Prediction of Busulfan Clearance by Predose Plasma Metabolomic Profiling

Jeannine S. McCune[1,*], Sandi L. Navarro[2], K. Scott Baker[3,4], Linda J. Risler[5], Brian R. Phillips[5], Timothy W. Randolph[3], Laura Shireman[5], H. Gary Schoch[3], H. Joachim Deeg[3,6], Yuzheng Zhang[3], Alex Men[5], Loes Maton[7] and Alwin D. R. Huitema[7,8,9]

Intravenous busulfan doses are often personalized to a target plasma exposure (targeted busulfan) using an individual's busulfan clearance (BuCL). We evaluated whether BuCL could be predicted by a predose plasma panel of 841 endogenous metabolomic compounds (EMCs). In this prospective cohort of 132 hematopoietic cell transplantation (HCT) patients, all had samples collected immediately before busulfan administration (preBU) and 96 had samples collected 2 weeks before busulfan (2-week-preBU). BuCL was significantly associated with 37 EMCs after univariate linear regression analysis and controlling for false discovery (< 0.05) in the 132 preBU samples. In parallel, with preBU samples, we included all 841 EMCs in a least absolute shrinkage and selection operator–penalized regression which selected 13 EMCs as predominantly associated with BuCL. Then, we constructed a prediction model by estimating coefficients for these 13 EMCs, along with sex, using ordinary least-squares. When the resulting linear prediction model was applied to the 2-week-preBU samples, it explained 40% of the variation in BuCL (adjusted R^2 = 0.40). Pathway enrichment analysis revealed 18 pathways associated with BuCL. Lysine degradation followed by steroid biosynthesis, which aligned with the univariate analysis, were the top two pathways. BuCL can be predicted before busulfan administration with a linear regression model of 13 EMCs. This pharmacometabolomics method should be prioritized over use of a busulfan test dose or pharmacogenomics to guide busulfan dosing. These results highlight the potential of pharmacometabolomics as a precision medicine tool to improve or replace pharmacokinetics to personalize busulfan doses.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

☑ Busulfan area under the plasma concentration-time curve (AUC) is associated with various clinical outcomes, but ~25% of patients achieve their target AUC with current dosing methods. Although pharmacokinetic (PK)–guided busulfan dosing reaches the target AUC in the majority of patients, it is resource-intensive and time-intensive. Thus, alternative approaches to estimate busulfan clearance are needed to achieve the target AUC with the first dose.

WHAT QUESTION DID THIS STUDY ADDRESS?

☑ This study addresses the question regarding whether endogenous metabolomic compounds in the plasma before busulfan administration are associated with busulfan clearance.

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?

☑ We found that 13 endogenous metabolomic compounds in plasma obtained 2 weeks before busulfan administration had a strong ability to predict intravenous busulfan clearance, particularly among patients with slow clearance.

HOW MIGHT THIS CHANGE CLINICAL PHARMACOLOGY OR TRANSLATIONAL SCIENCE?

☑ Of the three methods to predict busulfan clearances, pharmacometabolomics should be prioritized over the use of a busulfan test dose or pharmacogenomics to guide busulfan dosing.

Many allogeneic hematopoietic cell transplantation (HCT) recipients receive high-dose busulfan (busulfan) as part of their HCT conditioning regimen.[1] Busulfan has a narrow therapeutic index. Busulfan exposure, expressed as area under the plasma concentration-time curve (AUC),[2] is associated with the efficacy and toxicity of many busulfan-based conditioning regimens.[1] Low

[1]Department of Hematologic Malignancies Translational Sciences, City of Hope, Duarte, California, USA; [2]Public Health Sciences Division, Fred Hutchinson Cancer Center, Seattle, Washington, USA; [3]Clinical Research Division, Fred Hutchinson Cancer Center, Washington, Seattle, USA; [4]Department of Pediatrics, University of Washington, Washington, Seattle, USA; [5]Department of Pharmaceutics, University of Washington, Seattle, Washington, USA; [6]Department of Medicine, University of Washington, Seattle, Washington, USA; [7]Department of Pharmacy & Pharmacology, Netherlands Cancer Institute, Amsterdam, The Netherlands; [8]Department of Pharmacology, Princes Maxima & Pharmacology, Netherlands Cancer Institute, Amsterdam, The Netherlands; [9]Department of Clinical Pharmacy, University Medical Center Utrecht, Utrecht, The Netherlands. *Correspondence: Jeannine S. McCune (jmccune@uw.edu)

Received August 11, 2022; accepted November 7, 2022. doi:10.1002/cpt.2794

busulfan AUC, caused by underdosing or rapid clearance (or both), is associated with reduced efficacy (e.g., increased risk of relapse or rejection). Conversely, high busulfan AUC, caused by overdosing or a slow clearance, is associated with an increased risk of liver toxicity and nonrelapse mortality.[1]

These clinical outcomes are improved by pharmacokinetic (PK)-guided busulfan dosing, in which busulfan doses are personalized to clinician-chosen target AUC using the individual patient's (i.e., patient-specific) busulfan clearance.[1] Ideally, the target AUC is achieved with the initial busulfan dose. The current clinical practice of using various age-based and weight-based dosing schemes for the initial busulfan dose achieves the clinician-chosen target busulfan AUC in a minority (22.6%[3] to 24.3%[4]) of patients. PK-guided busulfan dosing is often successful: personalizing intravenous (i.v.) busulfan doses using PK-guided dosing results in over 85% of patients achieving the clinician-chosen target busulfan AUC at the end of 4 days of busulfan therapy.[3] Although PK-guided busulfan dosing is feasible and improves clinical outcomes, the current process (**Table S1**) is resource-intensive, labor-intensive, and time-sensitive, in part because of the short (up to 4 days) administration schedule of busulfan. In addition, PK-guided dosing is rarely feasible with the shorter (i.e., ≤ 2 day) busulfan courses included in reduced-intensity or gene therapy conditioning regimens.

It is anticipated that omics techniques could improve—or ideally replace—PK-guided busulfan dosing to achieve the target busulfan AUC with the first dose and to decrease resource intensity. Here, we sought to validate our previous global[5] or targeted[6] pharmacometabolomic analyses of endogenous metabolomic compounds (EMCs) associated with i.v. busulfan clearance (abbreviated busulfan clearance hereafter) and extend our findings using a larger set of identified EMCs. In our previous global analysis, we found that tertiles of increasing busulfan clearance were associated with 21 ions ($R^2 \geq 0.3$), two of which were identified as linoleic acid and the bile acids deoxycholic acid and/or chenodeoxycholic acid. Bile acids have been shown to inhibit glutathione-*S*-transferases *in vitro*.[7] Using targeted analysis, we found that the EMCs glycine, N-acetylglycine, 2-hydroxyisovaleric acid, creatine, serine, and tyrosine were significantly associated with busulfan clearance, and that pathway analyses were consistent, revealing that the glycine, serine, and threonine pathway was also associated with busulfan clearance. Glycine is conjugated to busulfan during detoxification, indicating biological plausibility of our findings in predicting busulfan clearance. Against this background, we sought to validate that these EMCs were associated with busulfan clearance and expand the repertoire of EMCs using a global pharmacometabolomics approach. We evaluated 841 identified EMCs in plasma samples from two longitudinal timepoints, 2 weeks before and immediately before i.v. busulfan administration in HCT recipients receiving i.v. busulfan.

