to underline the non-linear relationship. The vertical blue lines at 2DOB score=1 mark the threshold between normal and impaired detoxification capacity

prognostic tool in advanced liver diseases (Lau and Ahmad 2013).

**Robustness of the 2DOB test and further test optimization**

**Age-differences between controls and patients**

The subjects in the control group of our study were significantly younger (mean age = 42.4 years) than the subjects with liver disease (mean age = 55.1 years). Aging has been

reported to cause moderate declines in the Phase I metabolism of certain drugs (Cieslak et al. 2016; Schmucker 2005; Schmucker and Wang 1980) and thus could have affected the classifications. Therefore, we split the group of the 22 control subjects into 2 sub-groups, comprising 11 healthy subjects aged from 22 to 49 (mean age = 30.7) and 11 healthy subjects aged from 50 to 59 (mean age = 54.5). The older control group matches approximately the age of the diseased subjects. There was a slight albeit statistically insignificant tendency towards higher LiMAx scores (from 428 to 458) and  $k_L$  values (from 0.044 to 0.054), suggesting a marginal age-related increase of CYP1A2-dependent metabolization rates in normal livers. Interestingly, the mean value of the parameter  $M_L$  dropped from 62.5 in the group of younger controls to 50.7 in the group of older controls, suggesting an age-dependent decline of the storage capacity of the liver-associated compartment Y.

The classifications performed using the two age-corrected control groups separately (see Table 3, Supplementary Information), yielded identical statistical measures for the 2DOB test. Of note, the cut-off values for  $k_L$  and  $M_L$  were identical irrespective of the control group used. Changes in the ROC statistics of the LiMAx test were also very little affected by the choice of the control group. However, the cut-off value separating normal and diseased livers was significantly affected (282 versus 363) when using either the younger or older controls as reference.

We also tested the robustness of LiMAx- and 2DOB-based classifications against random variations of the training set of healthy and diseased subjects using the method of Bootstrap resampling (Mossman 1995). The frequency distribution of AUC values shown in Fig. 9 was constructed by choosing randomly with replacing 38 subjects from the original set, i.e. the randomly sampled set contained some data sets more than once (overrepresented) but some data sets were missing. For this synthetic data set, values of AUC

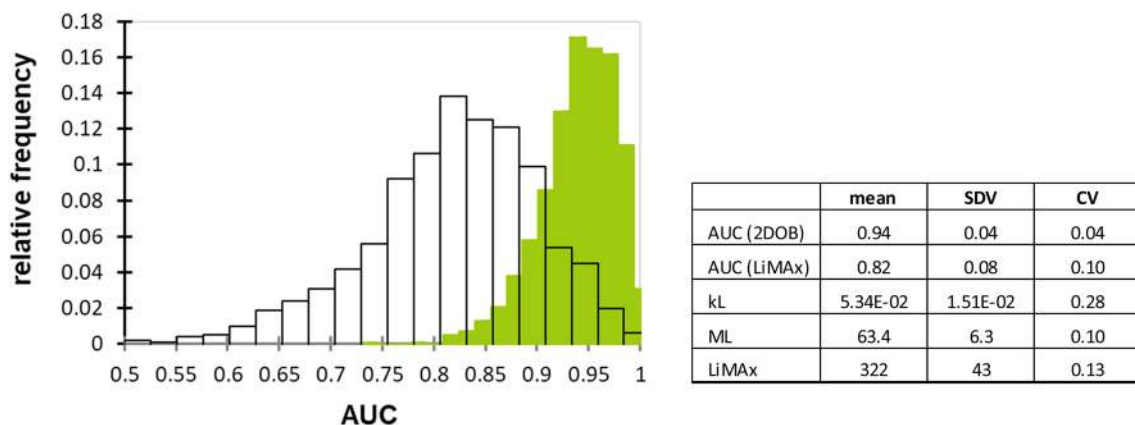
and the cut-off parameters were computed. This resampling was applied 1000 times. The Bootstrap variance of AUC and of the cut-off values are also given in the legend to Fig. 9. In less than 2% of trials, the  $AUC_{2DOB}$  was below 0.8, the mean  $AUC_{2DOB}$  was 0.94.

Regarding the safety and operational feasibility of the test protocol, there were no complications with any participant. Nevertheless, to keep the time interval for the investigation as short as possible, we checked whether the duration of the test can be reduced to about 30 min in total. To this end, we used only a part of the DOB data recorded for 15 min after administration of  $^{13}C$ -bicarbonate and for 15 min after administration of  $^{13}C$ -methacetin for the adjustment of the compartment model and estimation of kinetic parameters. With the short-time 2DOBtest, the 2DOB score yielded almost the same classification quality for the group of 38 subjects ( $AUC = 0.94$ , 4 misclassifications). The results of the 15 min test are given in Supplementary Table 1.

