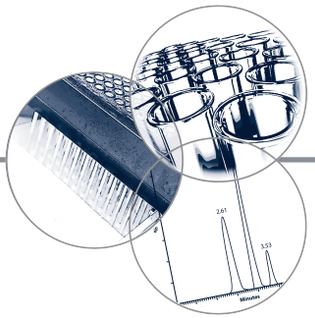
RESEARCH ARTICLE



Identification of drug metabolites in human plasma or serum integrating metabolite prediction, LC–HRMS and untargeted data processing

Background: Comprehensive identification of human drug metabolites in first-in-man studies is crucial to avoid delays in later stages of drug development. We developed an efficient workflow for systematic identification of human metabolites in plasma or serum that combines metabolite prediction, high-resolution accurate mass LC–MS and MS vendor independent data processing. Retrospective evaluation of predictions for 14 ^{14}C-ADME studies published in the period 2007–January 2012 indicates that on average 90% of the major metabolites in human plasma can be identified by searching for accurate masses of predicted metabolites. Furthermore, the workflow can identify unexpected metabolites in the same processing run, by differential analysis of samples of drug-dosed subjects and (placebo-dosed, pre-dose or otherwise blank) control samples. To demonstrate the utility of the workflow we applied it to identify tamoxifen metabolites in serum of a breast cancer patient treated with tamoxifen. **Results & Conclusion:** Previously published metabolites were confirmed in this study and additional metabolites were identified, two of which are discussed to illustrate the advantages of the workflow.

Background

Drug metabolism is the biochemical modification of pharmaceutical substances upon intake, generally resulting in the formation of more soluble compounds that are excreted via bile or urine. Circulating metabolites may contribute to pharmacological activity on the primary target, as well as off-target activity [1,2]. Although drug metabolism generally results in detoxification, it sometimes leads to toxic intermediates [3]. Such toxicity is screened for in preclinical *in vitro* and animal studies.

Because of its high selectivity and sensitivity, LC–MS/MS has become the method of choice for the quantitative determination of metabolites in biological samples [4–7] and high-resolution accurate mass LC–MS the standard for metabolite identification [8,9].

"Metabolites identified only in human plasma, or at disproportionately higher levels than in any of the animal test species, should receive special consideration in safety assessment". This is laid down in the US FDA/ICH Metabolites in Safety Testing (MIST) regulation, which states that *"nonclinical characterization of a human metabolite(s) is only warranted when that metabolite(s) is observed at exposures >10% of total drug-related exposure and at significantly greater levels in humans than the maximum exposure seen in the toxicity studies. Such studies should be conducted to support Phase III clinical trials"* [10,101,102].

A holistic strategy for characterizing the safety of metabolites through drug discovery and development has been the topic of a recent review [11], with other reviews discussing strategies on preclinical safety studies [12] and the role of radiolabelled mass-balance excretion and metabolism studies after the introduction of the FDA/ICH guidelines [13].

Current limitations

Major human metabolites may be missed by traditional 'cold' or 'hot' *in vitro* studies, as well as in 'cold' first-in-man studies as the analysis depends on the knowledge and focus of the operator. MS vendor software packages for metabolite identification, including AB SCIEX LightSight, Waters MetaboLynx XS and Thermo MetWorks, provide user-adaptable lists of mass differences corresponding to common biotransformations. In such generic lists, >1000 entries are needed to cover 80% of the unique elemental formula from literature-published biotransformations [14]. The large variety of dealkylations and hydrolysis reactions that might occur within the chemical space of metabolism of new chemical entities

Peter L Jacobs[1,2], Lars Ridder[3,4], Marco Ruijken[5], Hilde Rosing[6], Nynke GL Jager[6], Jos H Beijnen[6], Richard R Bas[1], William D van Dongen*[1]

[1]TNO Triskelion BV, Utrechtseweg 48, 3704 HE Zeist, The Netherlands
[2]MSam, Oss, The Netherlands
[3]Laboratory of Biochemistry, Wageningen University, Wageningen, The Netherlands
[4]Netherlands eScienceCenter, Amsterdam, The Netherlands
[5]MsMetrix, Maarssen, The Netherlands
[6]Slotervaart Hospital, The Netherlands Cancer Institute, Amsterdam, The Netherlands
*Author for correspondence:
E-mail: wdvandongen@gmail.com

FUTURE SCIENCE part of fsg

Key Term

Expected metabolite searches: Searches based on *in silico* prediction, frequently observed mass differences or spectral characteristics (biased).

in drug discovery would make such a list 'endless' and prone to frequent updates. Although generic lists are helpful for the operators, such long lists increase the risk of false positive results in **expected metabolite searches** as they also contain many entries irrelevant for the chemical structures of candidate drugs studied in a specific project. False-negative results for major human metabolite(s) in the cold first-in-man study result in serious delays during clinical Phase III, when they are ultimately identified in ^{14}C human ADME studies and would require further toxicological assessment.

Proposed workflow

We describe a systematic workflow for human metabolite identification in plasma or serum from first-in-man studies that reduces the risk of missing metabolites at this important stage in the drug-development process [15]. The workflow can be divided into several key steps (Figure 1):

The first step in the presented workflow is metabolite prediction, a systematic *in silico* generation of potential metabolites. Integrated with MS vendor-independent data processing, a key difference when compared with most workflows used in metabolite identification laboratories.

Using structure-based metabolite prediction software to generate lists of changes in the elemental formula has the advantages of covering more biotransformations and yielding a more rational list of mass differences, limited to those relevant for the structure at hand [16–18], and thereby reducing the risk of false-positive hits. Numerous methods for the prediction of metabolite structures have been reviewed recently [19]. Some of these methods are used in MS data processing, including Meteor Nexus (restricted to members of Lhasa Ltd) and Mass-MetaSite (Molecular Discovery Ltd) for prediction of metabolic transformations related to cytochrome-mediated reactions in Phase I metabolism. Here, we use an updated and extended set of the SyGMa biotransformation rules [16], which allow efficient enumeration of Phase I and II metabolites that can result from several subsequent reaction steps. The predicted metabolites are ranked on the basis of 'success rates' (scores) of the corresponding rules applied on a training set of observed human biotransformation reactions.

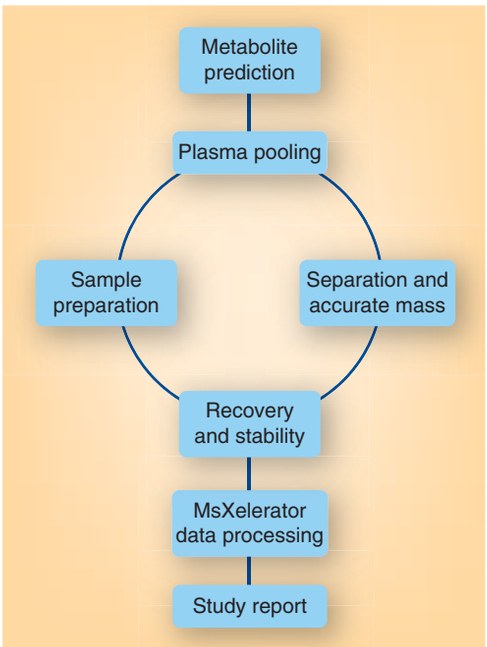


Figure 1. *i*Humite® workflow for plasma (or serum) metabolite identification.

The changes in elemental formula of the predicted metabolites, as well as the calculated monoisotopic mass differences, are used for a systematic and very sensitive targeted screening in pooled human plasma or serum and their coverage in preclinical species.

In the second step of the workflow 'Hamilton' pooling is used to create $AUC_{0\text{-}24h}$ pools of plasma or serum of drug-dosed subjects and controls [20]. In a typical first-in-man study design, $AUC_{0\text{-}24h}$ pools of multiple subjects are combined and metabolite profiles for pools on day 1 and, for example, day 10 are compared to check accumulation of circulating metabolites after multiple dosing [21].

