Clinical Mass Spectrometry 4–5 (2017) 11–18

Contents lists available at ScienceDirect

Clinical Mass Spectrometry

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

Detection of drugs of abuse in urine using the Bruker Toxtyper™: Experiences in a routine clinical laboratory setting

M. Ott, K. Berbalk, T. Plecko, E. Wieland, M. Shipkova*

Zentralinstitut für Klinische Chemie und Laboratoriumsmedizin, Klinikum Stuttgart, Stuttgart, Germany

ARTICLE INFO

Keywords:
Drugs of abuse screening
Assay performance
Ion trap mass spectrometry
Routine laboratory

ABSTRACT

Urine screening can be used to detect misuse of illicit drugs and validate opioid replacement therapy compliance. It is common that immunochemical assays are combined with GC-MS for these applications. Bruker has recently released an ion trap mass spectrometer, called Toxtyper™, with the potential to replace current screening algorithms to detect drug misuse.

Here, we compare our current strategy of urine screening for misuse of cannabis, amphetamines, cocaine, opiates, benzodiazepine, methadone, sufentanil, and pregabalin to the Toxtyper protocols provided by the manufacturer.

The analytical performance of the instrument was determined on a selected drug panel and with 188 urine samples being compared to establish concordance between our currently established approach and the Toxtyper.

The lower limits of detection and identification for acetylcodeine, amphetamine, benzoylecgonine, methadone, and nordiazepam were below the common cut-offs for immunological screening assays and comparable to GC-MS. Imprecision and accuracy, both within- and between-series, were consistently < 25%. Toxtyper screening for pregabalin and sufentail was less sensitive than a targeted LC-MS/MS assay. Concordance met the predefined criterion of > 90% for all drugs, except for pregabalin. Cannabis misuse could not be detected due to the limited sensitivity of the Toxtyper assay protocols used and the inherent imprecision of the assay.

Our study has revealed that a considerable portion of our current time-consuming protocol for screening drugs of abuse in urine, based on the combination of multiple analytical methods, could be consolidated by the Toxtyper for a majority of the most-relevant substances in our patient population.

1. Introduction

Drugs of abuse (DOA) screening in urine is usually performed by immunoassay. However, the number of validated immunoassays is limited, being largely restricted to well-known substances. The catalogue of misused drugs is continually in flux as new DOA are continually being developed [1,2]. Although the number of illicit drugs is growing, the development of applicable immunoassays has not kept up with their rapid development [3]. These factors present a challenge for clinical laboratories serving emergency rooms and hospitals that are dedicated to treating patients with suspected drug abuse or who are under opioid substitution treatment. To address this challenge, chromatographic methods such as gas chromatography (GC) or liquid chromatography (LC) coupled to mass spectrometry (MS) can be employed to detect substances for which no immunoassays are commercially available, as well as to confirm immunoassay results.

Specialised laboratories focussed on forensic drug screening and confirmation have established sophisticated LC and MS assay protocols and provide dedicated analytical equipment and staff [4]. The analytical approaches of these labs cover non-target analysis, suspect analysis, and target analysis with LC-MS equipment, such as ion-traps, triple-stage quadrupoles and, sometimes, high resolution mass spectrometry [2,5]. While it is outside the scope of this paper to discuss the advantages and limitations of the various MS approaches used in forensic toxicology, an excellent review article has been assembled by Meyer et al. [6]. Routine clinical laboratories in maximal care hospitals commonly rely on immunological screens for a limited number of analytes and then send out the samples to specialized laboratories for confirmation analysis, as well as extended screening. This send out approach suffers from an extended turnaround time, which generally impedes incorporation of the screening results into the clinical decision algorithm. An alternative is the establishment of a larger central laboratory that is equipped with well-trained staff and contemporary analytical instrumentation. However, maintaining a specialised analytical unit locally is a challenge due to the level of effort required to develop, validate, and continuously update the laboratory developed

* Corresponding author at: Central Insitute for Laboratory Medicine and Clinical Chemistry, Klinikum Stuttgart, Kriegsbergstrasse 62, D-70174 Stuttgart, Germany.
E-mail address: m.shipkova@klinikum-stuttgart.de (M. Shipkova).

http://dx.doi.org/10.1016/j.clinms.2017.08.002
Received 17 April 2017; Received in revised form 17 August 2017; Accepted 18 August 2017
Available online 19 August 2017

tests (LDT) in order to maintain analytical competency for the continually evolving products offered on the illicit drug scene [2,7]. Furthermore, while GC-MS is generally the most affordable and well-established MS equipment available, it is also more labour intensive than immunoassays and more time-consuming due to the need for complex sample extraction and derivatization steps [8].

Any analytical advance that would minimize and/or standardize these efforts, without compromising analytical quality, would be of value. An instrument recently released by Bruker, the Toxtyper, appears to provide these desired characteristics [9]. The Toxtyper is an LC-MS instrument based on an ion trap mass analyser (IT-MS) coupled to an ultra-high-performance-liquid-chromatography (UHPLC) device. In addition to standard operational software that facilitates set-up of LDTs, it is also equipped to enable operation by non-LC-MS experts for routine purposes. Additionally, ready-to-use assay protocols, which include UHPLC and MS set-ups, as well as algorithms for automatic report generation, are provided. Interpretation of MS spectra relies on matching to reference spectra found in libraries that are either provided by the manufacturer (e.g., Drugs of Abuse Library, DOAL or Toxtyper Library) or are commercially available (e.g., Maurer/Wissenbach/Weber; MWW, Wiley-VCH, Weinheim, Germany, 2014) [10]. Furthermore, customized libraries that include spectra of newly identified substances of interest can be developed, and the acquired spectra shared between laboratories to increase the ability of the lab network to address the appearance of novel DOAs.