METHODS

Study population

This was an ancillary prospective study of 132 participants who received HCT conditioning with i.v. busulfan from December 2014 to November 2018 under the aegis of a protocol approved by the Fred Hutchinson Cancer Research Center (Fred Hutch) Institutional Review Board (clinicaltrials.gov protocol number NCT02291965). The inclusion criterion was: scheduled to receive i.v. busulfan (any dose, any number of doses, any dosing frequency) as part of their HCT conditioning. The exclusion criteria were: inability to read English; female patients who were pregnant or breastfeeding; life expectancy severely limited by diseases other than malignancy. All participants provided written informed consent prior to study procedures. The conditioning regimen and postgraft immunosuppression were not affected by participation in this study. Standard practice for prophylaxis of busulfan-induced seizures was phenytoin (adults) or levetiracetam (children). Antiemetics, antibiotics, and antifungals were given per Institutional Standard Practice Guidelines. Demographic data were taken from the participants' medical charts (age, sex, height, total body weight [i.e., actual body weight (ABW)], body surface area, and clinical information (disease, conditioning regimen)).

Busulfan dosing, PK sampling, and quantitation

The initial busulfan dose, chosen by the treatment protocol or the participant's treating physician, was based on ABW or body surface area. The busulfan Dose 1 was calculated using ABW if it was less than ideal body weight, or adjusted ideal body weight (which equals 0.25 (ABW − ideal weight) + ideal weight) if it was greater than ideal body weight. The ideal body weight in adults was calculated as follows: for males = 50 kg + (2.3 kg for each inch over 5 feet); for females = 45.5 kg + (2.3 kg for each inch over 5 feet). Because of the circadian variability of busulfan PK,[8] all busulfan PK sampling occurred in the morning. A detailed description of the PK sampling and quantitation is included in **Supplemental Methods 1**.

Busulfan clearance estimation

Busulfan clearance (mL/min/kg of normal fat mass) over the entire duration of busulfan treatment was the primary end point. The patient-specific busulfan clearance, not the busulfan AUC, is the relevant end point because the goal with PK-guided dosing is to obtain the patient-specific busulfan clearance to be used for dose personalization to achieve the clinician-chosen target busulfan AUC.

This cohort had busulfan clearance over the entire duration of busulfan treatment as the end point of interest, as compared with Dose 1 busulfan clearance in our retrospective cohort.[5,6] This change was made for two reasons: (i) The busulfan clearance did not consistently change over time in this cohort; (ii) An increasing number of HCT programs do PK-guided busulfan dosing after multiple doses, so the clearance over the entire duration of treatment is of increasing importance.[1]

Busulfan clearance was estimated using a population approach based on nonlinear mixed-effects modeling. Nonlinear mixed-effect modeling was applied using the software package NONMEM (version 7.3.0, Icon, Hanover, MD). Pirana (version 2.9.5, Amsterdam, The Netherlands) and R (version 3.3.3, Vienna, Austria) were used for workflow management and data handling and visualization, respectively.

Estimation of busulfan AUCs was accomplished using maximum *a posteriori* probability Bayesian estimation of PK parameters, incorporating a blend of individualized PK data and a population parameter prior. The maximum *a posteriori* probability Bayesian method used individual patient busulfan concentration-time data, together with a PK model and mean parameter values plus their variance (derived from a prior population PK analysis of 1,610 HCT recipients).[9] All reported clearances are based on normal fat mass, the optimal body metric for busulfan clearance over a population of pediatric to adult allogeneic HCT recipients.[9] The use of normal fat mass to characterize busulfan clearance was based on the findings of McCune *et al.*,[9] who reported a population PK model based on normal fat mass and maturation based on postmenstrual age built from 12,380 busulfan concentration-time points obtained after i.v. busulfan administration in 1,610 HCT recipients, aged 0.1–66 years. All clearance and volume parameters were scaled for body size and composition using allometric theory and predicted fat free mass. The population prior parameters were not changed during this study. The estimated busulfan clearance was calculated by dividing the Bayesian AUC estimate by the

busulfan dose. **Supplemental Methods 2** contains the NONMEM code. **Figures S1 and S2** show the goodness of fit and simulated vs. actual concentration-time data, respectively.

Pharmacometabolomics sample collection

Longitudinal blood samples (3 mL/sample) were scheduled to be collected in sodium heparin tubes before i.v. busulfan dosing: up to 2 weeks prior to the first conditioning dose (2-week pre-busulfan sample), and immediately before administration of i.v. busulfan Dose 1 (pre-busulfan sample).

Because we sought to identify EMCs predictive of busulfan clearance, samples were obtained at two timepoints before busulfan administration. The 2-week pre-busulfan sample collection time was the earliest feasible time within the final HCT workup (i.e., time period, typically up to 2 weeks, in which patients undergo final assessment to determine whether they can receive an HCT). If a predictive association was found and pharmacometabolomics-guided (EMC-guided) i.v. busulfan dosing was subsequently used clinically, this sample would allow for more time for pharmacometabolomic quantitation and data interpretation. The immediately pre-busulfan sample was the latest feasible time before busulfan administration. If a predictive association was found and EMC-guided i.v. busulfan dosing was subsequently used clinically, this sample necessitates a rapid quantitation and data interpretation. A total of 228 pharmacometabolomic samples were obtained: 96 participants had 2-week pre-busulfan samples and 132 participants has pre-busulfan samples.

The pharmacometabolomic samples were immediately refrigerated at 4°C, stored for up to 4 hours from time of collection, centrifuged to plasma, and immediately stored at −80°C. The samples underwent at most one freeze–thaw cycle before pharmacometabolomic analysis (i.e., the analysis was conducted after the first or the second thaw).

Global pharmacometabolomics analysis

EMCs profiling of plasma was completed by Metabolon (Durham, NC), as we have previously described.[10] The samples were shipped on dry ice to Metabolon's facility and stored at −80°C upon receipt. Samples were divided into five aliquots; one was held in reserve while each of the others was analyzed by one of four different mass spectrometry methods. Raw data were extracted, peak-identified, and quality control processed; then Metabolon's proprietary software was used to confirm the consistency of peak identification across the various samples. EMCs were identified by comparison of the processed data to Metabolon's library entries of purified standards or recurrent unknown entities. Library matches for each EMC were checked for each sample and corrected if necessary. Using the criteria established by the Chemical Analysis Working Group as part of the Metabolomics Standards Initiative,[11] most EMCs met the level 1 standards for identification with the remainder meeting the level 2 standards per the same criteria. ANCOVA contrasts were performed to examine differences in EMC levels between the samples at each timepoint. See **Supplemental Methods 3** for more details about EMCs quantitation and quality control.

