A discovery biotransformation strategy: combining in silico tools with high-resolution mass spectrometry and software-assisted data analysis for high-throughput metabolism

Daniel J. Weston\textsuperscript{a}, Mehul Dave\textsuperscript{a}, Kevin Colizza\textsuperscript{b}, Steve Thomas\textsuperscript{a}, Laura Tomlinson\textsuperscript{c}, Richard Gregory\textsuperscript{c}, Claire Beaumont\textsuperscript{a}, Jill Pirhalla\textsuperscript{b} and Gordon J. Dear\textsuperscript{a}

\textsuperscript{a}GSK, DMPK, Disposition and Biotransformation, Stevenage, UK; \textsuperscript{b}GSK, DMPK, Disposition and Biotransformation, Collegeville, PA, USA; \textsuperscript{c}GSK, DMPK, Discovery DMPK, Stevenage, UK

ABSTRACT

1. Understanding compound metabolism in early drug discovery aids medicinal chemistry in designing molecules with improved safety and ADME properties. While advancements in metabolite prediction brings increased confidence, structural decisions require experimental data. In vitro metabolism studies using liquid chromatography and high-resolution mass spectrometry (LC–MS) are generally resource intensive and performed on very few compounds, limiting the chemical space that can be examined.

2. Here, we describe a novel metabolism strategy increasing compound throughput using residual in vitro clearance samples conducted at drug concentrations of 0.5 \textmu M. Analysis by robust ultra high-performance liquid chromatography separation and accurate-mass MS detection ensures major metabolites are identified from a single injection. In silico prediction (parent cLogD) tailors chromatographic conditions, with data-dependent tandem mass spectroscopy targeting predicted metabolites. Software-assisted data mining, structure elucidation and automatic reporting are used.

3. Confidence in the globally aligned workflow is demonstrated with 16 marketed drugs. The approach is now implemented routinely across our laboratories. To date, the success rate for identification of at least one major metabolite is 85%. The utility of these data has been demonstrated across multiple projects, allowing earlier medicinal chemistry decisions to increase efficiency and impact of the design–make–test cycle thus improving the translatability of early in vitro metabolism data.

Introduction

A comprehensive insight into the metabolic fate of a molecule is a key determinant of the absorption, distribution, metabolism and excretion (ADME) characteristics for new chemical entity (NCE) drugs. Metabolite identification has become firmly established within DMPK (drug metabolism and pharmacokinetics) departments across the pharmaceutical industry and is underpinned by analytical proficiency in chromatography, high-resolution mass spectrometry (Zhu et al. 2011), and nuclear magnetic resonance (NMR) spectroscopy (Dear et al. 2008). During the past two decades biotransformation science has received significant focus from industry, consortia, and regulatory agencies in relation to detecting, characterising and quantifying metabolites to define safety thresholds in clinical development relative to toxicology studies, namely Metabolites in Safety Testing (MIST). These concepts are reinforced in the most recent regulatory guidelines (ICH M3 2009; FDA 2020). MIST has been the subject of multiple seminal reviews (Smith et al. 2009; Smith and Obach 2010; Nedderman et al. 2011; Schadt et al. 2018) and will not be discussed further here.

Outside the attention to MIST, it is important to recognise that metabolite identification also has a critical role in the design–make–test cycle for small molecule medicinal chemistry (Nassar et al. 2004; Cerny et al. 2020), through amelioration of hepatic clearance for drug scaffolds. This can improve subsequent drug candidates in lead discovery and lead optimisation for efficacious human dose optimisation (Smith et al. 2019). Early metabolite identification also has an influence on metabolite risk assessment; for example, informing on bioactivation, mutagenicity, and toxicity potential can be mitigated through chemical redesign (Kalgutkar 2020). For these reasons, metabolite identification expertise has been applied increasingly in drug discovery, using the same analytical capability that supports MIST and clinical development. However, a fundamental difference between discovery and development is a need for analytical rigour that is...
proportionate with the discovery timescale and specifically aligned with the throughput of early discovery in vitro hepatic clearance assays (Luippold et al. 2010). If there is a disconnect between the generation of the metabolic stability data and associated metabolite identification, the latter becomes redundant as chemistry resource and design strategy move quickly in parallel with biological activity screening.

High-throughput metabolic stability assays are primarily conducted in liver microsomes or cultured hepatocytes. Both use automated methods and generic liquid chromatography-mass spectrometry (LC-MS) analyses on triple quadrupole instruments to accelerate data readout on the quantitative disappearance of parent drug in a timescale that can inform on the medicinal chemistry strategy (Shah et al. 2016). To advance early chemical design, it is vital that this quantitative data is supplemented with qualitative information on major metabolites. Providing information on the routes of metabolic clearance informs on the structure-activity relationship that can guide chemistry; such as sites of metabolism to modify drug clearance (Nassar et al. 2004; Cerny et al. 2020), for example. This has presented an analytical paradox because metabolite identification is typically an iterative and investigative process, often requiring numerous analytical experiments. These may include different MS scan functions (e.g. full-scan MS, product-ion scan, precursor-ion scan), ionisation polarities or chromatographic methods, along with metabolite purification, derivatisation, and NMR spectroscopy to enable unequivocal characterisation, which are not compatible with the speed of drug discovery.

Some groups have attempted to address this disparity through the incorporation of biotransformation prediction tools (Zelesky et al. 2013; Tzyack and Kirchmair 2019), automated data processing platforms (Mortishire-Smith et al. 2005; Bonn et al. 2010; Pähler and Brink 2013), the use of elegant data-dependent and data-independent scan approaches with hybrid Orbitrap instruments (Ma and Chowdhury 2013; Wilkinson et al. 2020; Ruan and Comstock 2021), and even integrated quantitative–qualitative (quan–qual) techniques using high-resolution LC–MS to achieve hepatic clearance screening (Backfisch et al. 2015; Paiva and Shou 2016; Paiva et al. 2017). Combinations of these approaches have been used with some success; however, key limitations often exist. For example, over-reliance on metabolite prediction software can introduce bias and cause unexpected metabolic routes to be overlooked. Quan–qual approaches can result in chromatographic resolution (data quality) compromises that hinder metabolite separation and identification (e.g. component co-elution, ion suppression). Moreover, technically achievable quan–qual methods that follow the classic approach of HRMS-only are often not practically feasible due to the high levels of automation, the capacity in modern metabolic stability screens, compromises in data quality, and the qualitative data load such an approach would generate with every sample to be analysed. Commentaries have addressed these issues in the wider context of progress for mass spectrometry within biotransformation, touching on scientific, economic, and cultural aspects, whilst suggesting next steps for the modern DMPK analytical laboratory (Cuykens 2018).