A disadvantage is that the combination of the analytical instrument and software package is not approved or registered as an in vitro diagnostic medical device; therefore, clinical laboratories must perform a full validation of all analytical protocols, in a way that is common for any other LDT, if the results are intended to be used for medical decision making. Additionally, standardized pre-analytical protocols for specimens such as urine, blood, or salvia are not available, hence, laboratories must establish and optimize their in-house sample preparation procedures in order to meet local analytical requirements.

Here, we present results following the establishment and validation of a Toxtyper-based urine DOA screening protocol in our laboratory. The protocol is intended to provide qualitative test results (i.e., positive/negative). It includes a simple protein precipitation step that is amenable to automation, and which can be easily performed in a routine setting by non-specialized lab technicians. We compared this new screening protocol with our established screening algorithms based on immunochemical methods applied to autoanalyzers in the general clinical chemistry laboratory, as well as on GC-MS and LC-MS/MS-based protocols used in our laboratory to screen for illicit drugs and their metabolites in urine samples. The method was validated using a panel of the most commonly encountered DOA in our routine service to spike drug-free urine samples; finally, concordance between the current screening algorithm and the new Toxtyper-based strategy was evaluated using urine from patients undergoing opioid replacement therapy.

2. Material and methods

2.1. Chemicals and reagents

(±)-Amphetamine, (±)-11-nor-9-carboxy-delta 9-THC, (±)-methadone, buprenorphine, sufentanil citrate were obtained from Cerilliant Corporation (Round Rock, Texas, USA). Nordazepam, benzoylecgonine, and acetylcodeine were from LGC Standards (Teddington, Middlesex, UK). Pregabalin was from TLC Pharmaceutical Standards Ltd (Aurora, Ontario, Canada). As internal standard (IS), Diazepam-D5 from Cerilliant Corporation (Round Rock, Texas, USA) was used. Liquichek Urine Toxicology Control level C1 and C3 quality control materials were from Bio-Rad Laboratories Inc. (München, Germany). LC-MS grade acetonitrile, methanol and formic were obtained from Biosolve (Valkenswaard, The Netherlands). ß-Glucuronidase from *Helix pomatia* was obtained from Sigma-Aldrich (Munich, Germany). Ammonium formate, sodium hydroxide and potassium dihydrogen phosphate were obtained from VWR (Darmstadt, Germany).

2.2. Toxtyper™ instrument and libraries

Analyses performed with the Toxtyper LC/IT-MS System (liquid chromatography with ion trap mass spectrometric detection) used protocols provided by the manufacturer [11,12]. Screening with this system is based on the LC-MS^n ion trap technology and MS^n library search. It consists of a Dionex™ Ultimate 3000 UHPLC (ThermoFisher Scientific, TF, Dreieich, Germany) coupled to an amaZon speed™ (Bruker, Bremen, Germany) ion trap with an electrospray ionisation source. The gradient elution was performed on a C_{18} column (Column Kit Screening Bruker, Bremen, Germany). According to the recommendations of the manufacturer the mobile phase A was composed of 1% acetonitrile and 0.1% formic acid in 2 mmol/l aqueous ammonium formate solution. The mobile phase B consisted of 0.1% formic acid and 2 mmol/l ammonium formate in acetonitrile. The run time per sample was 19 min and the injection volume was 25 µl. The flow rate was set to 0.5 ml/min with a gradient from 1% to 99% B [13]. DOA detection and identification was based on retention time, full scan MS, MS2, and MS3 information using two different library search approaches. The first (“MWW” screening) was a combined library search using the Toxtyper library (~900 compounds, mainly parent compounds) and the MWW library (> 4.500 entries including metabolites and conjugates). The second was the Drug of Abuse library screening (“DOAL” screening) with 83 entries of mainly parent compounds, adopted from the Bruker Toxtyper System. For this purpose, the same extracted samples were injected twice. Whereas the “MWW” method is configured as a non-targeted screening approach, the “DOA” method is a targeted screening approach and MS^n acquisition is data-dependent with a scheduled precursor list. Active exclusion after one spectrum and reconsideration if the intensity increased fivefold was used. The recording of the MS^n spectra was based on the “UltraScan” mode (i.e., 32,500 m/z sec^{-1}) with a mass range from 70 m/z to 800 m/z. The MS tuning was based on the SmartFrag™ technology and used 300 m/z as the target mass [11,12]. Continuous positive/negative ionisation mode switching was applied with the “DOA” screening, while “MWW” screening used positive ionisation mode only. For library search and reporting, the data sets were post-processed using the Bruker Compass® Data Analysis (DA) software package 4.1 and the processed spectra were submitted to the DA 4.1 library module [11,12]. Only hits identified with purity above 800 and deviation of ± 0.3 min from the target retention were considered positive. For the daily instrument performance check, a Toxtyper-QC test solution (SST BDAL MIX, Bruker, Bremen, Germany/LGC Standards, Teddington, Middlesex, UK) was injected followed by the injection of a blank sample to check for carryover. The Toxtyper QC test solution was analysed with a specific Toxtyper QC method and included six substances (i.e., paracetamol, hydrochlorothiazide, mirtazapine, risperidone, d4-haloperidol and desalkylflurazepam), which had to be reliably detected to meet specifications. In addition, an internal standard (i.e., diazepam-D5) was added to each sample and used to control for consistency of the sample preparation and for instrument fluctuations. Hystar® software (Bruker, Bremen, Germany) was used for peak integration and quantitative evaluation, but only for method validation.