In subsequent steps plasma or serum pools of drug-dosed subjects are protein precipitated or extracted and analyzed with high-resolution LC–MS, both to confirm predicted metabolites and to search for any additional metabolites. The latter is done with differential analysis where samples of drug-dosed subjects are compared with those of placebo-dosed subjects, pre-dose samples and/or blank matrix (controls). For major or disproportionately circulating human metabolites, Ion Trap MS^n and Orbitrap accurate mass experiments are performed to support follow-up activities, including chemical synthesis and/or activity testing. Parent and metabolite stability during sample clean-up and freeze–thaw experiments should be warranted and recovery goals met.

In the data-processing step MS vendor-independent software is used (MsXelerator, MsMetrix, Maarssen, The Netherlands), integrating this

part of the workflow with the metabolite prediction step. The data-processing software package reads the predicted metabolite set and reports the combined results of both the targeted (predicted) and nontargeted (unexpected) metabolite searches. All results from peak picking are checked in high resolution mode against a control pool. When necessary and applicable, the results can also be checked on the expected isotopic pattern and mass defect range based on the parent drug. Observed major and human specific metabolites, if any, are used to decide on the need for nonclinical characterization.

■ Tamoxifen

MIST regulation applies to recent and future new drug applications. As a consequence, human plasma metabolite identification experiments have not been performed to that standard for all drugs currently prescribed to patients. For some drugs, full-scan accurate mass LC–MS data with state-of-the-art equipment are lacking as initial metabolism data were obtained decades ago. Tamoxifen was selected for an evaluation of our workflow because of its extensive metabolism, while serum data from a state-of-the-art mass balance study were not present in the literature [22].

Tamoxifen is a first-generation selective estrogen-receptor modulator used in breast cancer treatment and as a chemopreventive agent. The metabolism of [^{14}C-]tamoxifen in female patients was first reported 40 years ago by ICI, now AstraZeneca [23]. Of at least 22 reported Phase I metabolites in humans, 4-hydroxytamoxifen and *N*-desmethyl-4-hydroxytamoxifen (endoxifen) are considered to be the most therapeutically active. Bioanalytical methods for determination of tamoxifen and its Phase I metabolites have been reviewed recently [24]. Assays to support therapeutic drug monitoring apply low-resolution SRM [25].

In this article, we apply the workflow to identify metabolites of tamoxifen in steady state serum samples (50 µl) of a breast cancer patient using high-resolution accurate mass LC–MS. We demonstrate that the workflow can identify both predicted and unexpected metabolites in a single data processing run.

Methods & materials

■ Drug metabolite prediction

Potential drug metabolites were generated on the basis of reaction rules representing most common Phase I and II human biotransformations, applied for three subsequent reaction steps. The rules were derived from biotransformations observed in man, extending the approach followed for the SyGMa method [16]. A total of 136 Phase I biotransformation rules were included, covering a wide range of reactions, including dealkylation, hydroxylation, dehydrogenation, dehydration, deamination, hydrolysis and several other types of oxidations. A total of 27 Phase II rules were included for conjugation of glucuronide, sulphate, phosphate, acetyl and methyl groups. The resulting list of metabolites and their calculated monoisotopic masses were used to perform a systematic (targeted) screen for the presence of possible metabolites in pooled human plasma and their coverage in preclinical species.

■ Accurate MS measurements

Accurate MS measurements have been acquired on a hybrid LTQ Orbitrap® mass spectrometer (Thermo Scientific, Breda, The Netherlands) at high mass accuracy (30,000 mass resolution in MS mode and 7500 in MS^2 mode) using the Orbitrap mass scan range 200–1200 Da in MS mode and 100–400 Da for HR Ion Trap MS^2. In most MS^2 LTQ experiments 35 eV CID was used with the mass scan range 105–400 Da. The number of sequential product ion scans with Orbitrap mass detection was limited to one or two on the first-generation LTQ Orbitrap used, to warrant sufficient full scans over the HPLC peak widths in the gradient used. Serum samples were re-injected to obtain good quality data for the precursor ions *m/z* 388, 374, 358 and 344. HPLC–MS and sample processing conditions for tamoxifen and its metabolites were based on a selective SRM HPLC–MS quantification serum assay developed at the Bioanalytical Laboratory of Slotervaart Hospital [25]. The HPLC gradient on the Thermo Accela was started at 10% acetonitrile in the HRM experiments, instead of 30% in the SRM method, to facilitate retention of the more polar metabolites including glucuronides.

Serum samples (50 µl) were obtained from a breast cancer patient, treated at The Netherlands Cancer Institute, Amsterdam, with a daily dose of 20 mg tamoxifen for at least 2 months to ensure steady-state concentrations. As a result, no $AUC_{0\text{-}24h}$ pooling of patient serum was needed and a single time point was used. Control samples were obtained from a serum pool of non-treated human subjects. This control serum was also used to spike a mix of available synthetic metabolite standards, including tamoxifen, *N*-desmethyltamoxifen, (*Z*)-*N*-desmethyl-4-hydroxytamoxifen (*Z*-endoxifen), 4-hydroxytamoxifen,

Key Terms

Non-targeted (unexpected) metabolite searches: Searches based on differential analysis (unbiased)

Major human plasma metabolites: Metabolites observed at exposure >10% of total drug-related exposure

Table 1. Selected ^{14}C-labeled compounds with human plasma data published.

No.	Compound	Mode of action	Main plasma metabolites	Ref.
1	Mirabegron	β3-Adrenoceptor agonist	*O*-glucuronide†	[34]
			N-carbamoyl-glucuronide‡	
			N-glucuronide†	
2	Brivanib alaninate	Dual inhibitor of VEGF and FGF	*O*-dealkyl, sulfate, carboxyl†	[28]
3	Peliglitazar	Dual α/γ peroxisome proliferator-activated receptor activator	1-*O*-β-acyl-glucuronide†	[36]
4	SB-649868	Orexin 1 and 2 receptor antagonist	Hemiaminal‡	[30]
5	Sunitinib	Oral multi-targeted tyrosine kinase inhibitor	Des-ethyl†	[33]
6	BMS-690514	ErbB/VEGF-receptor inhibitor	*O*-glucuronide†	[27]
			Hydroxylated rearrangement product§	
			O-glucuronide†	
			Glucuronide (position unknown)†	
7	INCB018424	Selective janus tyrosine kinase 1/2 Inhibitor	2-hydroxy-cyclopentyl§	[32]
8	Lersivirine	Next-generation non-nucleoside reverse transcriptase inhibitor	Glucuronide†	[35]
9	Stavudine	Orally active nucleoside reverse transcriptase inhibitor	+*O* +glucuronide†	[38]
10	Apixaban	Reversible and direct inhibitor of coagulation factor Xa	*O*-demethyl sulfate†	[37]
11	Bazedoxifene	SERM	Indole glucuronide†	[26]
12	Vildagliptin	Dipeptidyl peptidase 4 inhibitor	Carboxylic acid†	[29]
13	Brasofensine	Inhibitor of the synaptic dopamine transporter	*O*-demethyl†	[39]
			O-demethyl (isomer)†	
14	Brivaracetam	SV2A ligand	*N*-propyl side chain hydroxyl†	[31]

^{14}C in title of publication in Drug Metabolism and Disposition; 2007–January 2012.
†Predicted biotransformation.
‡Not predicted, elemental formula not present in the predicted set.
§Not predicted, elemental formula present in the predicted set.