Statistical analysis

All data transformations and analyses were carried out using Stata (version 17, College Station, TX) and using the R software package (version 4.1.1). A total of 841 EMCs were reliably measured and retained for analysis. The majority of EMCs were skewed to higher values and were therefore centered log-transformed to approximate a normal distribution. First, we performed univariate linear regression to assess marginal associations of each EMC individually on busulfan clearance (continuous) for each of the two samples, separately. In our previous evaluation of 1,610 HCT recipients (n = 904 male and n = 689 female) used to inform the present analysis, we found no effect of sex or weight on busulfan clearance, the end point of interest.[9] However, to ensure that sex or weight did not confound our results, sex was included in all univariate models and normal fat mass was used to normalize the busulfan clearance. Benjamini-Hochberg methods were used to control for false discovery rate (FDR).[12] Individual EMCs were considered significant at FDR < 0.05.

Second, we further evaluated the ability of EMCs to predict busulfan clearance by constructing a prediction model using linear regression. For this, we treated the 132 pre-busulfan samples as training data and fit a least absolute shrinkage and selection operator (LASSO)–penalized regression model to select the EMCs most associated with busulfan clearance. Here we applied the high-dimensional inference method with 500 subsample iterations using the stability selection method[13] to select among the 841 EMCs most associated with busulfan clearance. EMCs were retained if they were selected in at least 30% of the 500 subsample iterations. We then refit an ordinary linear regression model using the 13 selected EMCs. The coefficients from this model were used to define a linear prediction model and we applied this to the 96 samples obtained 2-week-pre-busulfan (see **Figure 1**). The percentage of variance explained, R^2, by the 13-EMCs linear prediction model was calculated on the 2-week pre-busulfan data.

Third, to consider EMCs that coordinately predict busulfan clearance, pathway analyses inputting all EMCs were carried out using MetaboAnalyst 5.0[14] integrating pathway enrichment analysis and pathway topology analysis for visualization. Within the pathway analysis module, EMCs were auto-scaled (mean-centered and divided by the standard deviation of each variable), and busulfan clearance was evaluated as a continuous outcome. EMCs were matched to those supported in the MetaboAnalyst compound library (derived from Kyoto Encyclopedia of Genes and Genomes,[15] Small Molecule Pathway Database,[16] and Human Metabolome Database[17]), which included EMCs in 49 metabolic pathways with sufficient coverage for meaningful pathway analysis. The Global test,[18] which evaluates changes among groups of EMCs, was used for statistical significance of pathway enrichment analysis, with FDR < 0.1 for multiple comparisons. Betweenness centrality (shortest path between nodes), based on an individual EMC's centrality in a given metabolic network, was used to calculate EMCs' importance.[19] Pathway impact was calculated as the sum of the importance measures of the pathway-specific EMCs, normalized by the sum of the importance measures of all EMCs in each pathway.[14]

RESULTS

Patient characteristics

Pretransplant characteristics and diagnoses of the study population are given in **Table 1**. Mean age was 47 years (range 1.7–70), and body mass index 29 kg/m^2 (range 15–56). Slightly more participants were male (64%); the majority (93%) had a hematologic malignancy. Mean busulfan clearance was 12.1 ± 3.88 mL/min/kg of normal fat mass, which is in agreement with previous studies.[9]

Pharmacometabolomics

All 132 participants had a metabolomics sample at the pre-busulfan sample and 96 participants also had a 2-week pre-busulfan sample.

In the univariate analysis, 37 EMCs at the pre-busulfan sample were associated with busulfan clearance at FDR < 0.05 (**Table 2**). For the 2-week pre-busulfan sample, 27 EMCs were associated with busulfan clearance at FDR < 0.05 (**Table 3**). Several androgen steroid conjugates were statistically significant in both samples (11 and 12 of 17 EMCs in this pathway for pre-busulfan and 2-week pre-busulfan, respectively). Of note, all significant associations with busulfan clearance were positive with the exception of 3 EMCs in the pre-busulfan sample: cortisol, cortisone, and methyl-4-hydroxybenzoate sulfate.

A total of 13 EMCs were selected by the LASSO training model (**Table 4**) in the pre-busulfan training data set: urate,

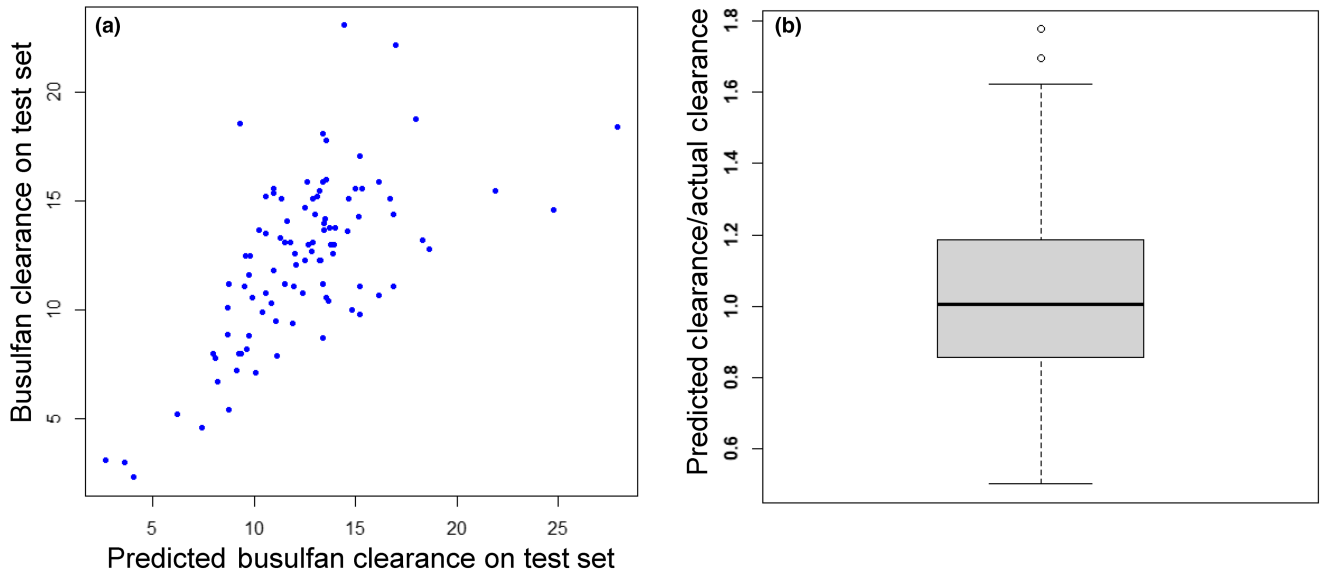

Figure 1 Accurate prediction of by the model of the EMCs in the 2-week pre-busulfan sample. Busulfan clearance is calculated using maximum *a posteriori* probability Bayesian estimation and reported in the units of mL/min/kg of normal fat mass. Panel (**a**) shows the predicted (*x*-axis) and the actual (*y*-axis) busulfan clearance model. The predicted clearance has a higher predictive power among patients with lower clearance. Panel (**b**) shows a box plot of the ratio between the predicted and actual clearance. Middle line is the median; top and bottom of the box show 25th and 75th percentile (i.e., the interquartile range or IQR); top whisker shows the 1.5×IQR above the 75% upper box, and bottom whisker means the minimum since it is within the 1.5×IQR below the 25% bottom box.