2.3. Sample preparation for the Toxtyper analysis

For the development of the sample pre-treatment procedure, a urine sample spiked with methadone, acetylcodeine, amphetamine, nordiazepam, 11-nor-9-carboxy-delta 9-THC, and benzoylecgonine (100 µg/L each), as well as 2.0 mg/L pregabaline and 5.0 µg/L sufentanil, was used. Various ratios of sample volume to added volume, of either

acetonitrile or methanol as precipitation reagent, were tested. In addition, the effect of adding formic acid (1% in acetonitrile) was investigated, but found to be without beneficial impact on sample clean-up. Based on these optimization experiments, the sample preparation protocol was established as follows:

100 µL of urine sample are added to a 1.5 mL polypropylene reaction tube (Sarstedt, Nümbrecht, Germany) with 10 µL ß-glucuronidase solution (10 kU/mL in 100 mmol/L phosphate-buffer, pH 5; Sigma-Aldrich, Steinheim, Germany) to break down glucuronide metabolites. After brief vortex mixing (i.e., 30 s) the samples are incubated at 60 °C for 60 min. 100 µL of the hydrolysed sample are then transferred to a clean reaction tube (Sarstedt). After addition of 125 µL IS solution (100 µg/L diazepam-d5 in acetonitrile) samples are vortex-mixed again for 30 s and centrifuged at 13,000*g* for 5 min. Finally, 200 µL of the supernatant are transferred to a clean 1.5 mL glass vial (Wicom, Heppenheim, Germany) and mixed with 800 µl mobile phase A followed by injection of 25 µL of the diluted sample onto the UHPLC column.

2.4. Validation

To validate the appropriateness of the analytical system for the demands of our local screening routine we investigated imprecision, inaccuracy, linearity, lower limit of identification (LLOI), lower limit of detection (LLOD), matrix effects, and carryover with a set of pre-selected test substances. This set was based on the prevalence of abused drugs detected in urine samples sent to our laboratory and included nordiazepam, acetylcodeine, benzoylecgonine, (±)-11-nor-9-carboxy-delta 9-THC, amphetamine, methadone, sufentanil, and pregabalin. Although the IT-MS method is intended for a qualitative screen, when using the Hystar® software the Toxtyper instrument permits numerical recording of peak areas; this allowed statistical evaluation of the analytical performance characteristics. Anonymized drug-free negative urine samples collected from willing laboratory staff were used for spiking experiments, except for the limit of detection and matrix effect studies, which were established with anonymized patient urines void of DOA. In addition, concordance between the current screening algorithm and the new Toxtyper-based strategy was evaluated using anonymized patient urines.

2.5. Precision and accuracy

Liquichek Urine Toxicology Control levels C1 and C3 (BioRad, München, Germany), which were spiked with test substances (i.e., buprenorphine, acetylcodeine, pregabaline, sufentanil) to achieve appropriate concentration levels, were used, as necessary, to test within- and between-batch precision and accuracy over six runs at two concentrations for each test compound (Table 1). Control materials were prepared from human urine and contained gravimetric concentrations of drugs and their metabolites. While the concentrations of C1 were below common cut-offs used for immunochemical screening tests, those of C3 were above; the only exception was amphetamine, for which C3 had the lower concentration level and C1 the higher. To determine imprecision, the coefficient of variation (CV %) was calculated. To establish accuracy, the bias was calculated as percent deviation from the assigned values of controls. The acceptance criteria were set at CV ≤ 25% for analytical imprecision (within and between batches) and ≤25% for the bias from the target value of controls. All single experiments to test between-batch imprecision and accuracy were run on different days.

2.6. Linearity

Linearity was assessed by fivefold analysis of a urine matrix spiked with a minimum of six concentration levels per test substance, plus a blank sample without analytes. The concentration ranged up to 1000 µg/L for amphetamine, nordiazepam, benzoylecgonine, acetylcodeine, and methadone; for sufentanil and pregabaline it ranged up to 50 µg/L and 10 mg/L, respectively. All five experiments to test the linearity were run on different days. For evaluation of linearity, the acceptance criterion of r ≥ 0.98 in a Spearman's Rank correlation was chosen.

2.7. LLOD and LLOI

To evaluate the detection performance of the "DOA" and "MWW" screening approaches on the test substances, the LLOD and the LLOI were first determined theoretically by calculation of the slope of the calibration curves and a significance level of 95% [14]. In a second step, the experimental LLOD was established using serial dilution of spiked urines with six different urine matrices (i.e., creatinine concentration varying between 50 and 350 mg/dl) and six runs per matrix. The LLOD was defined as the lowest concentration from the dilution series that was detected in at least 80% of the experiments. This mirrors the 20% CV acceptance limit for the lower limit of quantification (LLOQ) in quantitative methods [15]. We decided to utilize this two-step approach because the first step to establish the LLOI was only possible with quantitative results, which required use of the Hystar® software. However, this does not reflect the actual analytical workflow for qualitative screening with the "DOAL" and the "MWW" library. Therefore, we re-checked the LLOD with the six urine matrices as mentioned above.

2.8. Matrix effects

Matrix effects were tested using the post-column infusion method [15] and were based on four different urine sample matrices. To detect matrix effects a solution consisting of pregabalin, morphine, amphetamine, benzoylecgonine, methadone, sufentanil, morphine (50 mg/L) each and nordiazepam (10 mg/L) was infused between the UHPLC column and the ESI source to visualize ion suppression or enhancement effects.

2.9. Carryover

Carryover effects were evaluated by running blank samples immediately following analysis of the highest concentration drug-spiked urine samples during the linearity experiment (n = 8).

2.10. Concordance

To evaluate concordance between our established screening algorithm and the new Toxtyper screening strategy, based on the "MWW" or "DOAL" methods, urine samples from patients undergoing opioid replacement therapy were used. Our established screening strategy is based on automated group, as well as targeted, immunochemical assays run on Dimension Vista autoanalyzers (Syva EMIT, Siemens Healthineers, Eschborn, Germany), which are complemented by in-house validated GC-MS LDTs for amphetamine and methamphetamine, ecstasy, cocaine, opiates, and benzodiazepines. In addition, concordance was investigated for pregabalin versus an in-house LC-MS/MS method since the drug is not covered by immunoassays. For screening of methadone and its metabolite, 2-ethylidene-1,5-dimethyl-3,3-diphenylpyrrolidine (EDDP), we only used GC-MS because our immunoassay cannot detect EDDP, the presence of which is mandatory to prove compliance during opioid dependency maintenance treatment [16]. For the purpose of method comparison, a minimum of 20 positive and negative samples per analyte group (i.e., 188 urine samples in total) were analysed and an overall performance acceptance rate for concordance of ≥90% was predefined. Sufentanil was not detected in any of the patient urine samples and was, therefore, excluded from the comparison study. Between the analytical runs with the different methods, the urine samples were stored at −20 °C. Toxtyper screening

Table 1
Precision and accuracy of the Toxtyper "MWW" and "DOAL" methods.