Recently, efforts were made by the authors (Weston et al. as published in Walles et al. 2022) to encourage a shift in how the biotransformation community viewed the quan–qual paradigm, as described at the 2nd European Biotransformation workshop in collaboration with the DMDG (24–25 November 2021) (see Figure 1). This addressed how these methods are executed in support of drug discovery and what quan–qual could look like if perspectives were broadened, given that this discussion showed that most industry approaches did not analyse all clearance samples (i.e. not aligned to Example One, Figure 1). With this concept in mind, a more pragmatic solution would be to adopt a revised approach to address the analytical paradox and facilitate fit-for-purpose metabolite identification with sufficient throughput and scale to support fast-paced medicinal chemistry programmes. In this article, the authors describe an alternative quan–qual discovery metabolism workflow using residual clearance screening samples (aligned to Example Three). This approach not only maximises relevance of metabolism data through sample choice driven by clearance data, but also saves time as samples do not need to be regenerated for metabolite identification. The novel workflow uses automation for sample generation, in silico prediction, generic methods for data acquisition, and software-assisted data analysis and reporting on exemplar compounds to maximise the information for the design–make–test cycle without becoming the bottleneck of the process.

Figure 1. Schematic representing real-world alternatives to the traditional HRMS-based quan–qual approach.
This novel strategy combines elements from an in silico predict-first strategy, fast gradient chromatography method selection, software-assisted data analysis and reporting with the accuracy and precision of high-resolution liquid chromatography–tandem mass spectrometry (LC–MS/MS), without the need for additional sample incubations and minimised sample re-injection. Residual bar-coded in vitro hepatic clearance samples (from triple quadrupole-based ‘quan’ analyses) are analysed using high-resolution LC–MS/MS (to deliver structural ‘qual’ information), following a request through a LIMS scheduling system. A generic LC/MS method is selected based on empirical (calculated) parent drug LogD (cLogD). Metabolite predictions are used to inform on tandem MS inclusion lists to ensure the provision of targeted m/z values for MS/MS of expected metabolites. This provides cleaner, more relevant MS/MS data for automated interrogation of predicted metabolites using ACD MetaSense (Watanabe et al. 2017), for example. Together, these upfront in silico predictions help inform key decisions around generic method choice and overall approach, prior to commencing any experimental work, increasing confidence in getting the right data first time, thereby saving effort and time. Supplementation data interrogation by biotransformation scientists provide continued oversight to ensure any metabolite prediction-led bias is minimised and to help build confidence in the use of software-assisted data analysis. The detection of unexpected metabolites (non-targeted m/z) is achieved through manual interrogation of full-scan FTMS data, along with fragment-ion data from data-dependent MS/MS acquired automatically on the top-three most intense ions (not on an inclusion list) from each Fourier transform mass spectrometry (FTMS) spectrum. How the software is used to detect and integrate (relative MS response), interpret (MS/MS fragmentation), and report (metabolic scheme, spectra, chromatograms) the major metabolites is also described. Post-reporting, metabolite schemes are uploaded and stored within a central database to enable the search and recall of metabolites based on sub-structure or other criteria. As molecules continue into development, metabolic schemes may be updated when new (in vitro, in vivo, metabolite synthesis) data are generated, thereby providing a platform that acts as a repository for end-to-end results. Data can also be visualised using Spotfire by DMPK scientists and medicinal chemists throughout the global organisation, providing added context to clearance data, and in a format compatible with other data.

Herein, the workflow is described in detail, including a manually evaluated test data set used to build confidence in the analytical workflow prior to moving to software-assisted data interrogation. Examples of real-world project data are also shared to highlight the utility of this approach and its impact on chemical design. Importantly, the new quan–qual approach is built for throughput and scale, which provides the opportunity to explore more chemical space for greater impact on chemistry and drug clearance, whilst aligning more closely with the required cycle times to meet decision milestones and maximise the translatable data of these early in vitro metabolism data.

Materials and methods

In silico tools

Metabolism. Metabolite structure predictions were carried out using in silico software packages, namely MetaSite v6.0.5 (Molecular Discovery, Hertfordshire, UK) and Meteor Nexus Suite v3.1.0 (Lhasa, Leeds, UK), where predictions were then combined into a single SDF file. These data were used to create MS inclusion lists prior to LC–MS/MS analysis and as searchable input into the data-mining software.

LC gradient. The calculated LogD (cLogD) of each parent compound was used to select a preferred chromatographic method from a set of three generic polarity-based alternatives (polar, neutral, and non-polar). cLogD was predicted using Helium v3.0.3 (Ceiba Solutions) by inputting the compound structure as a SMILES string. Based on the nearest calculated cLogD (pH 2) to the selected mobile phase (0.1% (v/v) formic acid, approximately pH 2.8), gradient choice criteria were polar (cLogD < 0), neutral (0 < cLogD < 2), or non-polar (cLogD > 2). Analytical method details for all three ultra-high-performance liquid chromatographic (UHPLC) gradients are described below.

Chemicals

Drug compounds (7-ethoxycoumarin, dextromethorphan, diazepam, difusnil, domperidone, flumazenil, imipramine, ketoprofen, nifedipine, phenacetin, phenytoin, tenoxicam, tolbutamide, triprolidine, verapamil, warfarin) were provided either by GSK internal compound management services or purchased externally (Sigma Aldrich, Poole, UK). Structures of parent drug compounds are shown in the Supplemental Figure S1. Novel GSK compounds used in real-world application of this workflow were provided by GSK internal compound management. Mobile phases were made using 1 L of UHPLC grade water (Riedel-de Haen) or 1 L of UHPLC grade acetonitrile (Riedel-de Haen) in sealed glass bottles (both from Honeywell, Bracknell, UK) to which was added ultra-pure formic acid (1 mL) from break-neck vials (Waters, Manchester, UK) before manual mixing to give mobile phase A (0.1% (v/v) formic acid (aqueous)) and mobile phase B (0.1% (v/v) formic acid in acetonitrile), respectively. Other sample preparation solvents and chromatographic wash solvents consisted of water and acetonitrile (both as above) or methanol (Ultra-Pure Grade, CromSil, UK). A multi-analyte LCMS QC stock solution (Waters, Manchester, UK) was used for system suitability testing. The mass spectrometers were calibrated as required using Pierce LTQ-Velos Calibration Mix in either positive-ion or negative-ion mode (ThermoFisher Scientific, San Jose, CA).