androstenediol (3-beta, 17-beta)-disulfate (2), mannonate, arachidonoylcarnitine, 5-alpha androstane 3-alpha 17-beta diol monosulfate (2), N-acetylcarnosine, and lysine, cortisone, methyl-4-hydroxybenzoate sulfate, palmitoyl-arachidonoyl glycerol, and cortisol and N-acetyl aspartyl-glutamate and 2-butenoylglycine. The majority of these EMCs were also found in the univariate analysis (i.e., the results of these two methods overlapped); 7 (i.e., urate, androstenediol (3-beta, 17-beta)-disulfate (2), mannonate, arachidonoylcarnitine, 5-alpha androstane 3-alpha 17-beta diol monosulfate (2), N-acetylcarnosine, and lysine) overlapped with the univariate analysis in both the pre-busulfan and 2-week pre-busulfan samples; 4 (i.e., cortisone, methyl-4-hydroxybenzoate sulfate, palmitoyl-arachidonoyl glycerol, and cortisol) in the pre-busulfan sample overlapped with the univariate analysis; and 2 (i.e., N-acetyl aspartyl-glutamate and 2-butenoylglycine) were not associated with univariate analysis and were selected only by the LASSO training model (**Table 4**).

When the linear prediction model with these 13 EMCs was applied to the 2-week-pre-busulfan samples, the predicted vs. actual busulfan clearance values were strongly correlated ($R^2 = 0.40$). Plotting the EMCs-predicted busulfan clearance vs. the actual busulfan clearance shows the predictive power of this model, especially among patients with lower clearance (**Figure 1**). In particular, an R^2 of 0.53 is obtained when this model is restricted to the slower half ($n = 48$, busulfan clearance < 10.2 mL/min/kg of normal fat mass) of the 2-week-pre-busulfan samples. A predicted i.v. busulfan clearance was considered to be accurate for each participant if the value was within 80–125% of the observed i.v. busulfan clearance. This 80–125% range was used by the US Food and Drug Administration (FDA) to ascertain bioequivalence[20] and is a reasonable metric by which to determine prediction accuracy. The majority (63.5%, 61 of 96 participants) fall in this range (**Figure 1b**).

In the pathway enrichment analysis of all the EMCs in the pre-busulfan sample, 18 pathways were significant at FDR < 0.05 (**Table 5**). For the 2-week pre-busulfan sample, there was one significant pathway at FDR < 0.05. Lysine degradation was the most significant pathway in both samples. This pathway contained 7 EMCs: lysine; 4-trimethylamonio-butanoate; aminoadipate; N6,N6,N6, trimethyl-L-lysine; N6,N6, dimethyl-lysine; and pipecolate—all of which were positively associated with busulfan clearance—and hydroxy-lysine, which was inversely associated.

DISCUSSION

We designed this independent cohort of prospectively collected samples to replicate our findings from a retrospective cohort analyzed with global[5] or targeted[6] metabolomics assays, and to expand our repertoire of EMCs. The key findings reported in this manuscript are (i) 37 EMCs quantitated in the pre-busulfan samples were highly associated with busulfan clearance. (ii) A predictive model built using the pre-busulfan samples exhibited a strong ability to predict busulfan clearance when applied to the 2-week pre-busulfan samples ($R^2 = 0.40$), especially among the patients with low clearance ($R^2 = 0.53$). These results demonstrate the potential of a small panel of EMCs to predict busulfan clearance before busulfan administration. (iii) Validation of our previous findings, showing that the glycine, serine, and threonine pathway was significant in our retrospective cohort.[6] (iv) We identified novel pathways associated with busulfan clearance, including lysine degradation, which was most highly associated with busulfan clearance for both the 2-week pre-busulfan and the immediately pre-busulfan samples, and steroid biosynthesis.

Table 1 Demographic and clinical data for the HCT study population (*N* = 132)

Parameter	*N*[a]
Age (year)	46.4 (1.7–70)
Male sex	84 (64%)
Adjusted ideal body weight (kg)[b]	82.6 ± 28.7
Normal fat mass (kg)	68.8 ± 21.6
HCT conditioning[c]	
Cyclophosphamide/^{T}Busulfan ± others	64 (48%)
^{T}Busulfan/Cyclophosphamide ± others	35 (27%)
^{T}Busulfan/Fludarabine ± others	19 (14%)
CNS regimen (thiotepa, busulfan, cyclophosphamide)	7 (5%)
Fludarabine/Cyclophosphamide/^{T}Busulfan	5 (4%)
^{T}Busulfan/Melphalan	2 (2%)
Busulfan dosing frequency	
Every 6 hours	17 (13%)
Every 12 hours	1 (1%)
Every 24 hours	114 (86%)
Diagnosis	
Myelodysplastic syndrome (MDS)	71 (54%)
Acute myeloid leukemia (AML)	32 (24%)
Chronic myeloid leukemia (CML)	12 (9%)
Non-Hodgkin's lymphoma	8 (6%)
Nonmalignant diseases	7 (5%)
Neuroblastoma	3 (2%)

AUC, area under the plasma concentration-time curve; CNS, central nervous system; HCT, hematopoietic cell transplantation; PK, pharmacokinetic; ^{T}Busulfan: participants received targeted or PK-guided dosing of busulfan, in which the i.v. busulfan dose was personalized to a target AUC based on the individual's clearance.
[a]Data presented as: number (%) or mean ± standard deviation; percentages may not add up to 100 due to rounding. [b]Actual body weight was used for busulfan dosing if ABW was less than ideal body weight, whereas adjusted ideal body weight was used if ABW was greater than ideal body weight. [c]Listed in administration order.

Busulfan has a narrow therapeutic window, with many HCT recipients receiving PK-guided dosing of busulfan to achieve their target busulfan AUC.[2] Although PK-guided dosing improves clinical outcomes,[1] it is optimal to achieve the target AUC with the first dose—i.e., the "right-dose-first-time" paradigm.[21] Various methods to pre-emptively predict busulfan clearance have been evaluated (**Table S1**). One method is to estimate busulfan clearance after administration of a pre-HCT "test dose," which is a single small dose of busulfan ranging from 0.25 to 0.8 mg/kg. While the use of a test dose has been able to minimize subsequent dose adjustments during the actual conditioning, the test dose strategy does not predict clearance well enough to replace PK-directed busulfan dosing.[1] An additional method to predict an individual patient's busulfan clearance is to use pre-emptive pharmacogenomics of the genes regulating the glutathione *S*-transferase (GST) enzymes involved in busulfan metabolism.[1] GST isoenzymes A1-1 mainly catalyze this reaction; GSTM1-1 and GSTP1-1 have minor roles.[22,23] However, *GSTA1* and *GSTM1* polymorphisms are not consistently associated with busulfan clearance,[1] potentially due to redundancy in function across GST enzymes.[24] The third method is to predict busulfan clearance using an EMC panel from a pharmacometabolomic sample obtained pre-busulfan. The 2-week pre-busulfan sample was chosen as the test data set because its timing is logistically better than the pre-busulfan sample in the context of our long-range goal being to improve or replace PK-guided busulfan dosing. If a panel of EMCs quantitated 2-weeks before busulfan accurately predicts busulfan clearance (**Figure 1**), then there are 14 days to quantitate the EMCs and estimate the busulfan clearance such that the predicted busulfan clearance can be used to personalize the first busulfan dose. This is much simpler than the current process of administering busulfan and having < 36 hours to quantitate and model the busulfan concentration-time data.