Test Substance	Method	Liquichek C1				Liquichek C3			
		Concentration (µg/L)	Bias (%, n = 6)	CV (%, n = 6)		Concentration (µg/L)	Bias (%, n = 6)	CV (%, n = 6)	
			Between-Batch	Within-Batch	Between-Batch		Between-Batch	Within-Batch	Between-Batch
Acetylcodeine	MWW	100	8.78	4.74	8.29	350	3.88	3.98	10.43
	DOAL		11.79	6.77	4.10		4.34	5.82	7.17
Amphetamine	MWW	399	10.59	13.80	8.63	282	3.19	10.18	12.28
	DOAL		1.27	5.27	14.46		2.90	7.79	15.46
Benzoylecgonine	MWW	34	5.70	12.76	9.36	117	8.59	7.99	10.92
	DOAL		10.46	5.86	6.45		2.79	7.23	2.76
Methadone	MWW	90	12.40	9.33	23.02	315	12.58	4.87	8.74
	DOAL		10.22	8.33	12.34		8.76	7.54	8.30
Nordiazepam	MWW	70	17.10	6.96	7.37	308	8.52	1.66	12.62
	DOAL		20.80	2.59	13.00		−14.56	9.20	6.95
Sufentanil	MWW	3	2.17	5.33	7.88	6	−0.94	9.76	10.96
	DOAL		−1.83	9.32	11.92		−0.69	9.57	11.59
Pregabalin	MWW	2000	12.00	3.21	9.84	10000	−5.98	10.50	15.04
	DOAL		18.17	10.34	21.57		−6.63	9.33	10.82

was performed both with the "MWW" and with the "DOAL" screening method. The concordance of positive, negative, and overall results between the new approach and our established screening procedure was calculated based on 2 × 2 contingency tables constructed separately for each group of analytes and each Toxtyper method.

2.11. Statistics

Statistical data evaluation was supported using the MedCalc Software (Ostend, Belgium).

3. Results

Fig. 1 shows a typical chromatographic run generated with the Toxtyper using the chromatographic conditions applied to both the "DOAL" and "MWW" screening conditions. In the representative urine sample thirteen drugs and metabolites were reported.

3.1. Precision and accuracy

Precision and accuracy, as observed for the two Toxtyper methods, are shown in Table 1. For all analytes, imprecision was found to comply with the predefined acceptance criteria (i.e., CV ≤ 25%) for both "MWW" and "DOAL" methods. Furthermore, except for methadone and pregabalin, they were in agreement with acceptance criteria typically required for quantitative methods (i.e., CV ≤ 20%) [14,20]. The between-batch imprecision at the lower concentration level (C1) with the "MWW" application for methadone and with the "DOAL" application for pregabalin were 23% and 21.6%, respectively. At the C3 level both were ≤11%. In addition, the bias of 20.8% observed for nordiazepam with the C1 quality control material and the "DOAL" method was slightly outside the 20% limit. Reliable detection of 11-nor-9-carboxy-delta9-THC with the "MWW" method was not possible due to the exclusive positive ionisation; the "DOAL" method, which is more sensitive because it alternates between positive and negative mode demonstrated an imprecision and inaccuracy for the analyte of > 50%, vastly exceeding the acceptance criteria, hence, further experimentation for this analyte was not performed.

3.2. Linearity

For all compounds, as well as for both methods, the predefined acceptance criteria for linearity were met. Supplementary Fig. 1 demonstrates the results of the linearity evaluation for all test substances investigated with the "MWW" and the "DOAL" settings. The diagrams show the correlation between the drug concentrations and the peak area with a 95% confidence interval.

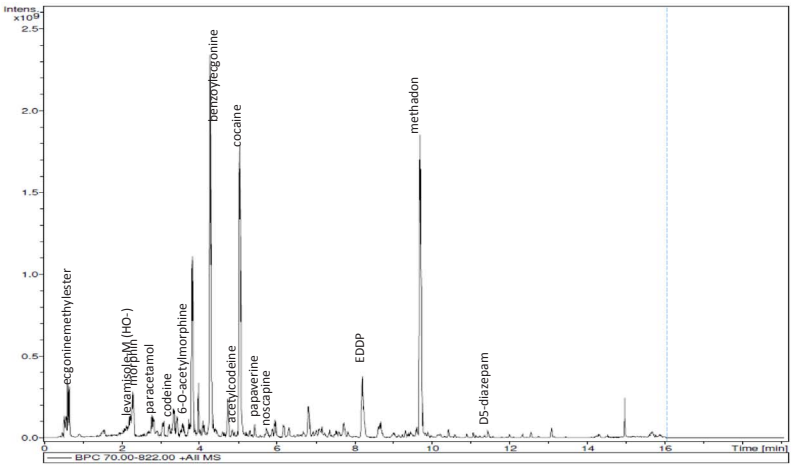

Fig. 1. Chromatogramm of a patient urine sample containing various drugs and metabolites.