Hepatocyte incubations for metabolic clearance assay

Cryopreserved human and rat (Han Wistar) hepatocytes (BioIVT, Brussels, Belgium) were defrosted according to manufacturer instructions and diluted to one million viable
cells per mL in Williams Media E (WME) (Lonza, Belgium), supplemented with 2 mM Glutamax (Gibco, UK) and 25 mM HEPES (Cytiva, USA). Incubations were conducted in a 96-well format using a Bravo automated liquid-handling workstation with Vworks software (both by Agilent, Cheadle, UK), performed in triplicate for each compound and species with a total incubation volume of 200 μL. Final incubation conditions consisted of 0.5 million viable cells per mL cell density, 0.5 μM compound concentration, incubated at 37°C with 5% CO₂ (v/v) for 4 h with orbital shaking (300 rpm). After 4 h, incubation samples were pipette-mixed and transferred (180 μL) into a barcoded sample plate containing 80:20 acetonitrile:ethanol (v/v) (360 μL), giving a maximum working concentration for parent drug of 0.17 μM. The sample plate was shaken vigorously (20 min) before centrifugation (3500 × g, 20 min). Samples were analysed in a metabolic clearance assay, after which residual samples were retained and frozen (−80°C) for subsequent metabolite identification.

**Preparation of hepatocyte samples for LC–MS/MS analysis**

Residual hepatocyte clearance samples (t = 4 h) were delivered from freezer storage following an electronic laboratory information management system (LIMS) request, then allowed to thaw (ambient temperature). Plate covers were removed and samples were mixed via manual pipette aspiration, before being transferred to Eppendorf tubes for centrifugation (10,000 × g, 5 min). Supernatants were transferred to Total Max Recovery HPLC vials (Waters, Manchester, UK) and capped. As the in vitro clearance assay did not automatically generate a representative t = 0 h sample, a surrogate was prepared using bulk rat or human control hepatocyte matrix (250 μL, already quenched as per t = 4 h samples) into which was spiked parent drug stock solution (1 μM, 50 μL) to give the same concentration as the t = 4 h samples (0.17 μM) prior to any metabolic turnover. Blank quenched hepatocyte matrix was taken (250 μL, vials as above) into which was spiked supplemented WME (50 μL) as a representative biological blank (no drug) control for each species. All samples were prepared within 30 min of thawing to minimise issues around sample stability or solvent losses due to evaporation. Aliquots of all 16 test set samples were taken in triplicate; one set was sent to each laboratory (USA and UK) with the third retained as a backup set. Samples for real project compounds were taken as a single aliquot directly from the relevant clearance plate.

**LC–MS/MS metabolite identification methods**

This workflow was developed for global drug discovery support, and therefore it was important to align the analytical instrumentation as closely as possible to allow flexible resourcing and global comparison of metabolite profiles. These methods were built around Q-Exactive instrumentation across both geographic sites, utilising Orbitrap technology to acquire accurate mass FTMS data and high-quality data-dependent tandem mass spectrometric information. A basic outline of each system is given below.

**LC–MS/MS instrumentation**

**UK site.** Analyses were performed using a Q-Exactive Plus mass spectrometer (ThermoFisher Scientific, Bremen, Germany) with integrated divert valve, hyphenated with a Vanquish Flex UHPLC liquid chromatograph system (ThermoFisher Scientific, Bremen, Germany) consisting of a Vanquish Binary Pump F (VF-P10-A-01); Split Sampler FT autosampler (VF-A10-A-02, held at 8°C (isothermal)) with a flow-through needle; dual-zone column oven (VF-C10-A-03) and a photo-diode-array UV detector (VF-D10-A). Instrument control was carried out using Xcalibur software v4.2.47 and SIL LC driver control software (both by ThermoFisher Scientific, Bremen, Germany); manual data analysis and writing of exclusion lists were performed using QualBrowser (Xcalibur, v4.2.47) and Compound Discoverer (v3.2) software, respectively (both by ThermoFisher Scientific, Bremen, Germany).

**US site.** Analyses were performed using a Q-Exactive mass spectrometer (ThermoFisher Scientific, San Jose, CA, USA) with integrated divert valve, hyphenated with an Agilent 1290 Series UHPLC liquid chromatograph system (Agilent Technologies, Palo Alto, CA) consisting of a binary pump (Agilent 1290 G4220A); autosampler (Agilent 1290 G4226A) with a flow-through needle; column oven (Agilent 1290 G1316C) and a photo-diode-array UV detector (Agilent 1260 G4212B). Instrument control and manual data interrogation were carried out using Xcalibur software v4.2.47 with relevant third-party LC drivers in SIL control (both by ThermoFisher Scientific, San Jose, CA, USA); manual data analysis and writing of exclusion lists were performed using QualBrowser (Xcalibur, v4.2.47) and Compound Discoverer (v3.2) software, respectively (both by ThermoFisher Scientific, San Jose, CA).

**LC–MS/MS methodology**

Samples were analysed at both sites in an identical manner, using one of three generic reverse-phase UHPLC gradients, prior to component elution into the mass spectrometer. Solvent was delivered (0.4 mL/min) into which samples were injected (1–5 μL) and separated using an Acquity Premier BEH C18 UPLC column (2.1 × 100 mm, 1.7 μm) with VanGuard FIT integrated guard column (Waters, Manchester, UK) held at 50°C (isothermal). UV diode-array data were acquired as default (190–400 nm, 0.2 s/scan). The three gradients were optimised; final conditions were: 0.0 min (start), 2.0% B; 0.5 min, 2.0% B; 8.0 min, 50% B (polar) or 70% B (neutral) or 90% B (non-polar); 9.0 min, 95% B; 10.0 min, 95% B; 10.2 min, 2.0% B; 12.0 min (end), 2.0% B; all gradient changes were linear. The autosampler needle was washed for 2 s (both pre- and post-injection) using 9:10 (v/v) water:methanol. Analytical blank injections (2:98 (v/v) A:B, 5 μL) were performed prior to injecting study samples. In total, there were six generic methods which represented either positive or negative-ion MS polarities for the three gradient choices. Based on the chosen gradient via LogD (pH 2) for an individual compound, two LC–MS methods (one for each MS polarity) were prepared for each compound using the generic
LC method, generic MS tune parameters, generic exclusion list and an imported compound-specific data-dependent analysis (DDA) MS/MS inclusion list.