## Discussion

### The need for non-invasive tests capable of detecting early-stage liver diseases

One of the main difficulties with liver diseases is that patients often do not present symptoms or signs until the disease becomes advanced. Around 50% of patients receive their first diagnosis when they arrive in accident and emergency units, typically in their 40 s or 50 s, with jaundice, gastrointestinal bleeding, abdominal swelling and disorientation. Detection of ongoing liver diseases at an early stage is required to prevent or reduce further disease progression by lifestyle changes and pharmacological treatment. Non-invasive diagnostic techniques currently used are serum biomarkers and transient elastography (TE). However, serum



**Fig. 9** Variability of AUC values for the 2DOB and the LiMAx. Variability was assessed by Bootstrap resampling (1000 trials). Green-shaded bars: LiMAx. White bars: 2DOB score

biomarkers are not liver specific and TE results require an expert clinician for interpretation (European Association for Study of Liver and Asociacion Latinoamericana para el Estudio del Hgado 2015). Regarding the use of breath tests for the detection of liver diseases, the elevation in the breath of naturally occurring volatile biomarkers such as limonene, methanol and 2-pentanone has been reported to reliably identify patients liver cirrhosis (Fernandez Del Rio et al. 2015). The same quality of discrimination between healthy and cirrhotic livers has been achieved with the LiMAX test (Buechter et al. 2018). However, differentiating between patients with and without cirrhosis alone has limited value, as this can be performed reliably with routine clinical methods. It is still a challenge to establish non-invasive liver functions tests that are sensitive and specific enough to detect the early onset of restrictions in the metabolic capacity of the liver. Taking for granted that the metabolization of drugs is a reliable indicator for the functional capacity of the liver as a whole, functional breath tests such as the LiMAX are promising test candidates provided that major confounding factors can be eliminated.

### Exhaled $^{13}\text{CO}_2$ as reliable indicator of hepatic detoxification capacity

Using the kinetics of exhaled  $^{13}\text{CO}_2$  produced during hepatic metabolization of a labeled test compound as indicator of the liver's detoxification capacity inevitably raises the problem to what extent this signal is influenced by the systemic distribution of  $\text{CO}_2$ . Measuring the rate of  $\text{CO}_2$  respiration doesn't really solve this problem as this measure does not capture the exchange of plasma  $\text{CO}_2$  with other body compartments. Therefore, we designed a test variant that allows to assess how a defined bolus of  $^{13}\text{CO}_2$  is eliminated from the plasma of the patient and to exploit this information for a more reliable assessment of the drug detoxification rate. To this end, we used a fit-for-purpose compartment model that condenses the numerous physiological processes involved in the hepatic uptake and metabolization of the drug and the systemic distribution of  $\text{CO}_2$  into a manageable number of phenomenological parameters. The model provided an excellent fit to measured DOB curve data of all 38 subjects included in this study (see Fig. 1, Supplementary Information). Based on the high variability of the individual kinetic parameters for systemic  $\text{CO}_2$  kinetics (illustrated in Fig. 4), one may expect variations in the maximum of the DOB curve by a factor of 2 for normal subjects and a factor of 1.5 for subjects with restricted detoxification capacity.

The direct evaluation of the capability of the novel 2DOB breath test to provide reliable estimates of the hepatic drug detoxification rate would ultimately require measurements of arteriovenous drug concentration differences across the patient's liver. As this has to be excluded for obvious reasons,

an indirect evaluation may consist in testing the capability of the 2DOB test to correctly discriminate between liver healthy and clinically diagnosed liver-diseased subjects. Taking into account the patient-specific  $\text{CO}_2$  kinetics in the estimation of the hepatic conversion rate of  $^{13}\text{C}$ -methacetin (parameter  $k_L$ ) provided a modest improvement of the discrimination between normal versus impaired detoxification capacity compared with the conventional LiMAX test. This is plausible because the likelihood that the "true" maximum of the DOB curve (the basis for the definition of the LiMAX score) of patients with severe liver dysfunction is falsely shifted up to values within the normal range just due to an extreme  $\text{CO}_2$  kinetics remains small: The coefficient of variation for the exemplary case shown in Fig. 4b was 0.14. The same holds for the likelihood of a strong down shift of the DOB maximum for subjects with high detoxification capacity. However, for patients with detoxification capacities close to the borderline between normal and mild metabolic dysfunction, the correction of DOB values for the influence of the individual  $\text{CO}_2$  kinetics may be relevant. 7 out of 8 the cases misclassified by the LiMAX, but correctly classified by the 2DOB parameter  $k_L$  are just in the middle range of  $k_L$  and LiMAX values (see grey-shaded region in Fig. 5) where a marginal shift in the maximum of the DOB curve may turn the classification result into the opposite.