Table 2
Lower limit of identification (LLOI) and lower limit of detection (LLOD) for the test substances as found with the Toxtyper "MWW" and DOAL" method compared to cut-off concentration of the respective immunochemical methods and the LLOD of the chromatographic methods with mass-spectrometric detection previously established in the laboratory.*

Analyte	LLOI/LLOD Toxtyper "MWW" [µg/L]	LLOI/LLOD Toxtyper "DOAL" [µg/L]	Cut Off Immunoassay [µg/L]	LLOD of GC- or LC with MS Detection [µg/L]
Acetylcodeine	11/40	10/20	300	10
Amphetamine	n.d.	100/200	500	150
Benzoylecgonine	22/50	9/10	300	75[1]
Methadone	3/5	2/5	300	40
Nordiazepam	80/200	50/100	200	125
Sufentanil	2/10	1/1	n.a.	1
Pregabalin	430/2000	430/500	n.a.	200

Urinary creatinine concentrations were between 50 and 350 mg/dl. n.d. = not determined; n.a. = not applicable. LLOI = lower limit of identification; LLOD = lower limit of detection.
[1] confirmation by GC-MS was based on cocaine detection.
* LLOI was calculated, LLOD was experimentally established.

3.3. LLOIs and LLODs

Table 2 summarizes the LLOIs calculated using the data generated with the Hystar® software and the LLODs based on the 20% CV acceptance limit for the test substances using the "MWW" and the "DOAL" method and compares them to the cut-off concentrations used with either the automated immunoassay or chromatographic (i.e., GC-MS and LC-MS/MS) LDTs, both current screening methods established in our laboratory. In general the calculated LLOIs were lower than the experimentally derived LLODs. As expected, the "DOAL" method achieved lower LLODs when compared to the "MWW" method for all substances tested. The LLODs found with the two Toxtyper methods were well below the cut-offs of the immunoassays we use and similar to our chromatographic LDTs, except for sufentanil and pregabalin, for which both Toxtyper methods were less sensitive. Overall, the "DOAL" procedure was superior compared to the "MWW" procedure.

3.4. Matrix effects

The post-column infusion experiment using four different patient urine samples, which were previously established to be drug-free, did not uncover any significant suppression and/or enhancement matrix effects over the analytical run. Fig. 2 shows the extracted ion chromatograms (EIC) of the infused analytes acquired during a post-column infusion experiment with one of the drug-free urine matrices, as an example.

3.5. Carryover

Carryover effects for methadone and benzoylecgonine were observed in analyses that followed urine samples with very high concentrations of the drugs. The origin appeared to be related to the IT-MS unit since hits detected did not match the expected retention time for these two substances; a fact that we used to prevent incorrect result interpretation.

3.6. Result concordance between our established screening strategy and the new Toxtyper approaches

Concordance of results between the previously established screening methods in our laboratory and the new Toxtyper protocols, including the "MWW" and the "DOAL" screening methods, was investigated using 188 urine samples from our routine service. For samples analyzed by both immunoassay and GC-MS, only those with consistent results are represented in Table 3. This selectivity was intended to increase the probability of true negative and true positive results since a third gold standard comparator assay was not available. Comparisons were focused primarily on amphetamines, methamphetamines, ecstasy, opiates (i.e., morphine, codeine, acetylmorphine and acetylcodeine), cocaine (i.e., cocaine and benzoylecgonine), benzodiazepines (i.e., diazepam, nordiazepam, temazepam, oxazepam, bromazepam, lorazepam, midazolam and 7-aminoflunitrazepam), methadone (plus EDDP), and pregabalin. Immunoassay and GC-MS results were available for most analyte groups, except for pregabalin, which was determined only by LC-MS/MS. Methadone and its metabolite EDDP were only determined by GC/MS. As shown in Table 3, there was, in general, high total concordance ranging from 95% to100% between the results generated with our established screening protocols and the new Toxtyper methods, which met our predefined acceptance criteria of a total concordance ≥90%. However, the "MWW" method showed a lower concordance with samples positive for amphetamines (i.e., 85% concordance) or pregabalin (i.e., 89% concordance). In general, the concordance match with the "DOAL" method was better than with the "MWW" method. Positive concordance for the "DOAL" method was 100% for MDA/MDMA, opiates, cocaine, and methadone. For amphetamines, benzodiazepines, and pregabalin it ranged from 97.5% to 98.6%. Regarding the detection of cocaine abuse, the Toxtyper found more samples to be positive than that by GC-MS or immunochemical examination, with the "MWW" and "DOAL" methods producing almost the same results. The identification of methadone via GC-MS and Toxtyper methods was comparable, independent of the Toxtyper procedure used. The detection of pregabalin by our LDT, based on LC-MS/MS, produced 55 positive and 133 negative findings. The "MWW" Toxtyper method showed a lower concordance with these results than the "DOAL" method, however the "DOAL" method produced a

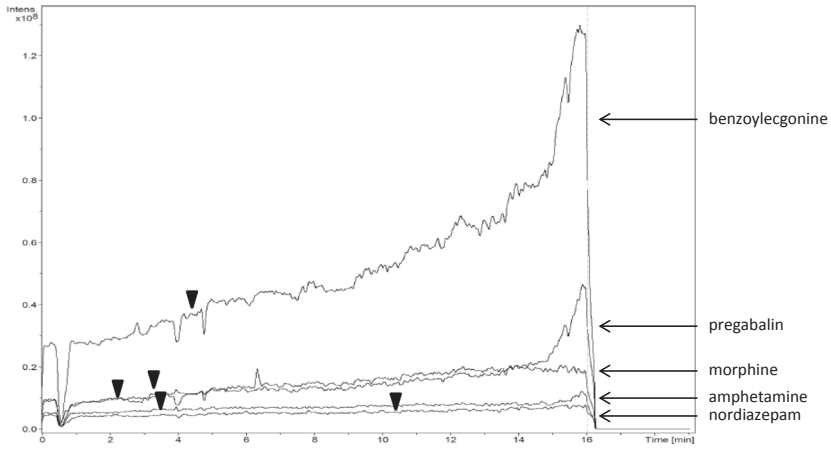

Fig. 2. Extracted ion chromatograms (EIC) of the infused analytes acquired during a post-column infusion experiment with one of the drug-free urine matrices as an example. Triangles show the retention times of the respective substances (A: benzoylecgonine, pregabalin, morphine, amphetamine, nordiazepam; B: sufentanil, methadone, acetylcodeine).