To our knowledge, there is no accepted framework for evaluating how well busulfan clearance is predicted by EMCs, pre-emptive pharmacogenomics, or test dose, despite the latter being posited over 20 years ago.[25] Future research should consider evaluating how well the target AUC is achieved with dose personalization using the busulfan clearance estimated with this panel of 13 EMCs. However, the "gold standard" in such analysis is not clear because of inaccuracies in the quantitation, PK modeling, and dose recommendations for PK-directed busulfan dosing in HCT recipients.[26] A recent busulfan proficiency exercise revealed that, of the four proficiency samples evaluated, between 67% and 85% of the busulfan quantitation results were accurate (i.e., within 85–115% of the reference value). Furthermore, 88% of round #1 and 71% of round #2 dose recommendation answers were correct.

Of the 13 EMCs selected by LASSO for the prediction model, 11 overlapped with the univariate regression analysis (**Table 2**), providing further confidence in our findings. Our previous assay included 5 EMCs in the glutathione pathway, i.e., glycine, pyroglutamic acid, ornithine, glutamic acid, and cadaverine.[6] However, our current assay contained only six of the 28 EMCs in the pathway: glycine, glutamate, cysteine, cys-gly, ornithine, and spermidine. This may have been insufficient for a complete evaluation. Higher EMC concentrations were associated with faster busulfan clearance for all the EMCs except glycine where lower glycine concentrations were associated with a faster busulfan clearance. It is tempting to speculate that more substrate for glutathione production may be driving the association between glycine and busulfan clearance. In fact, glutathione was one of the pathways significantly associated with increased busulfan clearance. We had previously found that deoxycholic or chenodeoxycholic acid[5] was associated with busulfan clearance, with the putative mechanism being inhibition of GST activity by bile acids.[7] However, i.v. busulfan was not associated with bile acids in the univariate or pathway analysis. Notably, linoleic acid was significant in our retrospective analysis[5] and the linoleic acid metabolism pathway was significant in this prospective cohort (**Table 5**). Our working hypothesis for the association of these pathways on busulfan clearance is presented in **Figure 2**.

Our data also supported the importance of mitochondrial dysregulation. Kim *et al.*[27] reported that urinary phenylacetylglutamine and two acylcarnitines were associated with busulfan AUC in 130 pediatric HCT recipients receiving PK-directed busulfan. They suggested that mitochondrial dysregulation influenced

Table 2 Pre-busulfan sample's EMC association with busulfan clearance (FDR < 0.05)

EMCs[a]	Beta coefficient[b]	*P* value[b]	FDR[c]
5alpha-androstan-3beta,17beta-diol disulfate	0.18	3.60E-11	3.03E-08
mannonate	0.06	6.60E-10	5.52E-07
androstenediol (3beta,17beta) disulfate (2)	0.11	1.10E-09	8.98E-07
N,N,N-trimethyl-alanylproline betaine (TMAP)	0.08	1.40E-08	1.16E-05
androstenediol (3beta,17beta) disulfate (1)	0.13	2.60E-08	2.16E-05
androstenediol (3beta,17beta) monosulfate (2)	0.11	9.50E-08	0.00008
5alpha-androstan-3beta,17alpha-diol disulfate	0.12	2.10E-07	0.0002
androsterone glucuronide	0.12	2.40E-07	0.0002
urate[d]	0.03	5.90E-07	0.0005
epiandrosterone sulfate	0.15	8.20E-07	0.0007
N-acetylcarnosine	0.06	1.30E-06	0.001
hydroxy-N6,N6,N6-trimethyllysine	0.05	1.60E-06	0.001
androsterone sulfate	0.16	1.70E-06	0.001
3-carboxy-4-methyl-5-propyl-2-furanpropanoate (CMPF)	0.13	2.10E-06	0.002
androstenediol (3alpha, 17alpha) monosulfate (3)	0.12	2.20E-06	0.002
5alpha-androstan-3alpha,17beta-diol monosulfate (2)	0.09	2.80E-06	0.002
picolinoylglycine	0.07	4.20E-06	0.003
creatinine[e]	0.03	4.30E-06	0.004
N-acetylalanine	0.03	4.90E-06	0.004
palmitoyl-arachidonoyl-glycerol (16:0/20:4)	0.10	4.90E-06	0.004
cortisol	−0.13	9.70E-06	0.008
N6,N6,N6-trimethyllysine	0.04	0.00001	0.01
cortisone	−0.12	0.00001	0.01
imidazole lactate	0.04	0.00002	0.01
methyl-4-hydroxybenzoate sulfate	−0.20	0.00002	0.01
deoxycarnitine[d]	0.03	0.00002	0.01
arachidonoylcarnitine (C20:4)	0.06	0.00002	0.02
1-methylhistidine	0.05	0.00003	0.03
pyroglutamine[d]	0.07	0.00004	0.03
5alpha-androstan-3alpha,17beta-diol disulfate	0.08	0.00004	0.03
palmitoyl-arachidonoyl-glycerol (16:0/20:4)	0.08	0.00004	0.03
lysine[d]	0.02	0.00004	0.03
gamma-glutamylisoleucine	0.04	0.00004	0.03
argininate	0.09	0.00005	0.04
1-myristoyl-2-palmitoyl-GPC (14:0/16:0)	0.05	0.00005	0.04
1-carboxyethylleucine	0.06	0.00005	0.04
xanthosine	0.07	0.00006	0.047

The bolded EMCs were evaluated in our previous targeted analysis.[6]
EMCs, endogenous metabolomic compounds; FDR, false discovery rate.
[a]Kyoto Encyclopedia of Genes and Genomes, Human Metabolome Database, or PubChem; EMCs with (#) represent isomers of the same compound. [b]Beta coefficient and *P* value from univariate linear regression models evaluating the association between log transformed EMCs and busulfan clearance. [c]False discovery rate (Benjamini-Hochberg). [d]Not statistically associated with busulfan clearance in our previous targeted pharmacometabolomic analysis.[6] [e]Statistically associated ($P = 0.006$) with busulfan clearance in our previous targeted pharmacometabolomic analysis.[6]

busulfan clearance; our findings are suggestive of this paradigm. In this prospective cohort, plasma acetylcarnitine, phenylacetyl carnitine, and phenylacetylglutamine were not statistically associated with busulfan clearance. However, the pathway enrichment analyses revealed that lysine degradation had the lowest *P* value (7.27×10^{-7}) in both the 2-week pre-busulfan and pre-busulfan sample. Lysine degradation occurs through two pathways: (i) via formation of saccharopine, which is a pathway confined to the mitochondria; (ii) via the pipecolic acid pathway, which is a pathway not yet fully elucidated and known enzymes are localized in the mitochondria, cytosol, and peroxisome. The lysine degradation pathway does include two sodium ion-dependent, low-affinity carnitine

