Clinical Mass Spectrometry 8 (2018) 1–7

Contents lists available at ScienceDirect

Clinical Mass Spectrometry

journal homepage: www.elsevier.com/locate/clinms

Reproducibility assessment for a broad spectrum drug screening method from urine using liquid chromatography time-of-flight mass spectrometry

Imir G. Metushi[a,*], Michael R. Wakefield[b], Kara Lynch[c], Judy Stone[a], Robert L. Fitzgerald[a]

[a] Department of Pathology, Center for Advanced Laboratory Medicine, University of California, San Diego Health Systems, San Diego, CA, United States
[b] Waters Corp., Pleasanton, CA, United States
[c] Department of Pathology and Laboratory Medicine, University of California San Francisco, San Francisco, CA, United States

ARTICLE INFO

Keywords:
Broad spectrum screening
Mass spectrometry
High resolution
Toxicology

ABSTRACT

During the reproducibility validation for a time-of-flight (TOF) high-resolution mass spectrometry (HRMS) method set up to detect 61 drugs of abuse commonly encountered in the toxicology laboratory, it was noticed that, a number of compounds were not identified correctly during the between run analysis; the most difficult compounds to identify were norpropoxyphene, morphine, norbuprenorphine, nortriptyline, EDDP and tramadol. In subsequent patient comparison studies, screening a panel of 338 analytes, the TOF-HRMS method correctly identified 211 analytes over two runs, but did not identify 127. A total of 11 false positive results were identified by manual review of the data to be the result of confirmation ion signal-to-noise ratio(s) < 3, although one false positive that was difficult to resolve (i.e., identification of maprotiline as amitriptyline) was due to similar fragment ions and retention times. The TOF-HRMS method showed reasonable agreement with LC–MS/MS results, but there were a number of discrepant results. Additionally, the TOF-HRMS did detect five compounds missed by the LC–MS/MS methods. This extensive validation effort highlights the difficulty of analysis for certain compounds that are likely to require additional follow up prior to reporting a positive result, especially at low and high concentrations, regardless of the type of instrumentation involved.

1. Introduction

Urine drug screening is among one of the most widely practiced procedures in the clinical toxicology laboratory. While immunoassays are typically used as an initial screen, confirmation is generally required to be made with liquid chromatography coupled to at least a unit resolution (i.e., low resolution) tandem mass spectrometer (LC–MS/MS) [1]. LC–MS/MS confirmation is made by comparing retention time (or relative retention time) and ion ratios between one or more pairs of precursor and product ion(s) detected in multiple reaction monitoring (MRM) mode [2]. A primary limitation of MRM-based analytical platforms is that they are restricted to a fixed panel of targeted analytes and are unable to perform non-targeted screening. Recently, high-resolution mass spectrometry (HRMS), including techniques such as time-of-flight mass spectrometry (TOF-MS), has been proposed as an alternative that would allow non-targeted drug screening [3,4]. Besides having the ability to detect a wide range of compounds, HRMS has a much greater specificity than immunoassays, and, therefore, does not require secondary confirmation.

In previous studies, we used a HRMS instrument to identify several novel psychoactive substances encountered in our emergency department [5,6], however, rigorous method validation is required to avoid false results [7,8]. To enable non-targeted screening our laboratory and others have utilized an “all-ions” approach for compound identification with fragment ions created in the collision cell [7,9–11]. From previous study we have determined that retention time, a precursor ion, and at least one fragment ion are necessary for positive analyte identification [7]. However, our initial study was limited in scope having only evaluated single spiked concentrations and patient comparisons; a more comprehensive analysis would have included run-to-run variability, which is an important consideration when evaluating acceptability criteria.

The resolving power of HRMS has made it possible to measure *m/z* to four decimal places, providing information that can aid in calculation of elemental composition of unknowns. However, the variability in exact mass measurements, especially when analyzing samples with a complex biological matrix, has not been widely reported. Understanding the variation in exact mass measurements that would be expected for an assay routinely used in a clinical setting is important since it is a critical parameter used for compound identification. Here,

* Corresponding author.
E-mail addresses: imetushi@mail.ucsd.edu (I.G. Metushi), Michael_Wakefield@waters.com (M.R. Wakefield), kara.lynch@ucsf.edu (K. Lynch), jastone@ucsd.edu (J. Stone), rfitzgerald@ucsd.edu (R.L. Fitzgerald).

https://doi.org/10.1016/j.clinms.2018.02.002
Received 11 May 2017; Received in revised form 29 January 2018; Accepted 7 February 2018
Available online 14 February 2018

we rigorously test the reproducibility and reliability of a broad spectrum TOF-HRMS method by performing within run and between run precision studies at multiple drug concentrations for 61 different drugs, and demonstrate, by varying the exact mass tolerance of our HRMS, how the variability in exact mass measurements caused by the analysis matrix can affect analyte identification. Carryover and patient comparison studies were also conducted in order to characterize the false positive and false negative rates using specimens submitted from various sources.

2. Materials and methods

2.1. LC conditions

LC conditions were as previously described [7]. Briefly, ultra-performance liquid chromatography (UPLC; Waters) with a BEH C_{18} guard column (1.7 μm, 5 mm) and a BEH C_{18} analytical column (1.7 μm, 2.1 × 150 mm) was used for separation. Flow rate was 0.4 mL/min at a column temperature of 50 °C. Two mobile phases were used: mobile phase A (5 mmol/L ammonium formate, pH 3) and B (0.1% formic acid in acetonitrile). A gradient was used for elution starting with 5% B for 0.5 min, increased linearly up to 50% B at 6 min, and then to 95% B at 7 min where it was held for 2 min, followed by re-conditioning of the column at 5% B at 9.05 min and held for 2 min. Total injection-to-injection time for one run was 14 min.