In summary, the already well-established LiMAX score appears to be robust against the bias caused by the systemic  $\text{CO}_2$  kinetics in subjects comprising either normal or drastically reduced hepatic detoxification capacities. However, in subjects with borderline capacities, the 2DOB test promises a significant improvement of the classification quality.

### Plasma clearance of methacetin = metabolization and temporary storage

Generally, owing to its chemical similarity with acetaminophen, methacetin should rapidly and evenly distribute throughout most tissues and fluids as reported for acetaminophen (Forrest et al. 1982). Therefore, it is difficult to specify the volume of model compartment Y to convert the apparent rate constants into true rate constants. An intriguing finding of our study was that a substantial part of the test drug should be transiently stored in a body compartment that is closely associated with the liver. Distinction between two modes of plasma clearance of methacetin, reversible storage (without detoxification) and detoxification, was possible because the reversible exchange of the drug with the storage compartment influences the shape of the declining part of the DOB curve (see sensitivity analysis in Fig. 5 of the Supplementary Information). The decline of the DOB curve is influenced by both the uptake of  $^{13}\text{CO}_2$  from the plasma into other body and the uptake of  $^{13}\text{C}$ -methacetin into the exchange compartments. Hence,

the share of the CO<sub>2</sub> kinetics in the shape of the declining part of the DOB curve has to be known to estimate reliable values for the relevant parameters,  $k_{+M}$  and  $k_{-M}$ , determining the magnitude of the parameter  $M_L$ . An issue with the correct estimation of  $M_L$  occurs if the DOB curve is very flat, i.e. for livers with very low detoxification capacity  $k_L$  (see Fig. 3b). Fortunately this is no impact on the 2DOB score because for very small values of  $k_L$  the estimated value of  $M_L$  is also small, i.e.  $M_L \ll M_{Lc}$ , so that the 2DOB score is identical with  $k_L$  and the liver is classified as functionally impaired.

The physiological and anatomical nature of methacetin-storing body compartments which we condensed into the single model compartment Y remains elusive. Owing to the chemical similarity of methacetin with acetaminophen (APAP) we searched the literature for known facts about the pharmacokinetics of APAP. A recently published compartment model of APAP clearance from the plasma worked well without introducing a reversible exchange of APAP (Mian et al. 2019). Reversible binding of acetaminophen to plasma proteins like albumin (the plasma level of which can be reduced in severe liver failure) can be excluded as this is a very rapid process in contrast to the rather slow uptake and release according to our modeling data. With an association rate constant of about  $5.8 \times 10^4$ /M/s as determined for binding of tryptophan to albumin (Talbert et al. 2002) and an average albumin plasma concentration of 40 g/l = 600 μM, the apparent first-order rate constant for drug binding would be about 2000/s which is about five orders of magnitude larger than typical values of the model parameter  $k_{-M}$ . Thus, it is tempting to assume that compartment Y is identical with the liver itself, whereby the transient storage without chemical conversion is either confined to a special fraction of hepatocytes or occurs in an intra-cellular compartment of all hepatocytes. Hepatocytes have cytosolic binding proteins (e.g. ligandin) that act as storage compartment for drugs and endogenous metabolites. Therefore, with loss of hepatocytes, there is loss of storage capacity. This explanation receives support from our observation that the storage capacity  $M_L$  was lower in the group of older controls compared with the group of younger controls (see section “Robustness of the 2DOB test” below), possibly reflecting the age-dependent decline of active liver mass (Wakabayashi et al. 2002). As binding of the drug to cellular binding proteins will be similarly fast as binding to serum proteins, the rate constants for the reversible uptake to and release of from hepatocytes should reflect the transport rates across the plasma membrane (Austin et al. 2005). The capacity of the liver to efficiently remove drugs and other xenobiotics from the circulation may also be determined by its capacity to transiently store the drug before it can be chemically converted. This may in particular hold in situations when the detoxifying enzyme systems become saturated. Testing this

hypothesis would mean to quantify the arteriovenous difference of labeled <sup>13</sup>C-methacetin and <sup>13</sup>CO<sub>2</sub> in a laboratory animal. Anyway, storage of the test drug without immediate chemical conversion in the liver appears to be a mechanism that contributes to the rapid clearance of the drug from the plasma.

