Qualitative method validation and uncertainty evaluation via the binary output
II – Application to a multi-analyte LC-MS/MS method for oral fluid

Brigitte Desharnais[a,b], Marie-Jo Lajoie[a], Julie Laquerre[a], Pascal Mireault[a], Cameron D. Skinner[b]

[a] *Department of Toxicology, Laboratoire de sciences judiciaires et de médecine légale*
1701 Parthenais Street, Montréal, Québec, Canada H2K 3S7
[b] *Department of Chemistry & Biochemistry, Concordia University*
7141 Sherbrooke Street West, Montréal, Québec, Canada H4B 1R6

Abstract

A study of impaired driving rates in the province of Québec is currently planned following the legalization of recreational cannabis in Canada. Oral fluid (OF) samples are to be collected with a Quantisal® device and sent to the laboratory for analysis. In order to prepare for this project, a qualitative decision point analysis method monitoring for the presence of 97 drugs and metabolites in OF was validated according to the guidelines presented in the first part of this paper (I – Validation guidelines and statistical foundations).

This high throughput method uses incubation with a precipitation solvent (acetone:acetonitrile 30:70 v:v) to boost drug recovery from the collecting device and improve stability of benzodiazepines (e.g. α-hydroxyalprazolam, clonazepam, 7-aminoclonazepam, flunitrazepam, 7-aminoflunitrazepam, N-desmethylflunitrazepam, nitrazepam). The Quantisal ® device has polyglycol in its stabilizing buffer but timed use of the mass spectrometer waste valve proved sufficient to avoid the glycol interferences for nearly all analytes. Interferences from OF matrices and 140 potentially interfering compounds, carryover, ion ratios, stability, recovery, reproducibility, robustness, false positive rate, false negative rate, selectivity, sensitivity and reliability rates were tested in the validation process. Five of the targeted analytes (olanzapine, oxazepam, 7-aminoclonazepam, flunitrazepam and nitrazepam) did not meet the set validation criteria but will be monitored for identification purposes (no comparison to a cut-off level).

Blind internal proficiency teting was performed, where six OF samples were tested and analytes were classified as "negative", "likely positive" or "positive" with success. The final validated OF qualitative decision point method covers 92 analytes, and the presence of 5 additional analytes is screened in this high throughput analysis.

1. Introduction

Scientific papers on the use of oral fluid (OF) in forensic toxicology were published as early as 1965 [1, 2], but the use of this alternative matrix has become more widespread specifically since the 2000s [3, 4]. There is no doubt that the numerous advantages of OF play a role in its increased use. Indeed, OF collection is easy and minimally invasive [4, 5, 6, 7, 8, 9, 10, 11], allowing roadside collection and testing [5, 8, 9, 10, 11], and reduces the legal burden to obtain a biological sample. Moreover, the risk of sample adulteration is considered to be lower than with urine [4, 5, 7, 8, 10, 11], which can be diluted or modified with an adulterant [5, 7, 8]. In contrast to urine, from which only past use can be inferred, OF analysis will inform the forensic toxicologist about recent use [5, 7, 8, 9, 10, 11]. Indeed, OF has a similar detection window to blood [4, 8], but has the noteworthy advantage of being easily collected at the roadside, shortly after arrest. Three main disadvantages are recognized with regards to OF use. First, contamination of the oral cavity with substances in direct contact with the mouth is to be expected, for example THC from smoked cannabis [7, 12]. Second, low saliva production in certain drug users might complicate an otherwise easy collection process and result in artificially increased concentrations [10]. Finally, drug concentrations in OF aren't as well documented as they are in blood, complicating the interpretation process for the forensic toxicologist.

Nevertheless, on the whole, OF might be the matrix of choice to test for individuals driving under the influence of drugs. This alternative matrix represents a worthwhile compromise between ease of collection and toxicological relevance to impairment, which is often associated with recency of use [5, 6, 7, 8, 9, 10]. With this application in mind, several OF screening devices [5, 9, 11] and OF collection devices for laboratory confirmation [4, 5, 6, 7, 9, 10, 11] have been developed since 1990 [8, 13]. The selection of a particular device relies on several considerations including cost, effectiveness and intended use [6].