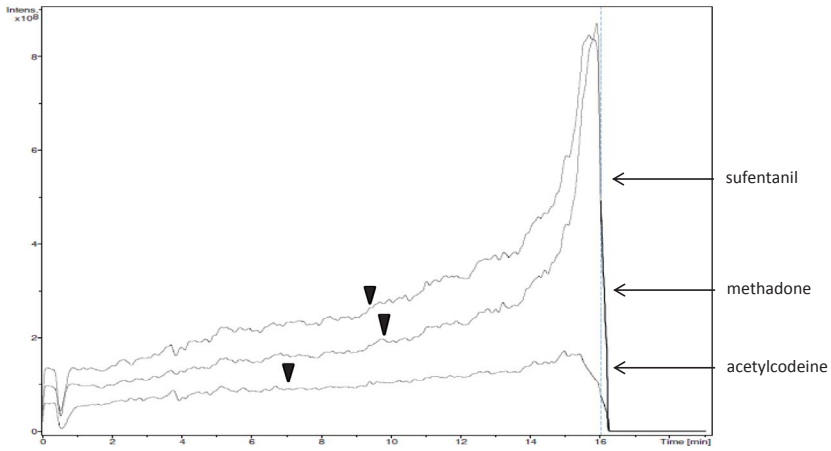

Fig. 2. (*continued*)

relatively high number of unclear findings (n = 19, qualified as tentative by the Data Analysis software). Tentative results were considered negative for data evaluation.

To avoid potential interpretation bias from combining immunoassay and GC-MS results, we selectively compared the Toxtyper methods against GC-MS alone for amphetamine/metamphetamine, MDA/MDMA, opiates, cocaine/benzoylecgonine, and benzodiapezpines. We found that the Toxtyper, with the "DOAL" application, detected more positive samples for all of these drug groups than GC-MS alone. This was also true when the "MWW" library search was used for cocaine/benzoylecgonine and MDA/MDMA, whereas for amphetamine/metamphetamine and opiates more samples were detected as positive with GC-MS versus the "MWW" method, demonstrating a positive concordance of 84% and 92%, respectively. The total concordance between the two techniques was ≥97% when using both the "DOAL" and "MWW" methods.

4. Discussion

In the present work, we evaluated the Toxtyper ion trap LC-MS instrument for screening of DOA in urine at a medical laboratory involved in the surveillance of drug addicts under substitution therapy. The laboratory also serves a tertiary care hospital with an emergency room and several intensive care units, as well as being involved in the recruitment of brain death organ donors, which requires a rapid clinical toxicological service. DOA testing with a relatively short turnaround time is commonly achieved using automated immunological screens on clinical chemistry analyzers, however, it is well-known that immunoassays suffer from false negative and false positive results, the rate depending on the cross-reactivity of the antibodies used by the assay manufacturer and on the recommended cut-off levels [17,18]. These immunoassay results, therefore, require confirmation by a chromatographic method [8]. Additionally, because immunoassays have not been developed for many drugs, particularly for new psychoactive substances, immunoassay screening is frequently complemented by GC-MS, as well as LC-MS/MS, techniques [19]. These chromatographic approaches are time consuming and require domain expertise. To accommodate analysis of the broad spectrum of illicit drugs available, our laboratory combines immunoassays, GC-MS, and targeted LC-MS procedures. While we are aware that this approach does not address newer psychoactive drugs, which can be reliably detected only by high resolution mass spectrometry [2,6], incorporation of this level of analysis would make it difficult to achieve timely delivery of results within 24 h on a 365 day basis.

The Toxtyper, which comes with predefined assay protocols and user-friendly software, is a step toward routine application of a MS-based DOA screening in urine. Analytical runs can be completed within 20 min and the instrument is easy to operate making it suitable for use by non-specialized laboratory technicians after a relatively short training period. For evaluation of MS spectra, targeted and untargeted approaches are available; however, the applications are neither certified for use under the European medicinal products law, nor cleared by the US FDA for clinical decision making. Therefore, each laboratory must validate the assays of interest and demonstrate their fitness for local analytical purposes. In principle, qualitative assays are only required to demonstrate that they are sufficiently sensitive and specific to detect the analytes of interest. The detection limits of the conformation assays should be lower than the cut-offs of the screening procedures [20]. When setting up our validation strategy we considered, and partly adopted, more general recommendations from guidelines for analytical

Table 3
Concordance of results between the Toxtyper procedures and the established screening protocol in the laboratory.

Test substance	Number of samples	Concordance of results using the "MWW" method (%)			Concordance of results using the "DOAL" method (%)		
		positive	negative	total	positive	negative	total
Amphetamine /methamphetamine[1)]	176	85.0	99.3	96.0	97.5	98.5	98.3
MDA /MDMA[1)]	173	95.8	100	99.4	100	100	100
Opiates (morphine, codeine, acetylmorphine, acetylcodeine)[1)]	186	93.4	100	97.8	100	100	100
Cocaine/benzoylecgonine[1)]	176	100	96.2	97.2	100	97.0	97.7
Benzodiazepines (diazepam, nordiazepam, oxazepam, temazepam, bromazepam, lorazepam, midazolam, 7-aminoflunitrazepam)[1)]	185	90.4	100	94.6	98.6	97.3	97.8
Methadone/EDDP[2)]	186	100	99.1	99.5	100	99.1	99.5
Pregabalin[3)]	188	89.1	100	96.8	98.1	99.3	98.9

[1] vs. immunoassay confirmed by GC/MS.
[2] vs. GC/MS.
[3] vs. LC-MS/MS.

method validation published by CLSI, EMA, or UNODC [15,21–23] and a publication from Chau et al. [24].