N-desmethyl-4′-hydroxytamoxifen and 4′-hydroxytamoxifen. No placebo-dosed serum or pre-dose serum of patients was used as control. All serum samples were processed and HPLC separated according to the method described by Jager *et al.* [25]. LC–MS samples were prepared in duplicate and injected (at least) twice with the instrument sequence repeated the next day.

■ MS data processing

Detected chromatographic peaks in LC–MS data of *in vitro* or *in vivo* metabolism studies were compared against a control sample using MsXelerator processing software (MsMetrix). All results from peak picking were checked in high-resolution mode against a control sample. Findings were judged and filtered on a number of characteristics, including chromatographic peak shape, isotope signature match, product ion and neutral loss ion scanning relative to the parent drug and mass defect range based on the parent drug. The sequential application of multiple filters ensured high confidence that detected peaks were true metabolites. This software package also reads the predicted metabolite set and reports the combined results of both the targeted (predicted) and **non-targeted (unexpected) metabolite searches**.

■ Availability

The complete workflow is offered as a commercial service by TNO Triskelion, Zeist, The Netherlands, using the *i*Humite® trade mark. Metabolite predictions are developed in a collaboration between Wageningen University and The Netherlands eScience Center and will become publically available via the internet [103].

Results

■ Evaluation of metabolite prediction

To estimate the expected coverage of our targeted approach, we compared the metabolite

predictions with major metabolites observed in 14 human ^{14}C-ADME studies (TABLE I) published by seven different companies in the period 2007–January 2012 [26–39]. For three of the evaluation compounds [31,37,38], no **major human plasma metabolites** were reported, and therefore the most abundant metabolites in plasma were included in our selection.

Table 2. Rank 1–10 predicted metabolites of lersivirine.

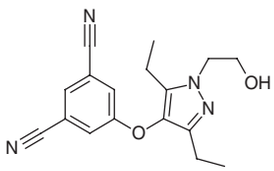
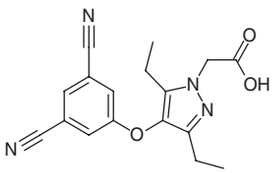
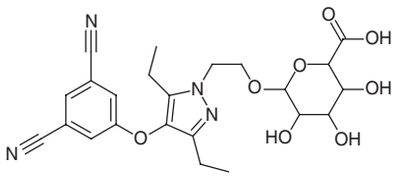
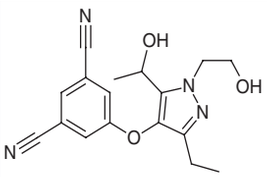
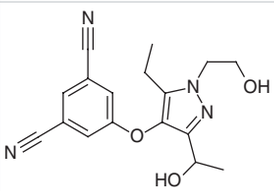
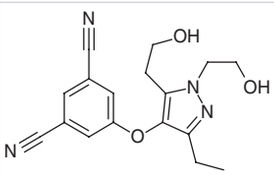
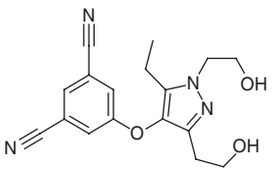
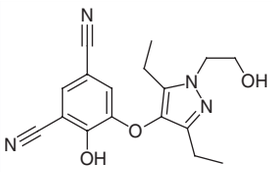
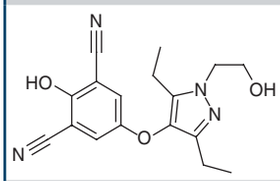
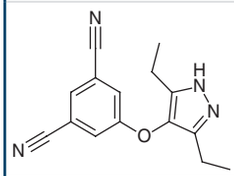
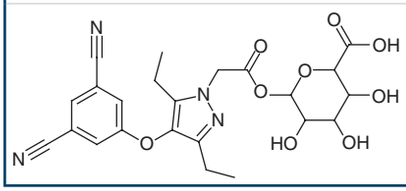
Structure	Biotransformations	Score [16]	Rank
	Parent		
	Primary alcohol oxidation	0.200	1
	*O*-glucuronidation	0.102	2
	Benzylic hydroxylation	0.073	3.5
	Benzylic hydroxylation	0.073	3.5
	Aliphatic hydroxylation	0.063	5.5
	Aliphatic hydroxylation	0.063	5.5
	Aromatic hydroxylation	0.060	7

Table 2. Rank 1–10 predicted metabolites of lersivirine (cont.).

Structure	Biotransformations	Score [16]	Rank
	Aromatic hydroxylation	0.059	8
	*N*-dealkylation	0.049	9
	Primary alcohol oxidation; *O*-Glucuronidation	0.032	10

As an example, Table 2 presents the top ten predicted metabolites for one of the evaluation compounds, lersivirine. The main observed human plasma metabolites, the *O*-glucuronide and the benzylic hydroxyl metabolite, are predicted on rank two and 3.5 (rank three and four shared by two biotransformations).

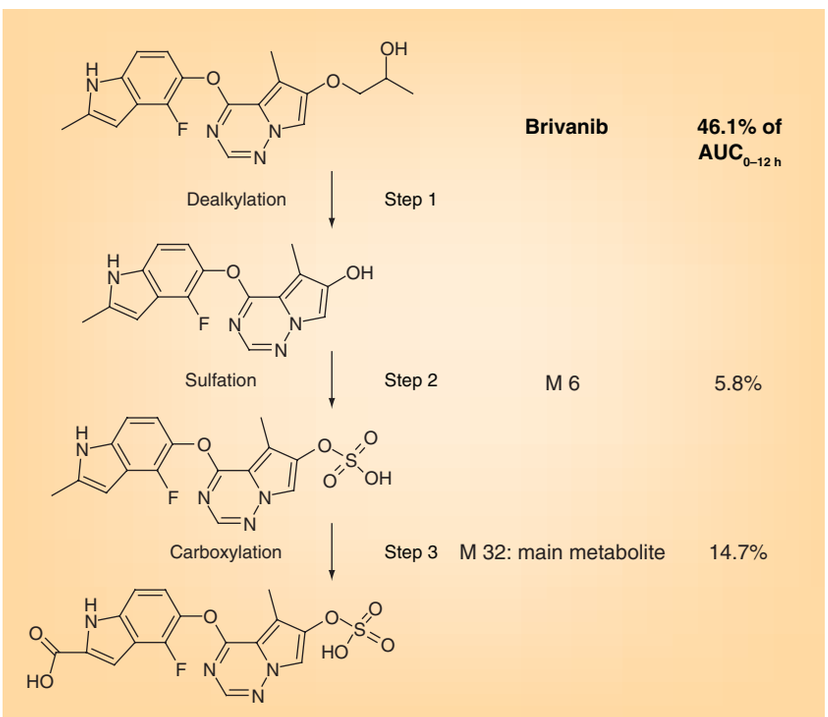


Figure 2. Multistep prediction for the main reported human plasma metabolite of ^{14}C-brivanib after oral administration of its alanine prodrug.

The main human plasma metabolite for Brivanib, M32, is the result of three subsequent biotransformation steps. Structures of both M32 and its intermediates are in the predicted metabolite set as indicated in Figure 2. In our approach, rank numbers will increase with the number of subsequent steps. As data processing is very fast, three-step metabolite prediction is routinely used for data processing.

Of the total set of 20 of the human plasma metabolites included in the evaluation, 18 (90%, including all of the 11 one-step metabolites) would have been found with extracted ion chromatograms derived from elemental formula of predicted metabolites when our workflow would have been applied to experimental data. Furthermore, 16 (80%) of the reported metabolite structures were in the predicted set (including all of the one-step metabolite structures).

■ Application to tamoxifen

Tamoxifen metabolites in human serum

High-resolution accurate mass full-scan LC–MS data files were obtained for steady state serum samples of a tamoxifen-treated patient, samples of a pool of blank control human serum, available standards, as well as a mix of standards spiked to the pool of blank human serum. In alternate experiments, ion trap product ions scans with accurate mass Orbitrap detection were obtained for *m/z* 388, 374, 358 and 344, to support structure assignments.