2.2. TOF MS conditions

TOF MS conditions were as previously described [7]. Briefly, we used the Xevo G2 TOF from Waters with a resolution of 20,000 (full width at half maximum at *m/z* 400); capillary voltage, 0.8 kV; cone voltage, 20 V; extraction cone, 4 V; source block temperature, 130 °C; desolvation temperature, 550 °C; gas flow, 25 L/h; desolvation gas flow 850 L/h. Instrument calibration was performed using 5 mmol/L sodium formate in 90:10 2-propanol:water at weekly intervals (or more frequently if needed). Data was acquired in profile mode through MassLynx software v4.1, SCN 869 (Waters) without real-time mass correction. The MS method consisted of 3 functions: 1) acquisition of data over the 50–650 *m/z* range with 6 eV collision energy (low energy); 2) acquisition of data over a 50–650 *m/z* range with a collision energy ramp of 10–50 eV (high energy); 3) acquisition of lockmass data over 50–650 *m/z* range. To process data we used UNIFI v1.7.1 (Waters), which involved mass correction by using leucine-enkephalin with each sample. The UNIFI database for the 61 analytes was built from the Waters Toxicology Library and included information about molecular formula, fragment ions and retention time for each analyte.

2.3. Method comparison studies

We used two different sample sources during validation: 1) compounds were divided into six groups as previously described (10 drugs/group with the last group having 11 drugs) [7] and spiked into drug free urine (UTAK laboratories, CA) at three different concentrations of 100 ng/mL, 1000 ng/mL and 5000 ng/mL; 2) patient samples that were confirmed positive using our in-house LC-MS/MS method and additional patient samples that were sent to us from the University of California, San Francisco (UCSF).

Our in-house confirmation included MRM-based LC–MS/MS (Waters UPLC-Xevo TQ-S) methods that are routinely used to confirm immunoassay urine drug screens. Confirmation methods from UCSF included 5600 ABSciex QTOF, 3200 LC–MS/MS, Thermo Exactive Orbitrap and patient prescription records [12]. Lastly, a few samples that we were not able to confirm in-house or through UCSF were sent out to NMS (Willow Grove, PA) or ARUP labs (Salt Lake City, Utah) for targeted LC-MS/MS analysis.

2.4. Sample preparation

Sample preparation was as previously described [7]. Briefly, 200 μL of urine were used for each sample. To this was added 400 μL of deionized water, 100 μL of internal standard solution (1000 ng/mL mix of amphetamine-D5, codeine-D3, diazepam-D5, oxazepam-D5 and venlafaxine-D6 in methanol), and 300 μL of β-glucuronidase solution (5000 U/mL from Helix pomatia; Sigma-Aldrich, CA) prepared in 1.0 mol/L sodium acetate buffer (pH 5). The mixture was incubated at 50 °C for 90 min and centrifuged at 2010*g* for 10 min. 20 μL of the supernatant were injected for analysis.

For between run studies, samples were prepared fresh each day from a stock solution that was stored at −20 °C for up to two months. For patient comparison studies, in-house samples were run within two months of storage.

2.5. Validation protocol

2.5.1. Spiking studies

Within and between run validation studies consisted of five injections of each sample within the same day and 20 injections of each sample over 20 days, respectively. Three different concentrations of drugs were spiked into drug free urine: 100 ng/mL, 1000 ng/mL and 5000 ng/mL. The 61 compounds were divided into six groups with the first five groups containing 10 drugs per group and the last group containing 11 drugs.

2.5.2. Carry-over studies

Carry-over studies were performed by spiking the 61 compounds (divided into six groups, as described above) at 30,000 ng/mL and analyzing them in the following order: Blank2 → Blank1 → High1 → High2 → Blank1 → Blank 2 → Blank3 → Blank 1. Where "Blank" refers to drug free urine with no drugs spiked, and "High" refers to drug free urine spiked with 30,000 ng/mL of drug. Each "Blank" refers to a different preparation of drug free urine.

2.5.3. Patient comparison studies and proficiency testing samples

Using an IRB approved protocol (UCSD HRPP protocol number 90188), a total of 112 patient samples were collected from existing clinical specimens. Patient samples were run twice. Run number one analyzed specimens in the forward direction (e.g., sample #1 → 20) while run number two (same sample preparation procedure) analyzed specimens in the reverse direction (e.g., sample #20 → 1). Both runs (i.e., run #1 and run #2) were performed on the same day for the selected batch of samples; with run #2 immediately following run #1. Proficiency testing samples were from previous challenges (i.e., year 2013–2015).

2.5.4. Sample set-up & identification criteria

For each sample batch, the following set-up was used: (1) Wash, (2) System Suitability Test (SST), (3) Negative QC, (4) Positive QC, (5) Samples, (6) Negative QC.

Wash: 10% methanol in LC–MS grade water. SST: all five internal standards described above were spiked at 1000 ng/mL in 10% methanol in LC–MS grade water. Negative QC: drug free urine (UTAK). Positive QC: UTAK custom made in drug free urine (codeine 300 ng/mL, doxepin 300 ng/mL, norhydrocodone 300 ng/mL, ketamine 300 ng/mL, meprobamate 300 ng/mL, methylphenidate 300 ng/mL, morphine-3-β-D-glucuronide 486 ng/mL, oxazepam glucuronide 486 ng/mL, phencyclidine 300 ng/mL, norpropoxyphene 300 ng/mL).

The criteria for a positive identification were as follows: retention time match within 0.2 min, accurate mass of precursor ion within 5 ppm, at least one fragment with 10 ppm and detector counts ≥200. Any compounds identified as false positive also met all of the above criteria. UNIFI v1.7.1 was used for data processing. UNIFI involved mass correction by use of leucine encephalin and used the "all in the RT