The Q-Exactive mass spectrometers were operated in either positive-ion or negative-ion heated electrospray ionisation (hESI) mode, acquiring FTMS (full-scan) data along with tandem mass spectrometric (MS/MS) data using DDA rules and associated inclusion (to ensure MS/MS triggering for metabolites) and exclusion lists (to reduce background triggering). Headline method details are shown herein, however finer method details (e.g. DDA triggering thresholds, AGC timings) are shown in Supplemental Material, favouring an overall basic description. Source conditions were the same in both polarities unless stated otherwise and were as follows: hESI spray voltage (+ve ion 3.5 kV, –ve ion –3.0 kV); maximum spray current (100 μA); probe heater (400 °C); heated capillary (320 °C); sheath gas (flow of 60 (arbitrary units)); auxiliary gas (flow of 20 (arbitrary units)); S-Lens RF (40 V); high-mass range mode (OFF); data mode (centroid). All FTMS data were acquired (m/z 100–1000; 35 000 resolution; acquire between runtime of 1.0 and 10.0 min; MS divert valve on from 0.0–1.0 min to 10.0–12.0 min) together with data-dependent MS/MS routines running in parallel in the same method. DDA parameters (see Supplemental Material, Section B) allowed the top-three ions from each FTMS scan (prioritised from an inclusion list) to be isolated and fragmented to give data-rich MS/MS spectra on-the-fly. Study samples were injected in positive-ion mode followed by negative-ion mode immediately afterwards, both batches bracketed by replicate system suitability tests (SSTs). Consistent background ions were used as internal lock masses in each polarity for reliable one-point mass correction in real-time to ensure mass accuracy.

**LC–MS SST methodology**

System suitability for the Q-Exactive LC-MS was established using the same mobile phases, UPLC column, MS tune file as used in the metabolite identification methodology above. Notable changes for MS method were as follows: resolution (17 500); FTMS mode only (scan range as above), no DDA-MS2 routine; polarity-switching (ON). Replicate injections of a multi-analyte LCMS QC stock solution (Waters, Manchester, UK) diluted 1:50 (v/v) in starting mobile phase composition (95:5 A:B, v/v) were injected (1 µL) and separated using a short five-minute gradient (0.5 mL/min) prior to full scan FTMS detection in both positive and negative-ion modes using polarity switching. Additional details are given in the Supplemental Material (Section C).

**Software-assisted data processing**

Software-assisted data interrogation of mass spectrometric data from in vitro incubations was carried out using MetaSense software (ACD Labs Suite v2021.2.2, Advanced Chemistry Development Inc., Toronto, Canada), where metabolite detection and componentisation across samples and species was performed. Blank (no drug) control incubation data were subtracted automatically from individual (r = 4 h for hepatocytes) data sets, where a multiplier setting was used for removal of endogenous peaks. Resultant data were interrogated for the presence of predicted metabolites using the IntelliTarget algorithm combined with comprehensive prediction data for potential metabolite structures generated from Meteor and Metasite (as described previously).

Full scan FTMS and DDA MS/MS data were automatically extracted for detected components. Metabolites were aligned and rationalised across species using m/z and retention time. Manual data interrogation for metabolite detection, fragment-ion interpretation, fragment-ion rendering, and metabolite naming were carried out at critical stages of review by a biotransformation scientist to ensure accuracy. A metabolic scheme was generated together with a histogram showing relative metabolite abundances. Agreed metabolite structures (either as Markush or preferred structures from in silico predictions) and metabolic schemes were written to a company-wide database using Spectrus, with reporting being accomplished via a user-customizable ChemSketch template (both by ACD Labs, as above).

**Results and discussion**

**Outline**

A workflow was developed which enabled identification of the major metabolites in residual rat and human in vitro clearance samples in a time-efficient manner using in silico predictions, together with streamlined analytical methodologies and data-reporting. The forthcoming section discusses key results from the optimised LC-MS/MS methodology, software-assisted data interrogation and reporting using a set of 16 commercially available compounds incubated in rat and human hepatocytes. These compounds were structurally diverse (Supplemental Figure S1) and had a range of intrinsic clearance (CLint) values in hepatocytes. The impact of this workflow for real-world drug project support is discussed, along with aspirations for future advances in sample throughput to further reduce expert input and turnaround times whilst running higher sample numbers and maintaining fit-for-purpose data quality at scale.

**Establishing system performance using an LC–MS SST**

To maintain data quality and optimise success from a single injection, the multi-analyte SST was injected before and after the analysis of study samples and control incubations. When performance criteria were satisfied, study samples were submitted for analysis using the metabolism workflow method. Additional details and acceptance criteria are given in the Supplemental Table S1 and Figure S2, showing typical output for replicate SST injections.
**LC–MS/MS data quality and fit-for-purpose workflow criteria**

The overall aspiration for the metabolism workflow was to demonstrate the capability to produce high-quality LC-MS/MS data for a minimum of one major metabolite for ≥ 80% of compounds analysed. The success of any software-based approach for data mining is highly contingent on the quality and suitability of the raw data. Prior to using any form of software-assisted data interrogation, it was important in the first instance to manually assess the quality and relevance of the LC–MS/MS data obtained using the described analytical approach. To minimise the need for additional reinjection and maintain an adequate overall duty-cycle of the mass spectrometer, several key considerations had to be addressed for the LC–MS/MS workflow to ensure detection of relevant metabolites:

- **Sensitivity:** To maintain biological relevance, intrinsic clearance determinations are performed using low sample concentration (0.5 μM), assuming this is below the \( K_m \) for any enzyme in the system and non-saturated first-order enzyme kinetics apply (Srinivasan 2021). Adequate analytical sensitivity had to be demonstrated when analysing these relatively low working concentrations (0.17 μM, post-quench) from clearance assay samples.

- **LC method choice:** The chromatographic performance for parent compound and detected metabolites needed to be fit-for-purpose with adequate retention and separation of metabolites without prohibitive fronting, splitting or early elution due to poor method choice (McCalley 2010). Sufficient organic content was also necessary for adequate ionisation (Nguyen and Fenn 2007). Marked changes in LogP can be observed for phase I and phase II metabolism (Parkinson et al. 2019) but estimation of LogP does not consider pH. To accommodate the chemical diversity of the GSK portfolio of chemical scaffolds, cLogD of the parent drug at pH 2 (closest to that of the LC mobile phase) proved a useful first step to help inform on the most appropriate LC conditions. Parent cLogD values (pH 2) were binned into three categories: polar, neutral, and non-polar.

- **Automated MS/MS:** To aid the capture of high-quality relevant MS/MS fragment-ion data for structural elucidation, data-dependent inclusion lists were produced based on parent drug biotransformation predictions from both MetaSite and Meteor software. The combined information was exported into a .SDF file format for transfer into Microsoft Excel. Following the manual removal of duplicate \( m/z \) values, positive and negative-ion lists were then imported into the Xcalibur LC/MS methods to act as an inclusion list for DDA routines. In order to ensure triggering of relevant MS/MS data on non-predicted metabolites in the same injection, exclusion lists were generated to exclude background ions (Guo et al. 2006), which can often monopolise data-dependent acquisition and decrease data relevance. Compound Discoverer was used to generate generic, pre-defined exclusion lists from analysis of blank (no drug) rat and human samples in both polarities for each gradient, which were embedded into all generic LC–MS methods as default.