Rule-based metabolite prediction assigned the highest probability to *N*-demethylation of

tamoxifen (TABLE 3), which was in agreement with the experimental LC–HRMS results obtained in serum of the tamoxifen-dosed patients.

Other metabolites observed in serum of a patient, and absent in control serum using a threshold of both %Peak Height and %Area ≥ 1 after differential analysis, included (*Z*)-*N*-desmethyl-4-hydroxytamoxifen (*Z*-endoxifen), 4-hydroxytamoxifen, *N*-desmethyl-4′-hydroxytamoxifen and 4′-hydroxytamoxifen, which were predicted at rank 13.5 ((*Z*)-*N*-desmethyl-4-hydroxytamoxifen; *N*-desmethyl-4′-hydroxytamoxifen) and 9.5 (4-hydroxytamoxifen; 4′-hydroxytamoxifen). Other minor metabolites, detected by the combination of metabolite prediction and differential analysis, included *N,N*-didesmethyl tamoxifen and the *N*-oxide (rank two and three, respectively). These metabolite findings are in agreement with published data, as summarized in recent reviews[22,24].

An additional metabolite, not published before, was found with $[M+H]^+$ corresponding to elemental formula $C_{33}H_{39}NO_9$, a mass difference of +222 Da ($C_7H_{10}O_8$) with respect to the parent compound tamoxifen. Based on the formula and the list of predicted metabolites, this could be a combination of *O*-glucuronide and *O*-methyl conjugation (multiple biotransformation steps: rank 78, 79 and above). Additional product ion MS^n and H/D exchange could be used to further characterize this minor metabolite (1 %Area relative intensity after differential analysis; FIGURE 3).

Tamoxifen and the eight metabolites described (TABLE 4) were also found as a result of the targeted, predicted metabolite search without the need for differential analysis. To increase confidence results indeed match predicted metabolite structures, additional filters can be used in MsXelerator. One such filter confirms the presence of expected isotopic patterns, for example, the correct $^{13}C/^{12}C$ ratio. Isotope patterns for $[M+H]^+$ ions of both parent and metabolites found are shown in FIGURE 4.

To illustrate signal to noise in both serum sample and control for one of the minor metabolites, the result is shown for the *N,N*-didesmethyl metabolite (FIGURE 5), with a relative intensity of 3% Area after differential analysis.

The differential analysis procedure allows an even more sensitive approach for tamoxifen in steady state serum as is illustrated with results obtained for a 10× lower intensity metabolite that was also not described in the literature

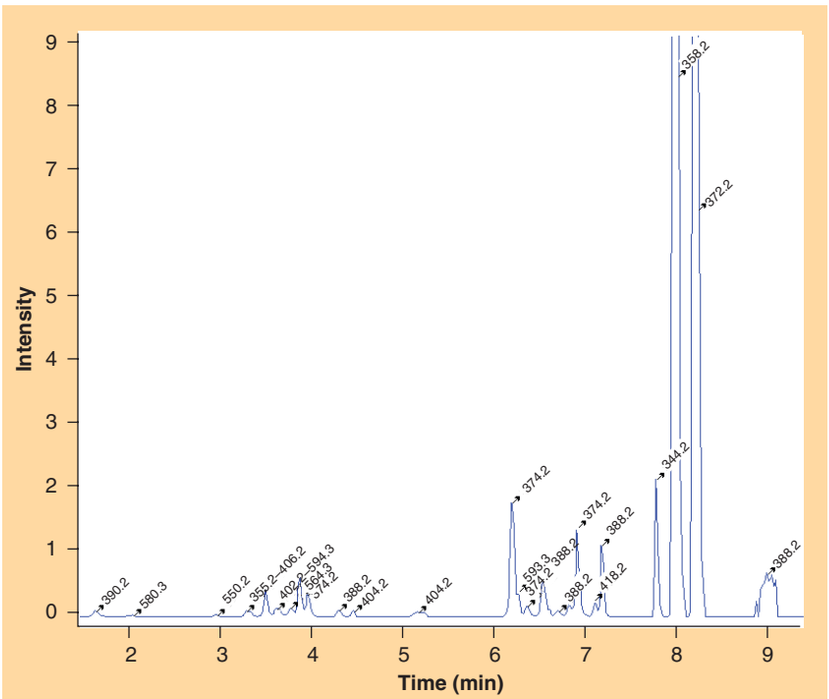


Figure 3. Reconstructed summed mass chromatogram after differential analysis of patient serum and serum controls (30 peaks).

before (0.3% Area after differential analysis, FIGURE 6). The elemental formula determined for the $[M+H]^+$, $C_{25}H_{26}NO_3$, is in agreement with a mass difference of +16 Da ($-CH_4 + O_2$) with respect to the parent compound (parent $[M+H]^+ = C_{26}H_{30}NO$). This could be the result of a combination of the modifications '*N*-demethylation' as well as '$-CH_3$ to $-COOH$' oxidation (rank 29).

To characterize this and other (minor) tamoxifen metabolites observed in more detail, additional product ion MS^n and H/D exchange experiments could be performed. However, final proof of chemical structure in drug metabolite identification is in a comparison with a synthetic standard as was conducted for six of the eight metabolites in TABLE 4. Synthesis also facilitates determination of the activity profile and decisions to be made for follow-up PK/PD or safety experiments.

Discussion

It is important to assure detection of metabolites in steady state serum, or $AUC_{0\text{-}24h}$ pooled plasma samples in general, with significant (e.g., >25%) contribution to the overall therapeutic and adverse effect. As potency of some circulating tamoxifen metabolites is 30–100-fold higher than the parent, it is a prerequisite for tamoxifen

metabolite identification to use a mass spectrometric approach with a detection limit <10% of the parent MS response.

For metabolite identification of drugs in general, differences in MS response (ESI efficiency) might further complicate reaching such

Table 3. Rank 1–14 predicted metabolites of Tamoxifen.