We found that the optimal sample preparation procedure for the Toxtyper, and the best compromise for all substances of interest, was protein precipitation using acetonitrile followed by high dilution of the protein-free supernatant in mobile phase A. Although a lower level of dilution could possibly result in a better LLOD for some of the drugs, working with a higher concentration of acetonitrile in the injected sample compromised chromatographic performance since the pre-defined gradient of the mobile phases starts with 1% acetonitrile. Evaporation of the organic solution was not a favourable alternative since it complicates sample pre-treatment and cannot be automated in our laboratory setting; amenability to automation was one of our key aims during optimization of the sample pre-treatment protocol.

The instrument showed some carryover effects, particularly when negative samples were analyzed following samples with high concentrations of methadone or benzoylecgonine. These carryover incidents could be recognised and accounted for due to retention times of the "false" hits being atypical. Cleaning of the MS source, part of the instrument maintenance protocol, was helpful in limiting carryover effects. Serious matrix effects were excluded by the post-column infusion method. However, we have assumed that some level of matrix effect was still present because, during our experiments to establish LLODs, the detectability of the majority of test substances was lower in urine matrices with higher creatinine concentrations than those with lower concentrations. Although a more powerful sample clean up (e.g., solid phase extraction) could be expected to more effectively remove matrix effects, it would complicate the procedure and increase cost; however, such a step might be necessary to accommodate local laboratory requirements.

Although our initial intent was to validate a qualitative screening method, we ended up using quantitative peak-area data from the Hystar® software to investigate linearity, assay precision and accuracy. We noted a linear relationship between the signal acquired and the concentration of all test substances. Based on this approach, it was possible to establish the low LLOI for selected substances. Because the Hystar® software is not part of the screening algorithm, we verified the LLOD experimentally in various patient urines. The concentrations used covered the expected ranges found in the urine of habitual, as well as single, drug users and were well below the common cut-offs for immunoassays used for DOA screening in urine. The predefined acceptance criterion of $r > 0.98$ for linearity was met for all test compounds. Based on these results, we considered the numeric results from the Toxtyper screening assays to be roughly indicative of a high or low concentration, if the ratio to the internal standard could be calculated.

The LLODs determined with the new assay protocol for the IT-MS instrument were consistently a factor of two better than the cut-offs used with the screening immunoassays. In general, more favourable results were seen with the "DOAL" method. However, when the qualitative Toxtyper methods were compared to our validated quantitative target LC-MS/MS methods for sufentanil and pregabalin, their performance was inferior. Since low concentrations are typical for sufentanil, the lower LLOD of the "DOAL" method is likely the best option for screening for this drug.

11-nor-9-carboxy-delta 9-THC could not be reliably detected with either the "MWW" or "DOAL" library due to insufficient sensitivity. Therefore, we did not perform a full validation and decided to retain our established immunoassay screen and LC-MS/MS confirmation to identify suspected cannabis consumption.

Because the Toxtyper methods allowed signal quantification, as demonstrated in the linearity experiments, we decided to use these data to evaluate imprecision and accuracy. However, because clinical toxicological screening is qualitative it is not required to meet the narrow acceptance criteria of quantitative methods. We, therefore, defined the acceptance criteria for our evaluation to be ≤ 2 5% for both performance characteristics; this is slightly above the target levels recommended for quantitative methods for bias and imprecision (i.e., $\leq 15\%$ and $\leq 20\%$, respectively, for the LLOD) [15,22]. This criterion was achieved for all test substances with the two Toxtyper methods; in fact, for most of the analytes the results were below 20% and, therefore, near the requirements for quantitative procedures. Contrary to our expectations there was no clear trend in favour of one of the two Toxtyper methods regarding precision and bias, as was seen regarding the LLOD in the case of the "DOAL" method.

Using 188 urine samples from our routine DOA screening programme we found excellent concordance between the established screening protocols in our laboratory and the new approach using the Toxtyper. Overall concordance ranged from 94.6% to 99.5% for the "MWW" method, and 97.7% to 100% for the "DOAL" method. The "MWW" method showed a tenuous concordance with positive results for amphetamines (i.e., 85%) and pregabalin (i.e., 89%). To address this, we decided to only use the "DOAL" method for confirmation of a positive result from the immunochemical screen of amphetamines, if the "MWW" approach fails confirmation. In the case of pregabalin the superiority of the quantitative LC-MS/MS method over the IT-MS screen must be also considered and we accepted the positive concordance (i.e., 89%), which only slightly violated our criterion of $\geq 90\%$ because, in our experience, this is still sufficient for our qualitative diagnostic purposes. When Toxtyper screening with the "DOAL" and "MWW" methods was compared only to the GC-MS screening results of the 188 urine samples we found that the "DOAL" method always detected more drugs in urine than the GC-MS method, while the "MWW" method was inferior for amphetamines and opiates.

It is plausible that the "MWW" method suffers from a limited diagnostic sensitivity due to the higher LLODs when compared to the "DOAL" method, which could be explained by the more comprehensive library and untargeted screening approach of the "MWW" method. Regardless, the true negative and positive rates cannot be determined because a gold standard reference method, such as high resolution LC-MS/MS, was not available [2,6].

The integration of the Toxtyper IT-MS into our screening algorithm has resulted in a consolidation of methods and instruments. Compared with our previous strategy, we are now able to waive GC-MS and in addition targeted LC-MS/MS methods for benzodiazepines, pregabalin, as well as synthetic opioides. From our experience with more than 9000 urine samples, the Toxtyper IT-MS is as robust as GC-MS screening. In addition, stepwise screening involving combination of the "DOAL" and the "MWW" library, achieved by injecting the same extracted sample twice, is a further advantage.