Table 3 Two-week pre-busulfan sample's EMC association with busulfan clearance (FDR < 0.05)

EMCs[a]	Beta coefficient[b]	*P* value[b]	FDR[c]
androstenediol (3alpha, 17alpha) monosulfate (2)	0.12	4.00E-10	3.37E-07
androsterone glucuronide	0.16	2.40E-09	2.04E-06
5alpha-androstan-3beta,17beta-diol disulfate	0.20	6.80E-09	5.72E-06
creatinine[d]	0.04	4.70E-08	3.95E-05
mannonate	0.08	1.50E-07	0.0001
N-acetylcarnosine	0.08	2.20E-07	0.0002
epiandrosterone sulfate	0.17	6.90E-07	0.0006
N,N,N-trimethyl-alanylproline betaine (TMAP)	0.05	8.00E-07	0.0007
androstenediol (3beta,17beta) disulfate (2)	0.11	8.40E-07	0.0007
1-methylurate	0.16	8.70E-07	0.0007
dehydroepiandrosterone sulfate (DHEA-S)	0.14	1.10E-06	0.0009
androsterone sulfate	0.19	1.40E-06	0.001
androstenediol (3beta,17beta) monosulfate (1)	0.15	1.60E-06	0.001
androstenediol (3beta,17beta) disulfate (1)	0.14	1.80E-06	0.001
5alpha-androstan-3alpha,17beta-diol monosulfate (2)	0.11	2.00E-06	0.002
hydroxyasparagine	0.04	8.20E-06	0.007
imidazole lactate	0.05	8.50E-06	0.007
androstenediol (3alpha, 17alpha) monosulfate (3)	0.13	0.00002	0.01
urate[e]	0.03	0.00003	0.02
5-acetylamino-6-amino-3-methyluracil	0.18	0.00003	0.02
5alpha-pregnan-3beta,20beta-diol monosulfate (1)	0.10	0.00003	0.03
pregnenediol sulfate (C21H34O5S)	0.09	0.00005	0.04
androstenediol (3beta,17beta) monosulfate	0.11	0.00006	0.048
N-(2-furoyl)glycine	0.20	0.00006	0.048
2,3-dihydroxy-5-methylthio-4-pentenoate (DMTPA)	0.03	0.00006	0.048
etiocholanolone glucuronide	0.11	0.00007	0.049
3-carboxy-4-methyl-5-propyl-2-furanpropanoate (CMPF)	0.15	0.00007	0.049

The bolded EMCs were evaluated in our previous targeted analysis.[6]
EMCs, endogenous metabolomic compounds.
[a]Kyoto Encyclopedia of Genes and Genomes, Human Metabolome Database, or PubChem; EMCs with (#) represent isomers of the same compound. [b]Beta coefficient and *P* value from linear regression models evaluating the association between log transformed EMCs and busulfan clearance. [c]False discovery rate (Benjamini-Hochberg). [d]Statistically associated ($P = 0.0006$) with busulfan clearance in our previous targeted pharmacometabolomic analysis.[6] [e]Not statistically associated with busulfan clearance in our previous targeted pharmacometabolomic analysis.[6]

(SLC) transporters (i.e., SLC25A21 and SLC25A29), which do not overlap with the putative SLC transporters possibly involved in busulfan transport (i.e., SLC7A8 (rs7141505) and SLC22A4 (rs1050152)).[28] The lysine degradation pathway has no clear association with glutathione, but it does result in glycine formation.[29] Glycine is a nonessential amino acid that can be endogenously synthesized from serine, threonine or choline. Glycine has a role in the production of purines, bile acids, creatine, heme, and glutathione. However, glycine was not associated with busulfan clearance ($P = 0.64$). However, consistent with our retrospective analysis,[6] the glycine, serine, and threonine pathway was associated with busulfan clearance ($P = 0.006$, FDR = 0.032) with an impact = 0.51.

We found that the steroid biosynthesis and was associated with busulfan clearance. The GSTs, specifically GSTA1-1 and GSTA3-3, catalyze the conversion of androst-5-ene-3,17-dione to androst-4-ene-3.17-dione,[30] the latter of which is part of the steroid hormone biosynthesis pathway. GSTA1-1 mainly catalyzes the metabolism of busulfan.[22,23] Furthermore, the GSTs are regulated by the nuclear receptors constitutive androstane receptor (CAR) and retinoid X receptor (RXR).[31] EMCs within the steroid biosynthesis pathway are agonists for CAR, while RXR is stimulated by EMCs within the retinoid pathway. We hypothesize that variation in the EMCs within these two pathways modulate the activity of these nuclear receptors, subsequently affecting GST activity and thus busulfan clearance. It should be noted that the majority of these participants received the antiepileptic medication phenytoin to prevent busulfan-induced seizures. Human CAR mediates induction of *CYP2B6* gene expression by phenytoin;[32] however, it is not apparent whether phenytoin can induce hepatic GST and busulfan clearance as we have previously reviewed.[1] It is unclear how busulfan clearance is associated with purine and pyrimidine metabolism pathways, which are related to energy metabolism.

In recent years, the use of pre-dose EMC profiling[33,34] to predict drug response in patients has been evaluated. There are an

Table 4 EMCs selected in LASSO cross-validation from the pre-busulfan sample with ≥30% frequency with 500-fold permutation testing

EMCs	β coefficient (SE)[a]	Frequency[a]
Urate	3.64 (1.81)	0.68
Androstenediol (3beta,17beta) disulfate (2)[b]	2.00 (0.72)	0.67
Cortisone	−3.07 (1.42)	0.65
Mannonate	2.55 (1.01)	0.56
Arachidonoylcarnitine	0.58 (0.75)	0.48
methyl-4-hydroxybenzoate sulfate	−0.06 (0.03)	0.46
5alpha-androstan (3alpha,17beta-diol) monosulfate (2)[b]	0.63 (0.32)	0.42
N-acetylcarnosine	0.57 (0.68)	0.38
N-acetyl-aspartyl-glutamate	−1.72 (0.71)	0.36
2-butenoylglycine	−1.16 (0.46)	0.36
Palmitoyl-arachidonoyl-glycerol	1.47 (0.48)	0.35
Cortisol	0.36 (1.26)	0.34
Lysine	3.10 (1.99)	0.30

The bolded EMCs were also significant in univariate analyses (**Table** 2).
EMCs, endogenous metabolomic compounds; LASSO, least absolute shrinkage and selection operator.
[a]Coefficients and frequency of EMCs selected with 500-fold permutation test. For example, urate was selected in 68 of the 500 permutations tested using LASSO. [b]EMCs with (#) represent isomers of the same compound.