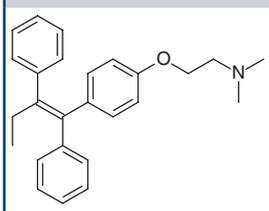
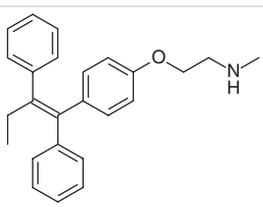
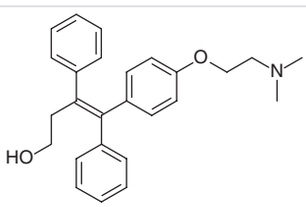
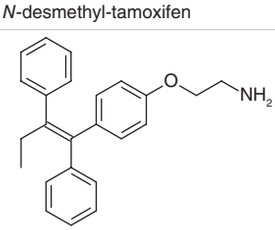
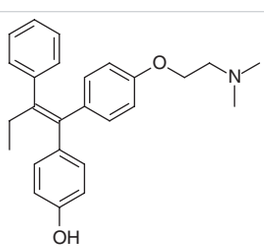
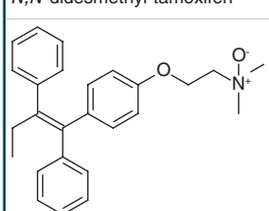
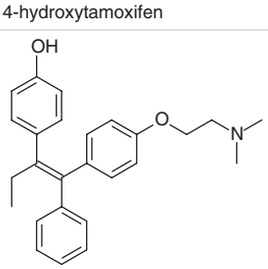
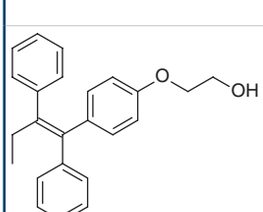
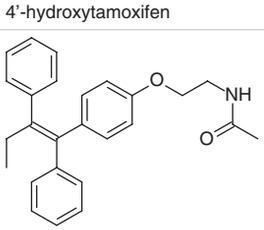
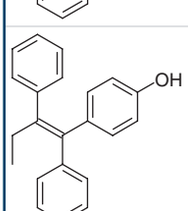
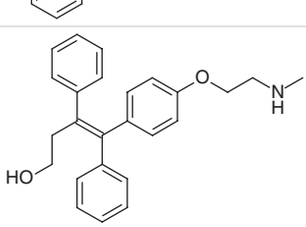
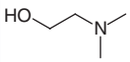
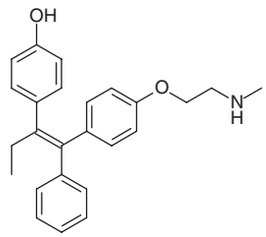
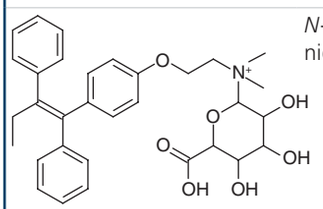
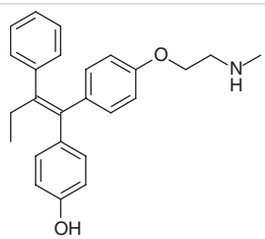
Structure	Biotransformations	Score	Rank	Structure	Biotransformations	Score	Rank
 Tamoxifen	Parent	1.000	0				
 *N*-desmethyl-tamoxifen	*N*-demethylation	0.590	1		Aliphatic hydroxylation	0.063	8.0
 *N*,*N*-didesmethyl-tamoxifen	*N*-demethylation *N*-demethylation	0.320	2	 4-hydroxytamoxifen	Aromatic hydroxylation	0.060	9.5
 Tamoxifen-*N*-oxide	*N*-oxidation	0.190	3	 4'-hydroxytamoxifen	Aromatic hydroxylation	0.060	9.5
	*N*-dealkylation	0.130	4		*N*-demethylation *N*-demethylation *N*-acetylation	0.049	11.0
	*O*-dealkylation	0.096	5.5		*N*-demethylation Aliphatic hydroxylation	0.037	12.0

Table 3. Rank 1–14 predicted metabolites of Tamoxifen (cont.).

Structure	Biotransformations	Score	Rank	Structure	Biotransformations	Score	Rank
	*O*-dealkylation	0.096	5.5	 *N*-desmethyl-4'-hydroxytamoxifen (*Z*-endoxifen)	*N*-demethylation Aromatic hydroxylation	0.035	13.5
	*N*-glucuronidation	0.095	7	 *N*-desmethyl-4-hydroxytamoxifen	*N*-demethylation Aromatic hydroxylation	0.035	13.5

detection limits for circulating metabolites. It has been shown that HPLC–ESI-MS response can overestimate metabolites by a factor of 20 when compared with radiochromatography. With nanospray LC–MS 'response normalization' was obtained applying a spray with constant solvent composition [40]. In a study of 14 drug candidates and >70 metabolites, it was shown that nanospray ESI–MS without 'response normalization' can overestimate metabolites by a factor of 12, and underestimate by a factor of 6, when compared with radiochromatography [41]. For some of the predicted metabolites of tamoxifen also (large) positive ion electrospray response differences can be expected as the result of deamination or dealkylations with loss of the nitrogen atom (e.g., metabolites with rank 4 resp. 5.5 in Table 3). These metabolites were not observed, with %Area ≥1 in our serum experiments.

Metabolite response can also depend on the biological matrix used, and therefore various matrix mixing and (radiolabeled) reference spiking methods have been described to minimize the ion suppression or ion enhancement effect on the quantitative estimation of circulating metabolite levels relative to the parent drug. In human plasma spiking experiments both samples from *in vitro* incubations and *in vivo* preclinical studies were used [21,42–44].

The good coverage of predicted metabolites, as shown in our literature exercise, results in less time to be spent on the much more difficult process of identification of unknown metabolites. One of the metabolites from the literature that was not predicted is an *N*-carbamoylglucuronide. This rare type of glucuronidation has been reported mostly during the last decade [45] and no examples were present in the training set on which the rules were based. Two other metabolites that were not predicted involved uncommon formations of a hemiaminal [30] and a hydroxylated rearrangement product [27]. On the other hand, one of the predicted metabolites in the evaluation set, *O*-demethylapixabansulfate [37], was disproportionately circulating in humans compared with animal species, illustrating the potential of the approach to identify compounds requiring additional experiments in an early clinical phase.

Conclusion

Evaluation of the metabolite predictions employed in the *i*Humite workflow, based on metabolite ID studies for 14 new chemical entities selected from the literature indicated a coverage of 90% of the major metabolites. Consequently, systematic and sensitive targeted screening of the predicted metabolites in pooled human plasma and in preclinical species is expected to result in a good coverage of the major metabolites. In an integrated approach the MsXelerator software reads the predicted set of metabolites and performs both targeted (predicted metabolites) and non-targeted (differential) data analysis. This data mining approach of LC–MS data obtained

Key Term

Active metabolites: Metabolites contributing to the therapeutic or adverse effect.

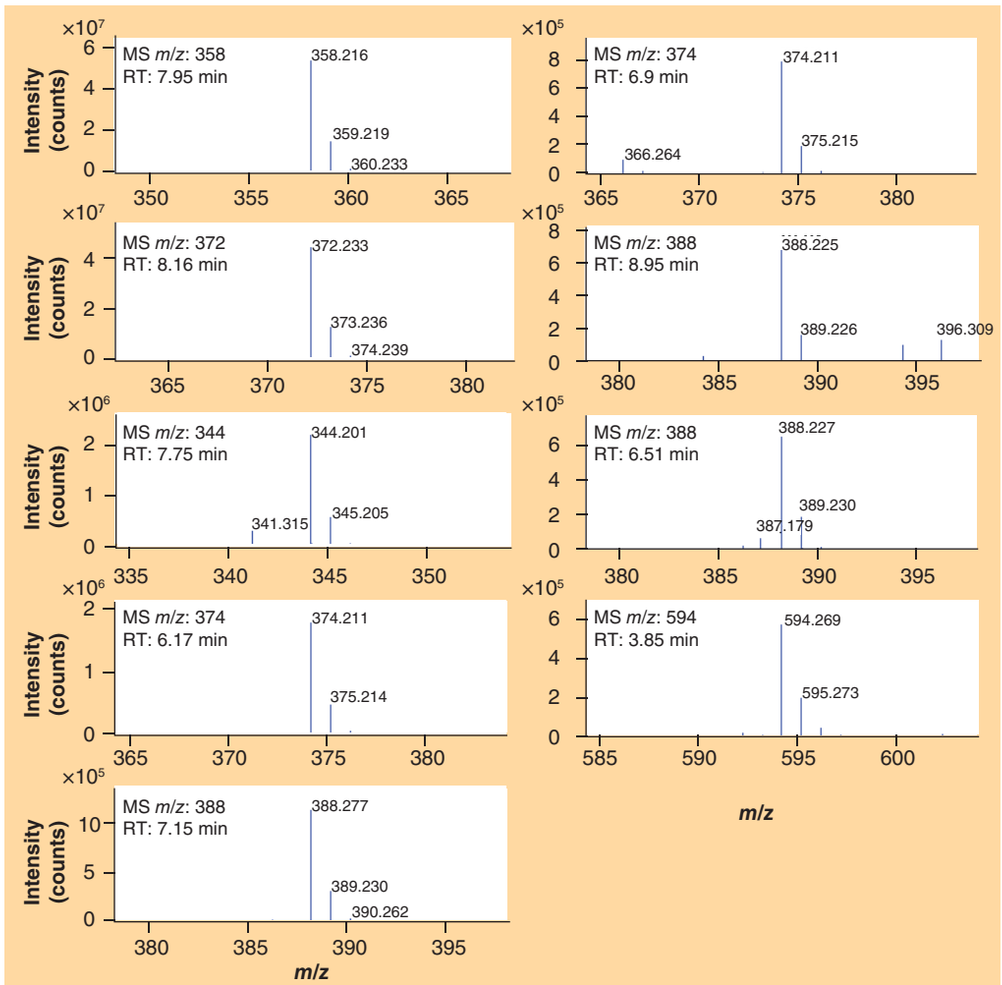


Figure 4. Partial mass spectra showing $[M+H]^+$ isotope patterns for tamoxifen (*m/z* 372) and eight metabolites found in the targeted approach after a search for predicted metabolites (sorted on ion intensity).