- **MS duty cycle:** Many Orbitrap-based mass spectrometers can perform polarity switching, but this takes a heavy toll on instrument duty cycle (>1 s). This is incompatible with maintaining six-to-eight data points per UHPLC peak which is desirable for robust peak detection, particularly where DDA analyses are required in the same sample injection. Using a single injection per polarity to trigger on only relevant (drug-related) ions for the acquisition of DDA–MS/MS data was critical. Therefore, separate injections were required in each polarity for each batch of samples to assure high-quality data for both acidic and basic metabolites.

The performance of the workflow was assessed by analysing 16 commercial drug compounds incubated in rat and human hepatocytes. Data were initially evaluated manually for detection of metabolites (observed vs. predicted) in FTMS mode. MS and MS/MS sensitivities for both parent and metabolites were assessed and the quality of the resulting MS/MS data was examined where DDA routines were triggered.

An example of typical method performance is shown for warfarin (\( t = 4 \) h), a low clearance compound in rat hepatocytes (Figure 2). Detection of the parent drug is demonstrated with good sensitivity in full-scan positive-ion FTMS mode, as shown in the extracted ion chromatogram (XIC) in Figure 2(a), along with detection of three chromatographically resolved oxidised metabolites (Figure 2(b)). These metabolites are all present at approximately <3% parent response, assuming equimolar MS response, where the least intense oxidised metabolite (RT 5.55 min) equates to approximately 4.8 pg on column (Figure 2(b)).

Good chromatographic performance was observed for both parent and metabolites when using the prediction-led choice of ‘non-polar’ LC gradient based on parent cLogD, and predicted warfarin metabolites were detected. Here, the mobile phases and column chosen for first intent performed well, giving robust chromatographic performance where, despite acetonitrile being known to suppress ionisation or protomer formation on oxygen atoms in positive-ion mode (Colizza et al. 2016; Zheng and Attygalle 2021), it was a favoured solvent choice over methanol or other alcohols due to unwanted pressure effects (Aburjai et al. 2011). Furthermore, DDA-MS/MS data were triggered for parent compound (not shown) and all three low-level metabolites (see total ion chromatogram (TIC), Figure 2(c)), providing high quality fragment-ion spectra on-the-fly (see Figure 2(d) and (e)) to enable confident structural localisation of the net biotransformation. Furthermore, using Compound Discoverer to produce retention-time-aligned exclusion lists greatly reduced redundant DDA triggering, allowing optimised use of MS duty-cycle.

Respective data for the same 16 replicate rat and human hepatocyte sample sets were interrogated manually and compared for analytical method performance across both sites (Table 1). For the 32 test sets (16 compounds in rat and...
human hepatocytes) between sites, we identified at least one major metabolite in all instances for a success rate of 100% (as shown in Table 1). Good success rates were observed for the automatic provision of usable fragment-ion data via effective and efficient triggering of DDA–MS/MS for detected metabolites, even when metabolites were detected at low (approximately single-figure nM) levels (e.g. Figure 2(c)). The number of total metabolites detected across the rat and human sample sets at US and UK sites (72 and 75, respectively) and subsequent DDA triggering on these metabolites (56 and 59, respectively) showed success rates of 78% (USA) and 79% (UK), greatly reducing the need for additional sample reinjection to acquire missing MS/MS data.

Here, DDA data were triggered for all major metabolites; only the most trace-level metabolites (i.e. those with an ion-count lower than the DDA threshold) did not trigger. In addition, the close agreement of these independent performance data highlights careful and continued alignment of the workflow across our global organisation. These data demonstrate the ability of upfront in silico prediction of cLogD to facilitate the correct choice of generic chromatographic method for each compound, along with delivering consistent, robust chromatographic selectivity and relative retention time for consistent global project support. The majority of test compounds performed better in positive-ion mode; however, one compound (diflunisal) was markedly improved in

Table 1. Data summary for the manual interrogation of raw data obtained from 16 commercial compounds incubated in rat and human hepatocytes (t = 4 h) analysed under the conditions of the metabolism workflow.

| Compound          | CLINT (human; rat) | cLogD classification (pH 2) | Optimum sensitivity | RRT alignment | Metabolites detected (US:UK) | DDA-MS2 triggered (US:UK) |
|-------------------|---------------------|-----------------------------|---------------------|---------------|------------------------------|---------------------------|
| 7-Ethoxycoumarin  | (H; H)              | Neutral                     | Positive-ion        | YES           | (3; 3)                       | (0; 1)                    |
| Dextromethorphan  | (M; H)              | Polar                       | Positive-ion        | YES           | (7; 6)                       | (6; 6)                    |
| Diazepam          | (L; H)              | Non-polar                   | Positive-ion        | YES           | (3; 3)                       | (2; 3)                    |
| Diflunisal        | (M; L)              | Non-polar                   | Negative-ion        | YES           | (6; 6)                       | (2; 4)                    |
| Domperidone       | (H; M)              | Polar                       | Positive-ion        | YES           | (4; 5)                       | (4; 4)                    |
| Flumazenil        | (M; M)              | Neutral                     | Positive-ion        | YES           | (4; 4)                       | (4; 4)                    |
| Imipramine        | (M; H)              | Non-polar                   | Positive-ion        | YES           | (5; 5)                       | (5; 5)                    |
| Ketoprofen        | (H; L)              | Non-polar                   | Positive-ion        | YES           | (2; 2)                       | (2; 1)                    |
| Nifedipine        | (H; H)              | Neutral                     | Positive-ion        | YES           | (5; 4)                       | (4; 3)                    |
| Phenacetin        | (M; M)              | Neutral                     | Positive-ion        | YES           | (3; 4)                       | (3; 4)                    |
| Phenytoin         | (L; L)              | Non-polar                   | Positive-ion        | YES           | (2; 2)                       | (0; 2)                    |
| Tenoxicam         | (L; L)              | Polar                       | Positive-ion        | YES           | (2; 2)                       | (2; 2)                    |
| Tolbutamide       | (L; L)              | Non-polar                   | Positive-ion        | YES           | (3; 3)                       | (3; 3)                    |
| Triprolidine      | (M; H)              | Polar                       | Positive-ion        | YES           | (9; 11)                      | (8; 7)                    |
| Verapamil         | (H; H)              | Non-polar                   | Positive-ion        | YES           | (8; 8)                       | (5; 5)                    |
| Warfarin          | (L; L)              | Non-polar                   | Positive-ion        | YES           | (6; 6)                       | (6; 5)                    |

Turnover classification relating to measured disappearance of parent compound at t = 4 h vs. T = 0 h in hepatocytes and scaled to % liver blood-flow (%LBF) as reported from intrinsic clearance assay, where L = low (< 30%); M = moderate (30–70%); and H = high (> 70%).