increasing number of pharmacometabolomic studies (summarized in **Table S1** of McCune *et al.*[10]) examining the association of EMCs with subsequent clinical outcomes among allogeneic HCT recipients. We recently reported that relapse was associated with the cysteine/methionine pathway and the glycine, serine, and threonine metabolism pathway.[10] The latter can be explained by the fact that glutathione S-transferases conjugate both busulfan and glutathione, which contains glycine as a component. The d-arginine and

Table 5 Top pathways, significance, and impact from pathway enrichment analyses of the pre-busulfan sample's EMCs

Pathway name	Total[a]	Matched[b]	*P* value[c]	FDR[d]	Impact[e]
Lysine degradation[f]	25	7	4.49E-07	2.92E-05	0.16
Steroid hormone biosynthesis	85	7	0.0002	0.003	0.05
Purine metabolism	65	11	0.0003	0.005	0.04
Glycine, serine, and threonine metabolism	33	12	0.002	0.03	0.72
Pyrimidine metabolism	39	11	0.003	0.03	0.35
Aminoacyl-tRNA biosynthesis	48	20	0003	0.03	0.17
Cysteine and methionine metabolism	33	7	0.003	0.03	0.26
Arginine and proline metabolism	38	11	0.004	0.03	0.40
D-Arginine and D-ornithine metabolism	4	2	0.005	0.03	0
Alanine, aspartate, and glutamate metabolism	28	14	0.006	0.03	0.72
Glutathione metabolism	28	6	0.007	0.03	0.18
Pantothenate and CoA biosynthesis	19	9	0.007	0.03	0.08
Fatty acid degradation	39	2	0.008	0.04	0
Glyoxylate and dicarboxylate metabolism	32	9	0.01	0.04	0.26
Valine, leucine, and isoleucine biosynthesis	40	7	0.01	0.045	0.04
Linoleic acid metabolism	5	2	0.01	0.045	1
Inositol phosphate metabolism	30	2	0.01	0.045	0.13
Fatty acid biosynthesis	47	3	0.01	0.045	0.01

Sorted by increasing *P* values. The bolded pathway was significantly associated with busulfan clearance in our previous targeted pharmacometabolomic analysis.[6]
CoA, coenzyme A; EMCs, endogenous metabolomic compounds; tRNA, transfer RNA.
[a]Total number of EMCs in the stated pathway. [b]Number of matched EMCs, explained in Statistical Analysis section. [c]Raw *P* value from the Global test for pathway significance. [d]False Discovery Rate (Benjamini-Hochberg). [e]Impact is the pathway impact value on busulfan clearance calculated from pathway topology analysis. [f]In the 2-week pre-busulfan sample EMCs, the lysine degradation was the only pathway that met the criteria for statistical significance (*P* value = 0.0003, FDR = 0.02, Impact = 0.16).

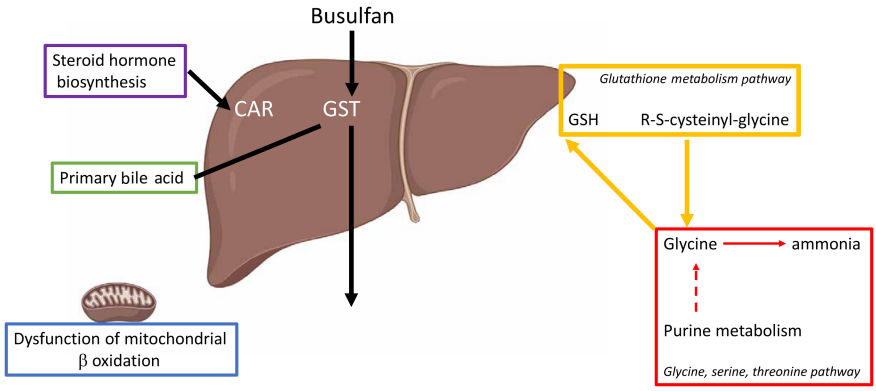

Figure 2 Working hypothesis for the pathways associated with busulfan clearance. Our working hypothesis is that glutathione metabolism (yellow box) continues to be central to busulfan clearance. We confirmed our prior findings of the association of the glycine, serine, and threonine pathway (red box) and bile acids (green box) with busulfan clearance. The impact of mitochondrial dysfunction (blue box) and reactive oxygen species was also found. CAR, constitutive androstane receptor; GST, glutathione transferase.

d-ornithine metabolism pathway and arginine and proline metabolism pathway were most associated with acute graft-vs-host disease.

Strengths of this work include the prospectively collected cohort of over 100 HCT recipients with a well-characterized i.v. busulfan PK data set, the global panel providing broad coverage of several pathways of interest, and consistency of robust results across three different statistical approaches. However, this study has limitations worth noting. Importantly, this prospective cohort had insufficient coverage of the glutathione pathway involved in busulfan metabolism. While 28 EMCs in the glutathione pathway were measured, only 6 had detectable signal in our plasma samples. Sufficient sensitivity was attained in three (i.e., spermidine, cysteine, and cysteinylglycine) EMCs not detected in our retrospective cohort. None of these EMCs were associated with busulfan clearance. However, five EMCs within the glutathione pathway (i.e., putrescine, spermine, glutathione, oxidized glutathione, and ascorbic acid) were not available in this prospective cohort.

In conclusion, we were able to select a subset of plasma EMCs measured pre-busulfan and construct a linear model that reasonably well predicts busulfan clearance among 2-week pre-busulfan samples, with the highest predictive power among patients with lower clearance. This suggests that pharmacometabolomics, quantitated *BEFORE* busulfan administration, could potentially estimate busulfan clearance to personalize the first dose to achieve the target busulfan AUC. We also discovered novel pathways associated with busulfan clearance. Further studies, including greater interrogation of the lysine degradation, steroid hormone biosynthesis, and glutathione pathway, are needed to validate these results.

SUPPORTING INFORMATION

Supplementary information accompanies this paper on the *Clinical Pharmacology & Therapeutics* website (www.cpt-journal.com).

ACKNOWLEDGMENTS

We are grateful to the patients who participated in this study. We are also grateful to the healthcare providers caring for the participants. The authors wish to thank the research staff (Sheri Ballard, Meagan Bemer, Genecelle Delossantos, Michael Donahue, and Jeffrey Stevens) involved in sample acquisition, sample transport, and data management.

FUNDING

This publication was supported by the National Institutes of Health under the Award Numbers: R01CA182963, R01GM129863, P01CA18029, 5P30CA015704, and UL1 TR002319. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

CONFLICT OF INTEREST

The authors declared no competing interests for this work.

AUTHOR CONTRIBUTIONS

All authors wrote the manuscript. J.S.M., K.S.B., and T.W.R. designed the research. J.S.M., K.S.B., L.J.R., B.R.P., L.S., G.S., H.J.D., and A.M. performed the research. J.S.M., S.L.N., T.W.R., Y.Z., and A.D.R.H. analyzed the data. No new reagents/analytical tools.