from both plasma or serum of drug- and placebo-dosed subjects, results in one combined table of observed metabolites. The *i*Humite workflow was successfully applied to steady state serum samples of a patient treated with tamoxifen. The combination of rule-based metabolite prediction, full-scan accurate mass and comprehensive data analysis resulted in the identification of known major and minor serum metabolites, as well as several additional minor metabolites (≤1 %Area relative intensity after differential analysis) that were not reported before.

Future perspective

Creation of a commonly accepted rule-base of biotransformations with application in first-in-man metabolite identification studies could help to assure that minimum criteria for metabolite identification are being met in all laboratories involved. As data processing results might become less biased by experience level of scientists involved, integration of quan/qual experimental bioanalytical groups could become standard practice in the future. A role for the bioanalytical and drug metabolism communities is foreseen to maintain and extend such a rule-base and discuss recommendation for implementation in revisions of the guidelines.

When recent hardware and software developments by the MS vendors mature, upfront recording of information-rich MS datasets, including derived spectra with in-depth structural information, might become standard practice in drug metabolite identification. A comparison of experimental MS^n trees with *in silico* generated fragments of chemical structures in the public domain was used as strategy to facilitate assignment of metabolite structures in metabolomics experiments [46,47]. In a similar way (confidential) compound databases within the pharmaceutical industry, together

with sets of predicted metabolite structures generated *in silico,* can be used to facilitate structural elucidation of drug metabolites.

The contribution of **active metabolites** to the overall therapeutic and adverse effect profile depends on its relative intrinsic potency and pharmacokinetic properties [1–3]. In the case of tamoxifen the most therapeutically active metabolites are (*Z*)-*N*-desmethyl-4-hydroxytamoxifen (*Z*-endoxifen) and 4-hydroxytamoxifen, being from 30- to 100-fold more potent than tamoxifen itself. The steady state level of endoxifen is a proposed predictor of the clinical outcomes of tamoxifen treatment [48]. However, this knowledge became only recently available, many years after initial application.

As a result of technological advances in MS in the recent past, detection limits as well as mass resolution improved and full-scan accurate mass LC–MS became feasible for the identification of circulating metabolites at levels only SRM detection could previously reach. This allowed ICH/FDA MIST guidelines to define the identification of major circulating human metabolites at levels >10% of total drug-related exposure. With further advances in the field it might become possible to aim at identification of metabolites with significant (e.g., >25%) contribution to the overall therapeutic (and adverse) effect. This could have a positive impact on future early-stage therapeutic decision making for drugs entering the market.

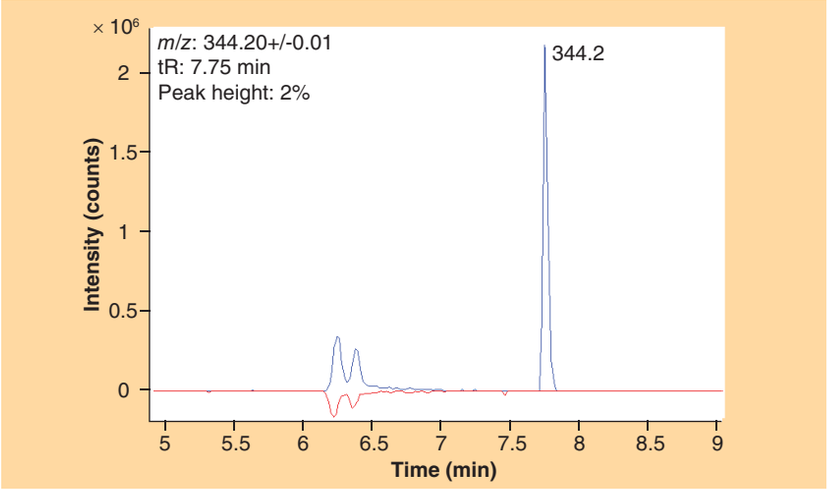


Figure 5. Example of differential analysis for patient serum (upper trace) and control (lower trace). Extracted-ion chromatogram analysis over the full mass range, resulting in the detection of *m/z* 344.20 that is assigned to *N,N*-didesmethyl (% Area = 3). Accurate EIC of control (lower trace), plotted in mirror mode.

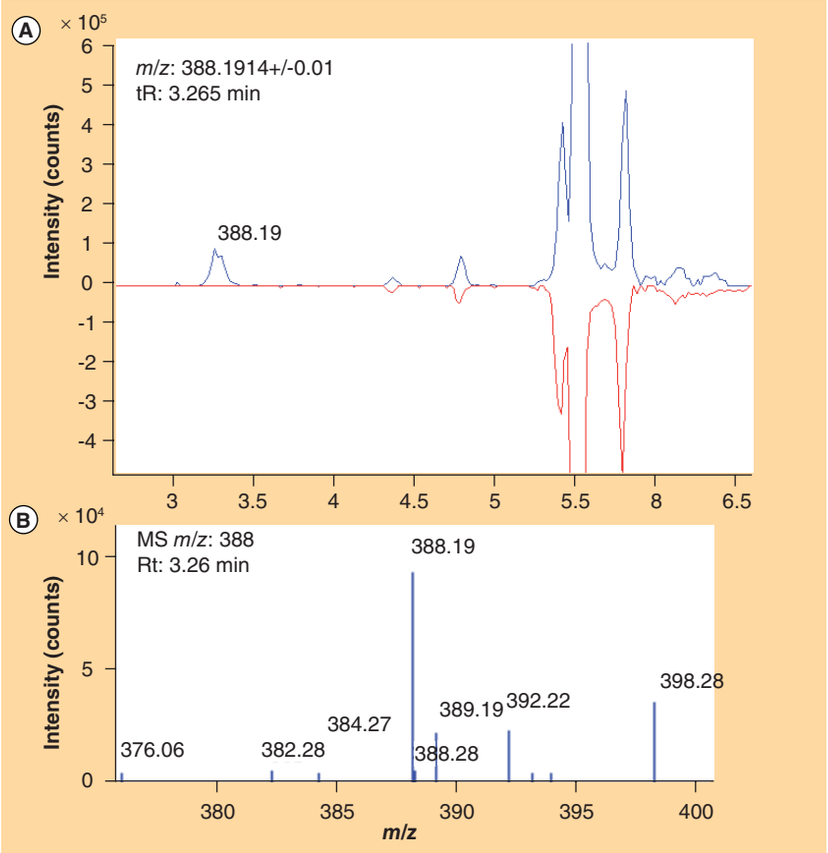


Figure 6. Example of differential analysis for patient serum (upper trace in upper pane) and control (lower trace). (A) Extracted-ion chromatogram analysis over the full mass range, resulting in the detection of *m/z* 388.19 that is tentatively assigned to *N*-demethylation in combination with R–CH_3 to R-COOH oxidation (%Area = 0.3). **(B)** Corresponding partial mass spectrum of the metabolite at Rt = 3.26 min.