The recent legalization of recreational cannabis in Canada [14] has been accompanied by several modifications to the impaired driving legislation [15]. Of particular interest here is the introduction of OF drug testing using point-of-collection screening devices [16] as an investigation tool in driving under the influence of drugs (DUID) cases. These instruments must be approved for use in Canada by the General Attorney and be able to detect at least one of the following: THC at 25 ng/mL, cocaine at 50 ng/mL and/or methamphetamine at 50 ng/mL [17]. Canada is not the first country to use these devices in this particular context and is following in the footsteps of Australia, Belgium, Finland, France, Germany, Italy and specific states in the United States [5, 6, 7, 8].

In the midst of these legislative modifications, a provincial study here in Québec of impaired driving rates and the relevance of OF point-of-collection screening devices is planned. At the checkpoint, the drivers are to be tested with a device currently approved

*B. Desharnais, M.-J. Lajoie and J. Laquerre contributed equally to this work and are named in alphabetical order.

**Author to whom correspondence should be addressed. Email: brigitte.desharnais@msp.gouv.qc.ca

for use in Canada [16, 18]. An OF sample is also to be collected from all drivers passing through this checkpoint and sent to the laboratory for testing. The Quantisal® collection device from Immunalysis, which has already demonstrated its effectiveness for drugs of abuse and therapeutics analysis [4, 6, 10], was selected for this purpose. This type of study has been carried out in several other jurisdictions such as Italy, Norway and the United States (Wisconsin) [19, 20, 21, 22].

This paper presents the development and validation of a qualitative decision point method fit for the aforementioned study. The method needed to be high-throughput, since 2 500 OF samples would be received for analysis over 28 days. It covers an extensive set of analytes (97 in total), including cocaine, benzoylecgonine, amphetamines, benzodiazepines, cannabis (THC) and opioids as they are the most prevalent DUID findings in Québec [23]. Samples were prepared using dilution with an organic solvent, and analyzed by liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS).

A qualitative decision point method was developed rather than a quantitative method, for three main reasons. First, toxicological interpretation of OF drug concentration(s) is not yet well-established; second, only a short time frame could be allotted for method development and validation, and finally, it permits comparison with the point-of-collection screening device(s). Validation of this method was performed according to the guidelines presented in the first part of this paper (I – Validation guidelines and statistical foundations) and ISO 17025:2005 [24] requirements. Following method optimization, the absence of interferences and carryover was confirmed. False negative, false positive, selectivity, sensitivity and reliability rates were determined and confirmed to be reproducible and robust. Ion ratios, stability and recovery were also validated. The production-ready method permits classification of samples as being below cut-off, likely above cut-off or above cut-off by taking into account the uncertainty of measurement (as required by ISO 17025:2017) [25].

2. Materials and methods

2.1. Development and optimization

2.1.1. Polyglycol interferences

Polyglycols are known to be present in the Quantisal® stabilizing buffer [26] and can be deleterious to the mass spectrometer and accuracy of the results. The presence and behaviour of polyethyleneglycol (PEG) under the selected chromatographic conditions was investigated. A blank OF sample was collected with the Quantisal® device, extracted as described below (Section 2.2.2) and analyzed using the two chromatographic methods (Section 2.2.3) in full scan mode rather than target multiple reaction monitoring (MRM) mode. The resulting total ion chromatograms (TIC) were investigated for the presence of characteristic PEG profiles, i.e. an envelope of peaks spaced 44 Da apart (mass of an ethylene glycol unit) [27].

2.1.2. Maximizing recovery

Analyte recovery for the 97 compounds targeted (see Supplementary Data 1 for a complete list) was estimated via the area ratios (analyte to internal standard) for spiked OF. Samples taken up with the collector stick and put in the Quantisal® tube were compared to the same reference OF added directly to the stabilizing buffer in the Quantisal® tube. Recovery was calculated as:

$$\text{Recovery}(\%) = \frac{\text{Area ratio with collector}}{\text{Area ratio directly in stabilizing buffer}} \times 100 \quad (1)$$

Recoveries $> 80\%$ for any given analyte were considered acceptable.

Recoveries were evaluated for several experimental conditions: collector stick equilibration times of 24, 48, 72, 96 and 120 hours (with the collector stick in the stabilizing buffer); inclusion of a 10 minute sonication step; addition of the organic precipitation solvent directly to the Quantisal® tube still containing the collector stick; added solvent volumes from 0.5 mL to 8.5 mL and solvent incubation times from 17 to 127 hours.