1. Palmer, J. *et al*. Personalizing busulfan-based conditioning: considerations from the American Society for Blood and Marrow Transplantation Practice Guidelines Committee. *Biol. Blood Marrow Transplant.* **22**, 1915–1925 (2016).
2. McCune, J.S. *et al*. Harmonization of busulfan plasma exposure unit (BPEU): a community-initiated consensus statement. *Biol. Blood Marrow Transplant.* **25**, 1890–1897 (2019).
3. Yeh, R.F. *et al*. Accurate targeting of daily intravenous busulfan with 8-hour blood sampling in hospitalized adult hematopoietic cell transplant recipients. *Biol. Blood Marrow Transplant.* **18**, 265–272 (2012).
4. McCune, J.S. *et al*. Variation in prescribing patterns and therapeutic drug monitoring of intravenous busulfan in pediatric hematopoietic cell transplant recipients. *J. Clin. Pharmacol.* **53**, 264–275 (2013).
5. Lin, Y.S., Kerr, S.J., Randolph, T., Shireman, L.M., Senn, T. & McCune, J.S. Prediction of intravenous busulfan clearance by endogenous plasma biomarkers using global pharmacometabolomics. *Metabolomics* **12**, 161 (2016).
6. Navarro, S.L., Randolph, T.W., Shireman, L.M., Raftery, D. & McCune, J.S. Pharmacometabonomic prediction of busulfan clearance in hematopoetic cell transplant recipients. *J. Proteome Res.* **15**, 2802–2811 (2016).

7. Singh, S.V., Leal, T. & Awasthi, Y.C. Inhibition of human glutathione S-transferases by bile acids. *Toxicol. Appl. Pharmacol.* **95**, 248–254 (1988).
8. Hassan, M. *et al*. Pharmacokinetics of high-dose busulphan in relation to age and chronopharmacology. *Cancer Chemother. Pharmacol.* **28**, 130–134 (1991).
9. McCune, J.S. *et al*. Busulfan in infant to adult hematopoietic cell transplant recipients: a population pharmacokinetic model for initial and Bayesian dose personalization. *Clin. Cancer Res.* **20**, 754–763 (2014).
10. McCune, J.S. *et al*. Prediction of acute graft versus host disease and relapse by endogenous metabolomic compounds in patients receiving personalized Busulfan-based conditioning. *J. Proteome Res.* **20**, 684–694 (2021).
11. Sumner, L.W. *et al*. Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics.* **3**, 211–221 (2007).
12. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate – a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B.* **57**, 289–300 (1995).
13. Meinshausen, N. & Bühlmann, P. Stability selection. *J. R. Stat. Soc. B.* **72**, 417–473 (2010).
14. Pang, Z. *et al*. MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. *Nucleic Acids Res.* **49**, W388–W396 (2021).
15. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
16. Jewison, T. *et al*. SMPDB 2.0: big improvements to the small molecule pathway database. *Nucleic Acids Res.* **42**, D478–D484 (2014).
17. Wishart, D.S. *et al*. HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res.* **46**, D608–D617 (2018).
18. Goeman, J.J., van de Geer, S.A., de Kort, F. & van Houwelingen, H.C. A global test for groups of genes: testing association with a clinical outcome. *Bioinformatics* **20**, 93–99 (2004).
19. Aittokallio, T. & Schwikowski, B. Graph-based methods for analysing networks in cell biology. *Brief. Bioinform.* **7**, 243–255 (2006).
20. Food and Drug Administration. Guidance for Industry: Bioavailability and Bioequivalence Studies for Orally Administered Drug Products — General Considerations. (2002). <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/bioavailability-and-bioequivalence-studies-submitted-ndas-or-inds-general-considerations>. Accessed November 22, 2022.
21. Minto, C. & Schnider, T. Expanding clinical applications of population pharmacodynamic modelling. *Br. J. Clin. Pharmacol.* **46**, 321–333 (1998).
22. Czerwinski, M., Gibbs, J.P. & Slattery, J.T. Busulfan conjugation by glutathione S-transferases alpha, mu, and pi. *Drug Metab. Dispos.* **24**, 1015–1019 (1996).
23. Ritter, C.A., Bohnenstengel, F., Hofmann, U., Kroemer, H.K. & Sperker, B. Determination of tetrahydrothiophene formation as a probe of in vitro busulfan metabolism by human glutathione S-transferase A1-1: use of a highly sensitive gas chromatographic-mass spectrometric method. *J. Chromatogr. B Biomed. Sci. Appl.* **730**, 25–31 (1999).
24. Deponte, M. Glutathione catalysis and the reaction mechanisms of glutathione-dependent enzymes. *Biochim. Biophys. Acta* **1830**, 3217–3266 (2013).
25. Bleyzac, N. *et al*. Improved clinical outcome of paediatric bone marrow recipients using a test dose and Bayesian pharmacokinetic individualization of busulfan dosage regimens. *Bone Marrow Transplant.* **28**, 743–751 (2001).
26. McCune, J.S. *et al*. Quality control of busulfan plasma quantitation, modeling, and dosing: an Interlaboratory Proficiency Testing Program. *Ther. Drug Monit.* **43**, 657–663 (2021).
27. Kim, B. *et al*. Pharmacometabolomics for predicting variable busulfan exposure in paediatric haematopoietic stem cell transplantation patients. *Sci. Rep.* **7**, 1711 (2017).
28. Lawson, R., Staatz, C.E., Fraser, C.J. & Hennig, S. Review of the pharmacokinetics and pharmacodynamics of intravenous busulfan in paediatric patients. *Clin. Pharmacokinet.* **60**, 17–51 (2021).
29. Wang, W., Wu, Z., Dai, Z., Yang, Y., Wang, J. & Wu, G. Glycine metabolism in animals and humans: implications for nutrition and health. *Amino Acids* **45**, 463–477 (2013).
30. Sharma, R., Shakell Ansari, G.A. & Awasthi, Y.C. *Physiological Substrates of Glutathione S-Transferases* (CRC Press, Boca Raton, LA, 2010).
31. Aleksunes, L.M. & Klaassen, C.D. Coordinated regulation of hepatic phase I and II drug-metabolizing genes and transporters using AhR-, CAR-, PXR-, PPARalpha-, and Nrf2-null mice. *Drug Metab. Dispos.* **40**, 1366–1379 (2012).
32. Wang, H., Faucette, S., Moore, R., Sueyoshi, T., Negishi, M. & LeCluyse, E. Human constitutive androstane receptor mediates induction of CYP2B6 gene expression by phenytoin. *J. Biol. Chem.* **279**, 29295–29301 (2004).
33. Clayton, T.A. *et al*. Pharmaco-metabonomic phenotyping and personalized drug treatment. *Nature* **440**, 1073–1077 (2006).
34. Clayton, T.A., Baker, D., Lindon, J.C., Everett, J.R. & Nicholson, J.K. Pharmacometabonomic identification of a significant host-microbiome metabolic interaction affecting human drug metabolism. *Proc. Natl. Acad. Sci. USA.* **106**, 14728–14733 (2009).