Financial & competing interests disclosure

PL Jacobs, RR Bas and WD van Dongen disclose that they work for a CRO that supports the industry in the topics discussed, and hence derive financial benefit from that relationship in terms of consultancy fees (PL Jacobs) and salaries (RR Bas & WD van Dongen). M Ruijken discloses that he develops software that supports the industry in the topic discussed, and hence derives financial benefit from software sales. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Ethical conduct of research

The authors state that they have obtained appropriate institutional review board approval or have followed the principles outlined in the Declaration of Helsinki for all human or animal experimental investigations. In addition, for investigations involving human subjects, informed consent has been obtained from the participants involved.

Table 4. Peaks with both %Peak height and %Area ≥1 after differential analysis of tamoxifen-treated patient serum and serum controls (parent and eight metabolites out of a total of 30 differential peaks sorted on %Peak height).

Retention time (min)	% Peak height	%Area	[Mass difference]	Formula	$[M+H]^+$ Parent and main metabolites	Identification on HRMS and retention time	Identified with synthetic reference
7.947	100	100	[-14.01565]	$C_{25}H_{27}NO$	358	*N*-desmethyl-tamoxifen	Y
8.164	82	89	[0.00000]	$C_{26}H_{29}NO$	372	Tamoxifen	Y
7.748	4	3	[-28.03130]	$C_{24}H_{25}NO$	344	*N,N*-didesmethyltamoxifen	N
6.165	3	4	[1.97926]	$C_{25}H_{27}NO_2$	374	(z)-*N*-desmethyl-4-hydroxytamoxifen (*Z*-endoxifen)	Y
7.147	2	2	[15.99491]	$C_{26}H_{29}NO_2$	388	4′-hydroxytamoxifen	Y
6.898	2	2	[1.97926]	$C_{25}H_{27}NO_2$	374	*N*-desmethyl-4′-hydroxytamoxifen	Y
8.953	1	9	[15.99491]	$C_{26}H_{29}NO_2$	388	Tamoxifen-*N*-oxide	N
6.505	1	2	[15.99491]	$C_{26}H_{29}NO_2$	388	4-hydroxytamoxifen	Y
3.851	1	1	[222.03757]	$C_{33}H_{39}NO_9$	594	*O*-glucuronide, *O*-methyl-tamoxifen†	N

†*Tentatively assigned*

Executive summary

- A workflow for efficient metabolite identification in first-in-man clinical studies is described.
- Metabolite prediction allows expected metabolite searching, including dealkylation and hydrolysis products.
- Integration with data processing allows structure-specific metabolite searches.
- Differential analysis complements metabolite prediction for unexpected metabolites.
- Literature and experimental data illustrate application for major circulating human metabolites.

References

Papers of special note have been highlighted as:
▪ of interest
▪▪ of considerable interest

1 Fura A. Role of pharmacologically active metabolites in drug discovery and development. *Drug Discov. Today* 11(3–4), 133–142 (2006).

▪▪ **Relevant examples of active metabolites.**

2 Smith DA, Schmid EF. Drug withdrawals and the lessons within. *Curr. Opin. Drug Discov. Dev.* 9(1), 38–60 (2006).

3 Uetrecht J. Idiosyncratic drug reactions: past, present, and future. *Chem. Res. Toxicol.* 21(1), 84–92 (2008).

4 Chen X, Zhong D, Xu H, Schug B, Blume H. Sensitive and specific liquid chromatographic–tandem mass spectrometric assay for dihydroergotamine and its major metabolite in human plasma. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 768(9), 267–275 (2002).

5 Chen X, Wang H, Du Y, Zhong D. Quantitation of the flavonoid wogonin and its major metabolite wogonin-7 beta-D-glucuronide in rat plasma by liquid chromatography–tandem mass spectrometry. *J. Chromatogr. B Analyt. Technol Biomed. Life Sci.* 775(2), 169–178 (2002).

6 Zhou S, Song Q, Tang Y, Naidong W. Critical review of development, validation, and transfer for high throughput bioanalytical LC–MS/MS methods. *Curr. Pharm. Anal.* 1(1), 3–14 (2005).

7 Li Y, Sun Y, Dong Y, Zhang Y, Zhong D. Identification and quantitative determination of benproperine metabolites in human plasma and urine by liquid chromatography–tandem mass spectrometry. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 830(1), 71–80 (2006).

8 Zhang H. Chapter 6: High-resolution LC–MS-based mass defect filter approach: basic concept and application in metabolite detection. In: *Mass Spectrometry in Drug Metabolism and Pharmacokinetics*. Ragu Ramanathan (Ed.). Wiley, NY, USA (2009).

9 Prasad B, Garg A, Takwani H, Singh S. Metabolite identification by liquid chromatography–mass spectrometry. *Trends Anal. Chem.* 30(2), 360–387 (2011).

10 Robison TW, Jacobs A. Metabolites in safety testing. *Bioanalysis* 1(7), 1193–1200 (2009).

11 Walker D, Brady J, Dalvie D *et al.* A holistic strategy for characterizing the safety of metabolites through drug discovery and development. *Chem. Res. Toxicol.* 22(10), 1653–1662 (2009).

12 Higgins J, Cartwright ME, Templeton AC. Progressing preclinical drug candidates: strategies on preclinical safety studies and the quest for adequate exposure. *Drug Discov. Today* 17(15–16), 828–836 (2012).

13 Obach RS, Nedderman AN, Smith DA. Radiolabelled mass-balance excretion and metabolism studies in laboratory animals: are they still necessary? *Xenobiotica* 42(1), 46–56 (2012).

14 Jacobs PL, Ridder L. Frequently reported mass differences and formula for list searching in drug metabolite identification. Proceedings of: *The 56th ASMS Conference on Mass Spectrometry and Allied Topics*, Denver, CO, USA, 1–5 June 2008.

15 Jacobs PL, Ridder L. Ruijken M, van Dongen WD. *i*Humite: a targeted metabolite profiling workflow. Proceedings of: *The 60th ASMS Conference on Mass Spectrometry and Allied*

Topics, Vancouver, BC, Canada, 20–24 May 2012.

16 Ridder L, Wagener M. SyGMa: combining expert knowledge and empirical scoring in the prediction of metabolites. *ChemMedChem* 3(5), 821–832 (2008).

▪ **Describes a relevant approach for rule-based metabolite prediction and ranking.**

17 Pelander A, Tyrkkö E, Ojanperä I. *In silico* methods for predicting metabolism and mass fragmentation applied to quetiapine in liquid chromatography/time-of-flight mass spectrometry urine drug screening. *Rapid Commun. Mass Spectrom.* 23(4), 506–514 (2009).

18 Trunzer M, Faller B, Zimmerlin A. Metabolic soft spot identification and compound optimization in early discovery phases using MetaSite and LC–MS/MS validation. *J. Med. Chem.* 52(2), 329–335 (2009).

19 Kirchmair J, Williamson MJ, Tyzack JD *et al*. Computational prediction of metabolism: sites, products, SAR, P450 enzyme dynamics, and mechanisms. *J. Chem. Inf. Model.* 52(3), 617–648 (2012).

▪▪ **Good overview of methods developed for the prediction of metabolite structures.**

20 Hop CE, Wang Z, Chen Q, Kwei G. Plasma-pooling methods to increase throughput for *in vivo* pharmacokinetic screening. *J. Pharma. Sci.* 87(7), 901–903 (1998).

21 Ma S, Chowdhury SK. Analytical strategies for assessment of human metabolites in preclinical safety testing. *Anal. Chem.* 83(13), 5028–5036 (2011).