## Conclusion

Individual variations in the systemic CO<sub>2</sub> kinetics may have a significant influence on the parameters of the DOB curve. The novel test variant 2DOB takes this confounding effect into account and promises a significant improvement in the assessment of impaired hepatic detoxification capacity compared to the well-established LiMAX test in cases where the detoxification capacity is at the borderline between the normal and moderately reduced level. The suitability of the test for the reliable characterization of the natural history of chronic liver diseases (fatty liver → fibrosis → cirrhosis) has to be assessed in further studies. Validation of the test in an animal model would be perfect to remove remaining uncertainties of the model-based analysis by direct measurement of hepatic CYP activities and concentration profiles of <sup>13</sup>CO<sub>2</sub> and <sup>13</sup>C-methacetin in different body compartments during the long-term progression of a liver disease (e.g. non-alcoholic fatty liver NAFLD).

**Acknowledgements** Open Access funding provided by Projekt DEAL.

**Funding** This research was funded by the German Systems Biology Programs “LiSyM”, grant no. 31L0057 and 031L0058, sponsored by the German Federal Ministry of Education and Research (BMBF), and the Max Planck Society.

## Compliance with ethical standards

**Conflict of interest** M. Stockmann is the inventor of the LiMAX test and has capital interest in Humedics, the company marketing the LiMAX test.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Austin RP et al (2005) The binding of drugs to hepatocytes and its relationship to physicochemical properties. *Drug Metab Dispos* 33(3):419–425
- Barstow TJ et al (1989) Changes in breath  $^{13}\text{C}$ CO $_2$ / $^{12}\text{C}$ CO $_2$  consequent to exercise and hypoxia. *J Appl Physiol* 66(2):936–942
- Barstow TJ et al (1990) Influence of increased metabolic rate on [ $^{13}\text{C}$ ]bicarbonate washout kinetics. *Am J Physiol* 259(1 Pt 2):R163–R171
- Bonfrate L et al (2015) Dynamic carbon 13 breath tests for the study of liver function and gastric emptying. *Gastroenterol Rep (Oxf)* 3(1):12–21
- Braden B et al (2007)  $^{13}\text{C}$ -breath tests: current state of the art and future directions. *Dig Liver Dis* 39(9):795–805
- Buechter M et al (2018) Liver maximum capacity (LiMAx) test as a helpful prognostic tool in acute liver failure with sepsis: a case report. *BMC Anesthesiol* 18(1):71
- Cieslak KP et al (2016) Liver function declines with increased age. *HPB (Oxford)* 18(8):691–696
- Craig H (1957) Isotopic standards for carbon and oxygen and correction factors for mass-spectrometric analysis of carbon dioxide. *Geochim Cosmochim Acta* 12(1–2):133–149
- Dinesen L et al (2008)  $^{13}\text{C}$ -methacetin-breath test compared to also noninvasive biochemical blood tests in predicting hepatic fibrosis and cirrhosis in chronic hepatitis C. *Dig Liver Dis* 40(9):743–748
- European Association for Study of Liver, Asociacion Latinoamericana para el Estudio del Hígado (2015) EASL-ALEH Clinical Practice Guidelines: non-invasive tests for evaluation of liver disease severity and prognosis. *J Hepatol* 63(1):237–264
- Fernandez Del Rio R et al (2015) Volatile biomarkers in breath associated with liver cirrhosis—comparisons of pre- and post-liver transplant breath samples. *EBioMedicine* 2(9):1243–1250
- Forrest JA, Clements JA, Prescott LF (1982) Clinical pharmacokinetics of paracetamol. *Clin Pharmacokinet* 7(2):93–107
- Gorowska-Kowolik K, Chobot A, Kwiecien J (2017) ( $^{13}\text{C}$ ) Methacetin breath test for assessment of microsomal liver function: methodology and clinical application. *Gastroenterol Res Pract* 2017:7397840
- Haycock GB, Schwartz GJ, Wisotsky DH (1978) Geometric method for measuring body surface area: a height-weight formula validated in infants, children, and adults. *J Pediatr* 93(1):62–66
- Holzhtutter HG et al (2013) Assessment of hepatic detoxification activity: proposal of an improved variant of the ( $^{13}\text{C}$ )-methacetin breath test. *PLoS ONE* 8(8):e70780