Figure 2. Representative data for warfarin in rat hepatocytes (t = 4 h) showing (a) warfarin parent drug XIC (m/z 309.1107), (b) oxidised warfarin metabolites XIC (m/z 325.1071), (c) DDA-MS/MS TIC for oxidised warfarin metabolites (MS/MS of m/z 325.1071), and resultant MS/MS spectra from DDA–MS/MS triggering of oxidised warfarin metabolites at (d) RT 5.91 min and (e) 5.55 min. All XICs constructed using 10 ppm m/z window.
negative-ion mode, which shows that being able to provide data from both polarities is critical in ensuring best performance first time (i.e. improved detection of structurally diverse metabolites, improved project relevance) and providing confidence in the workflow. Also, where metabolism alters the ionisation efficiency of a metabolite (e.g. dealkylation, addition of a sulphate conjugate), the resulting metabolites may ionise relatively poorly in positive-ion mode, requiring negative-ion data to provide a comprehensive, more relevant composite picture of metabolism.

Together, these results confirmed successful proof-of-concept and fit-for-purpose performance for sensitivity, chromatographic resolution, DDA triggering, and data quality from a globally aligned analytical methodology, allowing progression to using the rat and human hepatocyte LC–MS/MS raw data in the evaluation of software-assisted data interrogation and reporting.

Software-assisted data analysis and reporting for hepatocyte incubation test-set

ACD MetaSense was used for the software-assisted interrogation of the data-dependent LC–MS/MS data for the 16 compounds incubated with rat and human hepatocytes, followed by generation of a summary report for each compound. Software evaluation criteria were as follows:

- **Automated detection of metabolites**: Comparing results to those from the manual data interrogation, the software was expected to achieve a success rate of 80% or greater.
- **Automated identification of detected metabolites**: Correct alignment of MS/MS fragment ions for detected metabolites was expected to aid in localisation of biotransformation, which should not be inconsistent with manual data interpretation.
- **Automated reporting of metabolites**: The ability to produce a single report summarising the metabolites detected across each matrix, together with the underlying interpreted spectral data was desired.
- **Metabolite databasing**: Alignment with existing GSK working practices for databasing which would continue to allow both in silico and experimental biotransformation data (e.g. chromatographic data, MS and NMR spectra, together with interpretation) to be accessed by a single vendor-neutral platform.

Success rates were calculated for the automated detection of metabolites in rat and human hepatocyte incubations for the 16 test compounds using MetaSense software; these rates were then compared to those for detection and characterisation using manual data interrogation (UK data). All major metabolites were detected across all samples when using the software. When considering the detection of metabolites regardless of intensity, 57 out of 75 metabolites (76%) were detected as being consistent with data outcomes from the manual approach. Metabolites that were detected manually but not detected using the software were either below the threshold for peak integration (six metabolites), integrated poorly (eight metabolites), co-eluting with isobaric endogenous species (one metabolite), or were not on the predicted metabolite list (three metabolites). Conversely, three minor predicted metabolites which had been missed during manual data mining due to intensity or component coelution were detected using the software, highlighting the advantages of a non-biased semi-automated approach. A single false-positive metabolite was also reported, identified upon manual review as an in-instrument fragment ion of parent compound (tolbutamide), showing the continued need for manual oversight in contextualising and rationalising metabolite data. Importantly, all major metabolites were detected in all 32 test samples and, whilst some of the efficiencies of software-assisted data interrogation are inherent in this approach, these overall results highlight the need for improvement around the core performance of this software (see ‘Future directions’).

Once potential metabolites were detected using the software, a more manual approach was needed to localise each biotransformation. MS/MS fragment ion assignment was performed manually within the processing software to improve accuracy of overall localisation and fragment rendering. Minor or unrelated components were removed, prior to manual addition of relevant metadata, leaving rationalised and agreed summary data for drug-related material to be reported using a user-customizable template. An example report for warfarin in rat hepatocytes is shown in Figure 3. Report output included a metabolic scheme of detected metabolites and a comparison table of metabolites detected in each species (Figure 3(a)), combined XICs for individual species depicting detected drug-related material (Figure 3(b)), and an example summary of the proposed structure (Figure 3(c)) with MS/MS spectrum (Figure 3(d)) and fragment interpretation (Figure 3(e)) for the oxidised warfarin metabolite M03 (RT 5.98 min) is shown here. All other metabolites are reported in the same manner, along with parent drug (data not shown).

**Application of the metabolism workflow to the GSK portfolio**

This metabolite identification workflow has now been established globally, allowing the higher throughput of samples from multiple chemical classes to be investigated. Here, we discuss the overall impact and value added to drug projects from using this new workflow, along with improvements in turnaround time, overall efficiency, method suitability first-time-out, and data quality. Whilst our routine use of software-assisted data interrogation currently requires a measured degree of manual input and oversight by biotransformation scientists to ensure real-world success, we touch upon filling current gaps, in addition to any aspirations for the future, to scale this approach. Our ultimate aim here being to deliver fit-for-purpose data for more compounds, faster; in turn, further reducing the level of manual intervention.

A novel advantage to this approach, inherent in the intended design of this workflow and the resultant analytical
performance, is the routine use of residual clearance screening samples for metabolite identification. Not only does this allow direct correlation with measured clearance values and enable optimal compound choice, but data are derived from biologically relevant concentrations (0.5 μM) where non-saturated first-order kinetics exist, rather than those typical for in vitro metabolism studies (e.g. 10 μM). Being able to provide metabolism data on lower concentration in vitro samples means generated metabolite profiles are more relevant; thus, they are less readily influenced by enzyme saturation, which leads to poor correlation with in vivo metabolism data, as is often observed when using markedly higher in vitro incubation concentrations. In addition, the use of residual clearance samples negates the need to conduct incubations specifically for metabolite identification, thereby leading to resourcing efficiency gains together with a reduction in overall delivery times for metabolism data. Together, these factors allow for increased translatability of in vitro metabolism data when using this workflow, meaning that project decisions can be made earlier and with added confidence.