In conclusion, our study establishes that the Toxtyper IT-MS is suitable for screening the majority of substances encountered in our patient population. It opens the possibility for DOA screening at a high-quality level under routine conditions and will essentially replace the GC-MS methods in our combined algorithm, which includes immunoassay results. Although analytically less sensitive for some substances, we propose use of the "MWW" method as a first step because of its more comprehensive library. The "DOAL" method is valuable for targeted screening and when tentative results are generated with the "MWW" method. The Toxtyper is user-friendly, provides a reasonable solution for a 24 h/365 day emergency setting, and will not overburden the technical skills of laboratory technicians with minimal LC-MS experience. However, we believe it is mandatory that a dedicated well-trained staff be available during ordinary working hours for reagent preparation and instrument maintenance, as well as result interpretation. We are aware that neither our previous approach, nor the integration of Toxtyper low resolution mass spectrometry, will resolve the on-going challenge of detecting new substances with no hits in the libraries provided by the manufacturer; however, this issue could be alleviated if spectra acquired in different laboratories were shared, which is a feasible scenario given the tools provided by the manufacturer.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interest

None of the authors of the above manuscript has declared any conflict of interest within the last three years which may arise from being named as an author on the manuscript.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.clinms.2017.08.002.

References

[1] EMCDDA–Europol 2015 Annual Report on the implementation of Council Decision 2005/387/JHA, http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:35D0387:EN:HTML (accessed 06-14- 2017).
[2] M.R. Meyer, H.H. Maurer, LC coupled to low and high resolution mass spectrometry for new psychoactive substance screening in biological matrices. Where do we stand now, Anal. Chim. Acta. 927 (2016) 13–20.
[3] M.L. Snyder, C.R. Fantz, S. Melanson, Immunoassay-based drug tests are inadequately sensitive for medication compliance monitoring in patients treated for chronic pain, Pain Physician 20 (2S) (2017) SE1–SE9.
[4] J. Penders, A. Verstraete, Laboratory guidelines and standards in clinical and forensic toxicology, Accred. Qual. Assur. 11 (2006) 284–290.
[5] K.M. Rentsch, Knowing the unknown. State of the art of LCMS in toxicology. Trends, Anal Chem. 84 (2016) 88–93.
[6] G.M. Meyer, H.H. Maurer, M.R. Meyer, Multiple stage MS in analysis of plasma, serum, urine and in vitro samples relevant to clinical and forensic toxicology, Bioanalysis 8 (2016) 457–481.
[7] M.E. Nelson, S.M. Bryant, S.E. Aks, Emerging drugs of abuse, Emerg Med Clin North Am. 32 (1) (2014 Feb) 1–28, http://dx.doi.org/10.1016/j.emc.2013.09.001 Epub 2013 Oct 15.
[8] K. Aziz, Drugs-of-abuse testing Screening and confirmation, Clin Lab Med. 10 (3) (1990) 493–502.
[9] J. Kempf, J. Traber, V. Auwärter, L.M. Huppertz, 'Psychotropics caught in a trap' - adopting a screening approach to specific needs, Forensic Sci Int. 243 (2014 Oct) 84–89.
[10] D.K. Wissenbach, M.R. Meyer, D. Remane, A.A. Philipp, A.A. Weber, H.H. Maurer, Drugs of abuse screening in urine as part of a metabolite-based LC-MSn screening concept, Anal. Bioanal. Chem. 400 (10) (2011 Jul) 3481–3489, http://dx.doi.org/10.1007/s00216-011-5032-1.
[11] Bruker https://www.bruker.com/fileadmin/user_upload/8-PDF-Docs/Separations_MassSpectrometry/Literature/Brochures/1852451_Toxtyper_Brochure_05_2017_ebook.pdf (acessed 06-01-2017).
[12] Bruker Application Note # LCMS-91. https://www.bruker.com/fileadmin/user_upload/8-PDF-Docs/Separations_MassSpectrometry/Literature/literature/ApplicationNotes/1828629_AppNote_LCMS-91_Toxtyper_2014_ebook.pdf (acessed 06-01-2017).
[13] https://www.bruker.com/fileadmin/user_upload/8-PDF-Docs/Separations_MassSpectrometry/Literature/Poster/TIAFT_2014/TIAFT_2014_FAST_AND_RELIABLE_IDENTIFICATION_OF_DRUGS_pdf (accessed 06.01.2017).
[14] J. Vogelsang, J. Hädrich, Limits of detection, identification and determination: a statistical approach for practitioners, Accred Qual Assur. 3 (1988) 242–255.
[15] CLSI, Liquid Chromatography-Mass Spectrometry Methods; Approved Guideline. CLSI Document C62-A, Clinical and Laboratory Standards Institute, Wayne, PA, 2014.
[16] S. George, R.A. Braithwaite, A pilot study to determine the usefulness of the urinary excretion of methode and its primary metabolite (EDDP) as potential markers of compliance in methadone detoxification programs, J. Anal Toxicol. 23 (2) (1999) 81–85.
[17] S.E. Melanson, The utility of immunoassays for urine drug testing, Clin Lab Med. 32 (3) (2012) 429–447.
[18] K.E. Moeller, K.C. Lee, J.C. Kissack, Urine drug screening: practical guide for clinicians, Mayo Clin Proc. 83 (1) (2008) 66–76.
[19] A.H.B. Wu, D. French, Implementation of liquid chromatography/mass spectrometry into the clinical laboratory, Clin Chim Acta 420 (2013) 4–10.
[20] Mandatory Guidelines for Federal Workplace Drug Testing Programs. Federal Register /Vol. 80, No. 94 / Friday, May 15, 2015 /Notices).
[21] CLSI, User Protocol for Evaluation of Qualitative Test Performance; Approved Guideline - Second Edition. CLSI Document EP12-A2, Clinical and Laboratory Standards Institute, Wayne, PA, 2008.
[22] http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2011/08/WC500109686.pdf (accessed 06-14-2017).
[23] http://www.unodc.org/documents/scientific/validation_E.pdf (accessed 06-14-2017).
[24] C.H. Chau, O. Rixe1, H. McLeod, W.D. Figg, Validation of analytical methods for biomarkers employed in drug development, Clin. Cancer Res. 14 (19) (2008) 5967–5976.