▪▪ **Relevant overview of analytical strategies in Metabolites In Safety Testing.**

22 Teunissen SF, Rosing H, Seoane MD *et al.* Investigational study of tamoxifen Phase I metabolites using chromatographic and spectroscopic analytical techniques. *J. Pharma. Bio. Anal.* 55(3), 518–526 (2011).

▪ **Comprehensive overview of known Phase I metabolites of tamoxifen.**

23 Fromson JM, Pearson S, Bramah S. The metabolism of tamoxifen (I.C.I. 46,474).II. In female patients. *Xenobiotica* 3(11), 711–714 (1973).

24 Teunissen SF, Rosing H, Schinkel AH, Schellens JHM, Beijnen JH. Bioanalytical methods for determination of tamoxifen and its Phase I metabolites: a review. *Anal. Chim. Acta* 683(1), 21–37 (2010).

25 Jager NGL, Rosing H, Linn SC, Schellens JHM, Beijnen JH. Importance of highly selective LC–MS/MS analysis for the accurate quantification of tamoxifen and its metabolites: focus on endoxifen and 4-hydroxytamoxifen. *Breast Cancer Res. Treat.* 133(2), 793–798 (2012).

26 Chandrasekaran A, McKeand WE, Sullivan P, DeMaio W, Stoltz R, Scatina J. Metabolic disposition of [^{14}C]bazedoxifene in healthy postmenopausal women. *Drug Metab. Dispos.* 37(6), 1219–1225 (2009).

27 Christopher LJ, Hong H, Vakkalagadda BJ *et al.* Metabolism and disposition of [^{14}C] BMS-690514, an ErbB/vascular endothelial growth factor receptor inhibitor, after oral administration to humans. *Drug Metab. Dispos.* 38(11), 2049–2059 (2010).

28 Gong J, Gan J, Caceres-Cortes J *et al.* Metabolism and disposition of [^{14}C] brivanibalaninate after oral administration to rats, monkeys, and humans. *Drug Metab. Dispos.* 39(5), 891–903 (2011).

29 He H, Tran P, Yin H *et al.* Absorption, metabolism, and excretion of [^{14}C]vildagliptin, a novel dipeptidyl peptidase 4 inhibitor, in humans. *Drug Metab. Dispos.* 37(3), 536–544 (2009).

30 Renzulli C, Nash M, Wright M *et al.* Disposition and metabolism of [^{14}C]sb-649868, an orexin 1 and 2 receptor antagonist, in humans. *Drug Metab. Dispos.* 39(2), 215–227 (2011).

31 Sargentini-Maier ML, Espie P, Coquette A, Stockis A. Pharmacokinetics and metabolism of 14C-brivaracetam, a novel SV2A ligand, in healthy subjects. *Drug Metab. Dispos.* 36(1), 36–45 (2007).

32 Shilling AD, Nedza FM, Emm T *et al.* Metabolism, excretion, and pharmacokinetics of [^{14}C]INCB018424, a selective janus tyrosine kinase 1/2 inhibitor, in humans. *Drug Metab. Dispos.* 38(11), 2023–2031 (2010).

33 Speed B, Bu H-Z, Pool WF *et al.* Pharmacokinetics, distribution, and metabolism of [^{14}C]sunitinib in rats, monkeys, and humans. *Drug Metab. Dispos.* 40(3), 539–555 (2012).

34 Takusagawa S, van Lier JJ, Suzuki K *et al.* Absorption, metabolism and excretion of [(^{14})C]mirabegron (YM178), a potent and selective β(3)-adrenoceptor agonist, after oral administration to healthy male volunteers. *Drug Metab. Dispos.* 40(4), 815–824 (2012).

35 Vourvahis M, Gleave M, Nedderman ANR *et al.* Excretion and metabolism of lersivirine (5- benzene-1,3-dicarbonitrile), a next-generation non-nucleoside reverse transcriptase inhibitor, after administration of [^{14}C] lersivirine to healthy volunteers. *Drug Metab. Dispos.* 38(5), 789–800 (2010).

36 Wang L, Munsick C, Chen S *et al.* Metabolism and disposition of ^{14}C-labeled peliglitazar in humans. *Drug Metab. Dispos.* 39(2), 228–238 (2011).

37 Zhang D, He K, Raghavan N *et al.* Comparative metabolism of ^{14}C-Labeled apixaban in mice, rats, rabbits, dogs, and humans. *Drug Metab. Dispos.* 37(8), 1738–1748 (2009).

38 Zhou L, Kaul S, Liu-Kreyche P *et al.* Disposition of [1′-14C]Stavudine after oral administration to humans. *Drug Metab. Dispos.* 38(4), 655–666 (2010).

39 Zhu M, Whigan DB, Chang SY, Dockens RC. Disposition and metabolism of [^{14}C] brasofensine in rats, monkeys, and humans. *Drug Metab. Dispos.* 36(1), 24–35 (2007).

40 Ramanathan R, Zhong R, Blumenkrantz N, Chowdhury S, Alton K. Response normalized liquid chromatography nanospray ionization mass spectrometry. *J. Am. Soc. Mass Spectrom.* 18(10), 1891–1899 (2007).

41 Schadt S, Chen L-Z, Bischoff D. Evaluation of relative LC/MS response of metabolites to parent drug in LC/nanospray ionization mass spectrometry: potential implications in MIST assessment. *J. Mass Spectrom.* 46(12), 1281–1286 (2011).

▪ **Discusses a large set of nanospray examples of metabolite response.**

42 Cuyckens F, Pauwels N, Koppen V, Leclerq L. Use of relative ^{12}C/^{14}C isotope ratios to estimate metabolite concentrations in the absence of authentic standards. *Bioanalysis* 4(2), 143–156 (2012).

43 Yi P, Luffer-Atlas D. A radiocalibration method with pseudo internal standard to estimate circulating metabolite concentrations. *Bioanalysis* 2(7), 1195–1210 (2007).

44 Yu C, Chen C, Gorycki F, Neiss T. A rapid method for quantitatively estimating metabolites in human plasma in the absence of synthetic standards using a combination of liquid chromatography/mass spectrometry and radiometric detection. *Rapid Comm. Mass Spectrom.* 21(4), 497–502 (2007).

45 Gunduz M, Argikar UA, Baeschlin D, Ferreira S, Hosagrahara V, Harriman S. Identification of a novel *N*-carbamoyl glucuronide: *in vitro, in vivo*, and mechanistic studies. *Drug Metab. Dispos.* 38(3), 361–367 (2010).

46 Ridder L, van der Hooft JJJ, Verhoeven S, de Vos RCH, van Schaik R, Vervoort J. Substructure-based annotation of high-resolution multistage MS^n spectral trees. *Rapid Comm. Mass Spectrom.* 26(20), 2461–2471 (2012).

47 Ridder L, van der Hooft JJJ, Verhoeven S, de Vos RCH, Bino RJ, Vervoort J. Automatic

chemical structure annotation of an LC–MS based metabolic profile from green tea. Anal. Chem. 85(12), 6033–6040 (2013).

▪ **Relevant example of automatic metabolite structure annotation.**

48 Madlensky L, Natarajan L, Tchu S *et al.* Tamoxifen metabolite concentrations, CYP2D6 genotype, and breast cancer outcomes. *Clin. Pharmacol. Ther.* 89(5), 718–725 (2011).

■ Websites

101 ICH M3(R2) Guidance on Nonclinical Safety Studies for the Conduct of Human Clinical Trials and Marketing Authorization for Pharmaceuticals. ICH, Geneva, Switzerland (2010). www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm

102 US FDA, US Department of Health and Human Services, Center for Drug Evaluation and Research. Guidance for Industry. Safety Testing of Drug Metabolites. US FDA, Rockville, MD, USA (2008). www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm

103 The eMetabolomics research project. www.emetabolomics.org