A key measure of workflow performance in support of early discovery and chemical design is not only the impact of these data and their quality, but importantly the ability to

Figure 3. Output for software-assisted data interrogation and reporting of warfarin in rat hepatocytes (0.5 μM, t = 4 h), showing: (a) metabolic scheme and summary table; (b) combined XICs for all detected metabolites; (c) proposed structure for metabolite M03 (m/z 325.1069, RT 5.98 min), along with (d) respective MS/MS data and (e) fragment-ion interpretation.
provide the data in a meaningful timescale. Since the introduction of this workflow, average turnaround times have decreased markedly, reducing by almost two-thirds. This meaningful reduction in data delivery time forms part of a continuing trend in achieving faster turnaround times as this workflow continues to embed and be refined further; ultimately allowing the provision of critical metabolism data to align with the cycle time of the design-make-test process on more exemplar compounds. In the continued interest of both increased data relevance and shorter turnaround times when analysing samples from drug discovery projects, it is essential to define how to report metabolites in the most pragmatic way. Here, the provision of the most relevant data in the shortest time without limiting reporting to a fixed number of detected metabolites is desired, therefore reporting is limited to the most abundant metabolite (M1) detected, defined by its UHPLC peak area (10 ppm FTMS XIC window), along with any other metabolite(s) with a UHPLC peak area...
peak area greater than or equal to half of the peak area for metabolite M1. In order to best utilise *in silico* predictions, where metabolite prediction favours a single structure and MS/MS fragmentation data is supportive but not definitive, any predicted structures detected are reported as ‘preferred’. It should be noted that, whilst the ability to detect and characterise metabolites from low-clearance compounds at low incubation concentrations reflects the excellent performance of the analytical methodology, the workflow is almost exclusively applied to addressing and understanding issues around high intrinsic clearance for maximised relevance and utility to chemical design.

In terms of the relevance of generic LC-MS/MS method choice guided by upfront *in silico* prediction of cLogD of the parent drug, the success rate in this method giving fit-for-purpose method choice is >90% for the analysis of
real-world drug project samples to date. This shows the utility of our approach, bringing greater confidence and right-first-time consistency; in addition, method development is virtually eliminated, in turn ensuring the chromatographic method is far more suited to the chemical space being explored. Instances where the initial LC method choice was unsuitable (<10% of the time) was due to either poor prediction of parent drug cLogD (i.e. parent performed unexpectedly poorly when using chosen method) or the chromatographic performance of markedly polar metabolites was unacceptable (e.g. excessive peak fronting, splitting or analyte breakthrough). Together, the wide applicability of these generic UHPLC methods, along with flatter generic tuning of the Q-Exactive (wide sweet-spot for broad sensitivity across a diverse chemical space) and the robust overall performance of the workflow, mean that prohibitive issues or the need for additional method development have been a somewhat rare occurrence.

To date, the new workflow described above has been implemented for the analysis of over fifty sample sets, covering several therapeutic areas and multiple chemical entities, using residual samples from rat and human hepatocyte (t = 4 h) or liver microsome incubations (t = 45 min). The performance characteristics discussed above, along with extensive development and robust testing across global sites, established a high level of confidence in the performance of the workflow for real-world project samples. Consequently, we can report that at least one major metabolite has been detected and identified for 85% of all project samples submitted to date using the metabolism workflow. Derived data, metabolite schemes, spectral assignments, and metadata are uploaded to an ACD database, where all associated metabolism data are stored and made available for searching. Curated biotransformation summary data are also made visible from this database using Spotfire tools, to cross-link with other biological data and outcomes, creating a single and more widely available portal for DMPK scientists and medicinal chemists for use in facilitating project decisions.

Whilst discussion of proprietary information around therapeutic area or chemical space is not possible here, typical examples of project impact are described. A discovery project had reported marked differences in microsomal clearance values between rat (high) and human (low) for a structural series, whereupon GSK-A was submitted for metabolite identification using the workflow. Data reported following software-assisted interrogation (Figure 4) showed combined XICs for parent and metabolites detected in a reference (t = 0 min) incubation (Figure 4(a)) and metabolised (t = 45 min) samples for human and rat (Figure 4(b,c), respectively). Metabolites characterised in both species showed distinct differences in oxidative metabolism, where hydroxylated metabolites (M1, M2) were seen to be further oxidised to the carboxylic acid (M4), which was detected as a major response in rat (Figure 4(c)) but not detected at all in human (Figure 4(b)). Metabolite M3 represented a minor N-dealkylation pathway. The timely provision of these critical metabolism data was key in explaining clearance differences for this structural series; the utility of which helped refine decisions around drug design to further improve metabolic clearance for this project. In another example, where two novel scaffold substituents were explored, resultant metabolite identification data showed both substituents to be inherently unstable, demonstrating extensive hydrolysis and oxidative metabolism; these data guided the project team in shifting their design strategy away from these functionalities, leading to early attrition of this structural platform. Conversely, a project where the metabolic stability of linker moieties within the chemical scaffold were perceived to be a risk for high clearance, the rapid provision of metabolite identification showed that these linkers were stable, allowing increased confidence in retaining the current scaffold, driving the science around understanding metabolic clearance for this project.

Future directions

In order to further scale this approach to meet demand for higher compound numbers, whilst improving on current turnaround times but maintaining fit-for-purpose data quality and high impact, a process of constant improvement is embedded in the periodic global review of this workflow. This process allows for gaps, quick wins for increased efficiency and input from stakeholders (voice-of-the-customer, improved project relevance) to be identified, prioritised, and actioned.

Examples include better detection of predicted and non-predicted (unexpected) metabolites when using the software (e.g. more efficient automated comparison with no-drug control, improved peak integration), where better detection of unexpected metabolites (currently performed using manual data interrogation only) will help minimise any prediction-led biases; reduced manual input for fragment-ion structure rationalisation via improved automatic scoring and best-fit suggestion by the software; the ability of the software to automatically rationalise and discount false-positive responses (e.g. in-instrument fragments) which align with other drug-related ions and display sensible drug-related elemental formula; ability to perform automatic mass-defect filtering; broader in silico metabolite predictions to learn from experimental data within our proprietary database; improved reporting with less manual intervention; the embedding of shorter LC gradients for when instrument capacity eventually becomes critical; use of software to construct automated inclusion and exclusion lists for DDA routines to reduce scientist manual input and increased relevance of reporting (via voice-of-the-customer feedback). All suggested workflow changes will be tested using the 16-compound test set (either whole or in part) prior to the decision for their incorporation into a revised global workflow. Incidentally, this analytical workflow has also found use as a successful method of first intent for metabolite identification of other in vitro matrices (e.g. cytosolic S9 fractions) and in vivo samples (e.g. blood, plasma, urine) across our global organisation. Together, these iterative improvement activities continue to push towards our next step-change in biotransformation support.
for chemical design, whilst freeing up biotransformation scientists to maximise their impact through the most efficient use of their valuable expertise.