- Irving CS et al (1983) [ $^{13}\text{C}$ ]bicarbonate kinetics in humans: intra- vs. interindividual variations. *Am J Physiol* 245(2):R190–R202
- Jara M et al (2015) Prognostic value of enzymatic liver function for the estimation of short-term survival of liver transplant candidates: a prospective study with the LiMAx test. *Transpl Int* 28(1):52–58
- Jonderko K et al (2008) A systematic study of a neutral meal suitable for subjects undergoing ( $\text{CO}_2$ )- $\text{C}$ -13 breath tests. *Med Sci Monit* 14(10):Cr543–Cr546
- Klatt S et al (1997) Evaluation of the  $^{13}\text{C}$ -methacetin breath test for quantitative liver function testing. *Z Gastroenterol* 35(8):609–614
- Lau T, Ahmad J (2013) Clinical applications of the Model for End-Stage Liver Disease (MELD) in hepatic medicine. *Hepat Med* 5:1–10
- Mian P et al (2019) Population pharmacokinetic modeling of acetaminophen and metabolites in children after cardiac surgery with cardiopulmonary bypass. *J Clin Pharmacol* 59(6):847–855
- Mossman D (1995) Resampling techniques in the analysis of non-binormal ROC data. *Med Decis Mak* 15(4):358–366
- Razlan H et al (2011) Diagnostic value of the C methacetin breath test in various stages of chronic liver disease. *Gastroenterol Res Pract* 2011:235796
- Rubin T et al (2011) Quantitative determination of metabolization dynamics by a real-time  $^{13}\text{C}$ CO $_2$  breath test. *J Breath Res* 5(2):027102
- Rubin T et al (2016) Liver status assessment by spectrally and time resolved IR detection of drug induced breath gas changes. *Photonics* 3(2):31
- Schaffer J (2015) What not to multiply without necessity. *Australas J Philos* 93(4):644–664
- Schmucker DL (2005) Age-related changes in liver structure and function: Implications for disease? *Exp Gerontol* 40(8–9):650–659
- Schmucker DL, Wang RK (1980) Age-related changes in liver drug metabolism: structure vs function. *Proc Soc Exp Biol Med* 165(2):178–187
- Stockmann M et al (2009) Prediction of postoperative outcome after hepatectomy with a new bedside test for maximal liver function capacity. *Ann Surg* 250(1):119–125
- Stravitz RT et al (2015) Use of the methacetin breath test to classify the risk of cirrhotic complications and mortality in patients evaluated/ listed for liver transplantation. *J Hepatol* 63(6):1345–1351
- Talbert AM et al (2002) Determination of drug-plasma protein binding kinetics and equilibria by chromatographic profiling: exemplification of the method using L-tryptophan and albumin. *Anal Chem* 74(2):446–452
- Vergara IA et al (2008) StAR: a simple tool for the statistical comparison of ROC curves. *BMC Bioinform* 9:265
- Wakabayashi H et al (2002) Evaluation of the effect of age on functioning hepatocyte mass and liver blood flow using liver scintigraphy in preoperative estimations for surgical patients: comparison with CT volumetry. *J Surg Res* 106(2):246–253
- Zhu D et al (2014) Role of the ( $^{13}\text{C}$ )-methacetin breath test in the assessment of acute liver injury in a rat model. *World J Gastroenterol* 20(32):11305–11312

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Affiliations

Hermann-Georg Holzhütter<sup>1</sup> · Tilo Wuensch<sup>2</sup> · Robert Gajowski<sup>3,4</sup> · Nikolaus Berndt<sup>1,5</sup> · Sascha Bulik<sup>6</sup> · David Meierhofer<sup>3</sup> · Martin Stockmann<sup>2</sup>

<sup>1</sup> Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Institute of Biochemistry, Computational Systems Biochemistry Group, Charitéplatz 1, 10117 Berlin, Germany

<sup>2</sup> Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Department of Surgery, Augustenburger Platz 1, 13353 Berlin, Germany

<sup>3</sup> Max Planck Institute of Molecular Genetics, Mass Spectroscopy Facility, Ihnestr. 63-73, 14195 Berlin, Germany

<sup>4</sup> Department of Biology, Chemistry, and Pharmacy, Free University Berlin, Takustraße 3, 14195 Berlin, Germany

<sup>5</sup> Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Institute for Computational and Imaging Science in Cardiovascular Medicine, Augustenburger Platz 1, 13353 Berlin, Germany

<sup>6</sup> Federal Institute of Risk Assessment, Diedersdorfer Weg 1, 12277 Berlin, Germany