2.2. Final analytical method

2.2.1. Preparation of control and threshold samples

Samples of OF were collected from voluntary laboratory employees and anonymized. Aliquots (1.5 mL) were then spiked at the required concentrations with the 97 targeted analytes (see Supplementary Data 1 for threshold (cut-off) concentrations). All compounds were purchased from Cerilliant (Round Rock, Texas, USA), except for N-desmethyl diphenhydramine, procyclidine and rolicyclidine which were obtained from Toronto Research Chemicals (North York, Ontario, Canada) and 3,4-methylenedioxypyrovalerone metabolite which was secured from Cayman Chemicals (Ann Harbor, Michigan, USA). The OF samples were spiked in borosilicate glass tubes 16 x 100 mm (Fisher Scientific, 14-961-29, Fair Lawn, New Jersey, USA). The Quantisal® collector pad was inserted into the glass tube, in contact with the OF, until blue coloration on the collector indicated saturation of the pad with OF ($1mL \pm 10\%$, according to manufacturer's documentation). At this point, the collector was transferred into the Quantisal® tube (Immunalysis, QS-0025, Pomana, California, USA) containing 3 mL of stabilizing buffer. Samples were stored as is at 4 °C overnight prior to sample extraction to simulate the expected sample shipment delay of future checkpoint studies.

2.2.2. Sample extraction

In the Quantisal® tube, 4.5 mL of acetone:acetonitrile (30:70 v:v) (acetone: HPLC Grade, Fisher Scientific, A949, Fair Lawn, New Jersey, USA; and acetonitrile: HPLC Grade, $\geq 99.9\%$, EMD Millipore corporation, AX0156-1, Billerica, Massachusetts, USA) organic solvent was added, without removing the collection device. Tubes were capped, mixed by inversion and incubated for 72 h at 4 °C. Following incubation and vortexing, 600 μL of extract was transferred to a 2 mL square well 96-well plate (Fisher Scientific, AB-0932, Ottawa, Ontario, Canada). Stable isotope internal standards (IS) solution

(10.0 μL) were added to each well; the compounds and concentration are detailed in Supplementary Data 1. Following mixing (1 minute at 1500 rpm on Thermomixer, Eppendorf, Mississauga, Ontario, Canada) and centrifugation (5 minutes at 3200 $\times g$), two different supernatant dilutions were prepared for analysis with the two chromatographic methods. The first chromatographic method (general) covered 96 of the 97 analytes targeted; for this purpose, 25.0 μL of supernatant was transferred to a different 96-well plate equipped with 1 mL round bottom wells (Canadian Life Science, RT96PPRWU1mL, Peterborough, Ontario, Canada) and diluted with 180 μL of 0.2% formic acid (Fisher Chemical, A117-50, Fair Lawn, New Jersey, USA). The second chromatographic method (cannabinoid) was designed specifically for cannabinoids analysis; for this purpose, 200 μL of supernatant was transferred to an identical, round bottom, 96-well plate and diluted with 50 μL of 1.5% formic acid.

2.2.3. LC-MS/MS analysis

For the general chromatography, 5 μL of the extract was separated in 13 minutes on an Agilent Zorbax Eclipse Plus C18 (2.1 X 100 mm, 3.5 μm) maintained at 50 °C. Mobile phase A was an aqueous solution of ammonium formate (pH 3.0):methanol (98:2). Mobile phase B was methanol (EMD Millipore corporation, MX0486-1, Billerica, MA, USA). A 650 $\mu L/min$ step/ramp from the A to the B mobile phase was used. Detailed gradient, analyte retention times, Q1/Q3 identification and confirmation transitions, source and mass spectrometer parameters are available in Supplementary Data 1.

For the cannabinoid chromatography, a 10 μL aliquot of the extract was separated using the above conditions, except the column was 50 mm long and a 550 $\mu L/min$ flow rate was used for the 6.5 minute separation. Analytical details are available in Supplementary Data 1.

In both cases, an Agilent HPLC 1200 or 1260 Infinity coupled to a Sciex 5500 QTrap mass spectrometer were used. The data acquisition software used was Analyst® 1.6.2 build 8489. Data was analyzed using Multiquant® 3.0.1 (Version 3.0.6256.0) software.