Additionally, as the quality and relevance of these data continue to improve, cross-species comparisons using in vitro techniques (e.g. cultured hepatocytes, HuhREL® hepatocytes) may find that these markedly lower drug incubation concentrations (e.g. 0.5 μM) prove to be more biologically relevant than the traditional 5–10 μM used historically in generating metabolite profiles, with the utility of this novel workflow allowing for increased translatability of pivotal in vitro metabolism data. This would allow for more accurate predictions of the overall fate and disposition of the drug and its metabolites from in vitro systems, building better models for the future.

Conclusions

A robust, sensitive, and semi-automated analytical workflow has been developed for the identification of the major routes driving in vitro metabolic clearance in early drug discovery programmes. Underpinned by in silico predictions and robust analytical methodology, this novel approach delivers highly relevant data in a timely manner to aid the design of molecules with improved ADME properties and safety profiles. The routine use of residual clearance samples at low incubation concentrations maximises data relevance and the translatability of early in vitro metabolism data, whilst eliminating the need for bespoke incubations. Robust analytical conditions and high-quality in silico metabolite predictions give confidence that the major metabolite(s) will be detected and
identified (85% success to date). The impact of resultant data from this globally aligned workflow has provided early drug discovery projects with critical insights to help refine chemical design in a data-led manner. In addition, the continued enhancement of this approach will further reduce the need for expert input into method development, sample reanalysis, and metabolite structural assignment allowing for increased sample numbers and higher throughput, ultimately reducing design–make–test cycle times for drug discovery projects.

Acknowledgements

The authors wish to extend their thanks to the many colleagues across GSK Global DMPK for their assistance in delivering this workflow and their support in the analysis of real-world samples.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The author(s) reported there is no funding associated with the work featured in this article.

References

Aburjai T, Alzweiri M, Al-Hiari YM. 2011. Temperature and pressure tuned in this article.

Ma S, Chowdhury SK. 2013. Data acquisition and data mining techniques for metabolite identification using LC coupled to high-resolution MS. Bioanalysis. 5(10):1285–1297.

McCallery DV. 2010. The challenges of the analysis of basic compounds by high performance liquid chromatography: some possible approaches for improved separations. J Chromatogr A. 1217(6): 858–880.

Mortishire-Smith RJ, O’Connor D, Castro-Perez JM, Kirby J. 2005. Accelerated throughput metabolic route screening in early drug discovery using high-resolution liquid chromatography/quadrupole time-of-flight mass spectrometry and automated data analysis. Rapid Commun Mass Spectrom. 19(18):2659–2670.

Nassar AE, Kamel AM, Clarimont C. 2004. Improving the decision-making process in the structural modification of drug candidates: enhancing metabolic stability. Drug Discov Today. 9(23):1020–1028.

Nedderman AN, Dear GJ, North S, Obach RS, Higton D. 2011. From definition to implementation: a cross-industry perspective of past, current and future MIST strategies. Xenobiotica. 41(8):605–622.

Nguyen S, Fenn JB. 2007. Gas-phase ions of solute species from charged droplets of solutions. Proc Natl Acad Sci USA. 104(4):1111–1117.

Pähler A, Brink A. 2013. Software aided approaches to structure-based metabolite identification in drug discovery and development. Drug Discov Today Technol. 10(1):e207–e217.

Paiva A, Shou WZ. 2016. Recent developments in software tools for high-throughput in vitro ADME support with high-resolution MS. Bioanalysis. 8(16):1723–1733.

Paiva AA, Klaakousci C, Li S, Johnson BM, Shu YZ, Josephs J, Zvyaga T, Zamora I, Shou WZ. 2017. Development, optimization and implementation of a centralized metabolic soft spot assay. Bioanalysis. 9(7): 541–552.

Parkinson A, Ogilvie BW, Buckley DB, Kazmi F, Parkinson O. 2019. Biotransformation of Xenobiotics. In: Klaassen CD, editor. Casarett & Doull’s toxicology: the basic science of poisons. 9th ed. New York (NY): McGraw-Hill Education.

Ruan Q, Comstock K. 2021. A new workflow for drug metabolite profiling by utilizing advanced tribrid mass spectrometry and data-processing techniques. J Am Soc Mass Spectrom. 32(8):2050–2061.

Shah P, Krens E, Nguyen DT, Obach RS, Wang AQ, Zakharov A, McKew J, Simeonov A, Hop CE, Xu X. 2016. An automated high-throughput metabolic stability assay using an integrated high-resolution accurate mass method and automated data analysis software. Drug Metab Dispos. 44(10):1653–1661.

Schadt S, Bister B, Chowdhury SK, Funk C, Hop CECA, Humphreys WG, Igarashi F, James AD, Kagan M, Khojasteh SC, et al. 2018. A decade in the MIST: learnings from investigations of drug metabolites in drug development under the “metabolites in safety testing” regulatory guidance. Drug Metab Dispos. 46(6):865–878.

Smith DA, Obach RS, Williams DP, Park BK. 2009. Clearing the MIST (metabolites in safety testing) of time: the impact of duration of administration on drug metabolite toxicity. Chem Biol Interact. 179(1): 60–67.

Smith DA, Obach RS. 2010. Metabolites: have we MIST out the importance of structure and physicochemistry? Bioanalysis. 2(7):1223–1233.

Smith DA, Beaumont K, Maurer TS, Di L. 2019. Clearance in drug design: the MIST: learnings from investigations of drug metabolites in drug development under the “metabolites in safety testing” regulatory guidance. Drug Metab Dispos. 47(5):1297–1301.

Srinivasan B. 2021. A guide to the Michaelis-Menten equation: steady state and beyond. FEBS J. 289(20):6086–6098.

Tyzack JD, Kirchmair J. 2019. Computational methods and tools to predict cytochrome P450 metabolism for drug discovery. Chem Biol Drug Des. 93(4):377–386.
new synthetic opioids acetylfentanyl, acrylfentanyl, furanylfentanyl, and 4-fluoro-isobutrylfentanyl. AAPS J. 19(4):1102–1122.

Wilkinson SD, Martin S, Orton AL, Markandu R, Jones BC. 2020. Drug metabolite identification using ultrahigh-performance liquid chromatography-ultraviolet spectroscopy and parallelized scans on a tribrid Orbitrap mass spectrometer. Rapid Commun Mass Spectrom. 34(10):8735.

Zelesky V, Schneider R, Janiszewski J, Zamora I, Ferguson J, Troutman M. 2013. Software automation tools for increased throughput metabolic soft-spot identification in early drug discovery. Bioanalysis. 5(10):1165–1179.

Zheng Z, Attygalle AB. 2021. Impact of ambient vapors present in an electrospray ionization source on gas-phase ion structures. J Am Soc Mass Spectrom. 32(3):725–735.

Zhu M, Zhang H, Humphreys WG. 2011. Drug metabolite profiling and identification by high-resolution mass spectrometry. J Biol Chem. 286(29):25419–25425.