2.3. Validation procedures

Method validation was performed according to the principles established in the first part of this paper (I – Validation guidelines and statistical foundations), while considering SWGTOX's [28] recommendations and conforming to ISO 17025:2005 [24] and CAN-P-1578 [29] requirements.

2.3.1. Interferences

The presence of interferences from the matrix or exogenous compounds was assessed. To evaluate interferences from the matrix (specificity), 15 blank OF samples were analyzed with the final analytical method, including both chromatographies. Similarly, 140 potentially interfering compounds including caffeine, nicotine, cannabidiol, cannabinol, Δ8-THC, Exo-THC and cannabichromen were added to blank OF and analyzed using

both chromatographic methods; the full list of compounds tested and their concentrations are available in Supplementary Data 2. Interferences were considered to be present, from the matrix or potentially interfering compounds, if the analyte peak area ratio in the blank was greater than 25% of the average analyte peak area ratio of the cut-off sample; or if internal standard peak area in the blank was greater than 5% of the average internal standard area in cut-off samples.

2.3.2. Carryover

Three OF samples were prepared: one spiked with all compounds at a concentration 50 times higher than the cut-off; the second (analyte blank) was spiked with internal standards only and the third (full blank) was not spiked. These samples were extracted and analyzed in pairs (spiked followed by analyte blank or full blank) on three separate days with two replicates, for a total of 6 carryover tests. Carryover was considered to be present for an analyte when the analyte peak area ratio, in the analyte blank, was greater than 25% of the average area ratio of the cut-off samples; or for an internal standard when the internal standard peak area, in the full blank, was greater than 5% of the average internal standard area in cut-off samples.

2.3.3. Ion ratios

The ion ratio for each analyte was calculated as:

$$\text{Ion ratio} = \frac{\text{Area}_{\text{Transition 1}}}{\text{Area}_{\text{Transition 2}}} \times 100 \tag{2}$$

The ion ratio in all available samples spiked at 100% and 150% of the cut-off concentration (rates/reproducibility/robustness experiments, total of 60 samples from 15 different OF donors) were compared for each analyte to the average area ratio of 4 OF reference samples at 25 times the cut-off concentration. Ion ratios were expected to be within ±30% of the average ion ratio measured in the reference samples in $\geq$ 90% of cases.

2.3.4. Stability

Analyte stability at the cut-off concentration was evaluated both in the Quantisal® stabilizing buffer (i.e. the collector pad filled with OF in the Quantisal® tube, as the sample would be if it was shipped from the collection point to the laboratory) and in the organic solvent supplemented Quantisal® (i.e. after the addition of organic solvent for extraction purposes). In both cases, stability was evaluated at 4 °C over four weeks. Stability was calculated using the average of three OF samples as:

$$\text{Stability}(\%) = \frac{\overline{\text{Area ratio}}_{t=X} - \overline{\text{Area ratio}}_{t=0}}{\overline{\text{Area ratio}}_{t=0}} \tag{3}$$

with a target stability of 0% (a negative stability indicating degradation) and $t = 0$ wk being freshly spiked and extracted samples. To be considered stable at time $t = X$

wk, the analyte's stability should be $> -20\%$. However, in interpreting the stability results, the pattern of stability over the course of the study, not just one time point, was considered. A single out-of-criteria point at e.g. 2 weeks might be more reflective of a biased accuracy on that day when the next two time points showed stabilities closer to the 0% target.

2.3.5. Recovery

Recovery was evaluated according to the procedure described above using 5 OF samples spiked at the cut-off concentration.

2.3.6. False negative (FNR), false positive (FPR), selectivity (SLR), sensitivity (SNR) and reliability (RLR) rates

Method performance parameters were calculated on OF samples spiked with all analytes at 50% and 150% of the cut-off concentration. Results for the 15 samples at 50% of the cut-off were classified as true negative or false positive, whereas results for the 15 samples at 150% of the cut-off were classified as true positive or false negative. Results were used to calculate the performance parameters according to the equations detailed in the first part of this paper (I – Validation guidelines and statistical foundations). The acceptance criteria for the rates were as follows: $FNR \leq 7\%$, $FPR = 0\%$, $RLR \geq 93\%$, $SLR = 100\%$, $SNR \geq 93\%$.

2.3.7. Reproducibility and robustness

Reproducibility and robustness were evaluated by carrying out an evaluation of the rates as described in the previous paragraph on three different days, changing the solutions and mobile phases lots, HPLC columns, LC-MS/MS instruments and technical staff. The same acceptance criteria for the different rates apply.

2.3.8. Internal proficiency testing

In production, 5 samples of OF spiked at the cut-off concentration and 3 samples of OF spiked at 150% of the cut-off concentration were used to establish the classification bins. The average area ratio ($\overline{AR}_{CO}$) measured in the cut-off samples acted as the measurement threshold. Unknown samples with $AR_{UNK} < \overline{AR}_{CO}$ were classified as negative; samples with $\overline{AR}_{CO} \leq AR_{UNK} \leq \overline{AR}_{150\%}$ were classified as likely positive; and samples with $AR_{UNK} > \overline{AR}_{150\%}$ were classified as positive, as suggested in the first part of this paper (I – Validation guidelines and statistical foundations).

In order to test that the procedure performed as expected, six OF samples were fortified with different compounds at various concentrations, adsorbed onto the collector pad and stored in the Quantisal® tube. A laboratory member who was "blinded" to their concentrations analyzed these samples. The results (negative/likely positive/positive) were checked for accuracy by the laboratory member who performed sample spiking. This blind analysis was carried out in triplicate over two different batches.

3. Results and Discussion

3.1. Development and optimization

Because of the short time frame available for method development and validation, the most efficient path was to modify an existing LC-MS/MS targeted screening method [30] to suit the needs of this project (new matrix, qualitative analysis). In doing this, most of the method development concerns centered on the impact of using this new matrix: the OF in a Quantisal® stabilizing buffer. The Quantisal® device was selected based on literature reports and communications with other forensic toxicology laboratories. No other collection device was evaluated for this project, due to the short timeframe.

3.1.1. Polyglycol interferences

Although the exact contents of the Quantisal® stabilizing buffer are proprietary, the presence of polyglycol compounds such as polyethylene glycol (PEG) has been surmised in a few other studies [26, 31]. This family of compounds is known to be problematic for mass spectrometers by creating large ionic suppression and contaminating the quadrupoles [32]. It was crucial that this project, being secondary to the main laboratory production activity, did not jeopardize the performance of the mass spectrometry instrumentation. Presence of PEGs was, therefore, the first item examined in the course of method development. Full scan analyses of extracted blank OF in the Quantisal® stabilizing buffer were performed for both the general and cannabinoid chromatographic methods. Resulting total ion chromatograms (TICs) did not reveal the presence of PEGs in the general method; most likely, they eluted during the wash where the flow was diverted to waste. On the other hand, a typical polyethylene glycol pattern was identified in the cannabinoid method (Figure 1). The elution of PEGs in this separation coincided with the elution of 11-nor-9-carboxy-Δ9-tetrahydrocannabinol (THC-COOH) and 11-hydroxy-Δ9-tetrahydrocannabinol (THC-OH). The chromatographic conditions could have been modified to attempt to separate PEG interferences from THC-COOH and THC-OH; however, these metabolites have only a moderate importance to toxicological interpretation in OF [12]. Although the presence of THC-COOH in particular is relevant, it is present at very low levels (pg/mL) [33] in OF which, in any event, would be hard to detect with the present type of method. A separate, optimized method most likely using a pre-concentration step would need to be used [33]. Taking all of these elements into account, the final decision was to eliminate THC-COOH and THC-OH from the assay, which allowed the effluent to be diverted to waste during PEG's elution.

3.1.2. Maximizing recovery

Recovery from the whole collection device, including the collector pad and stabilizing buffer, should be estimated to be representative of analyzed samples [4, 5, 10]. Using this approach, it was quickly obvious that recovery from the device was an issue (Figure 2), with 55 out of 97 compounds achieving a recovery below the 80% threshold. Others have faced the same issue and increased the recovery by using a plunger to destroy the collection pad [5] or a serum separator to compress the collection pad [4]. The number of samples and time frame envisioned for the roadside study here called for a higher throughput solution, where individual collection pad crushing for each collected oral

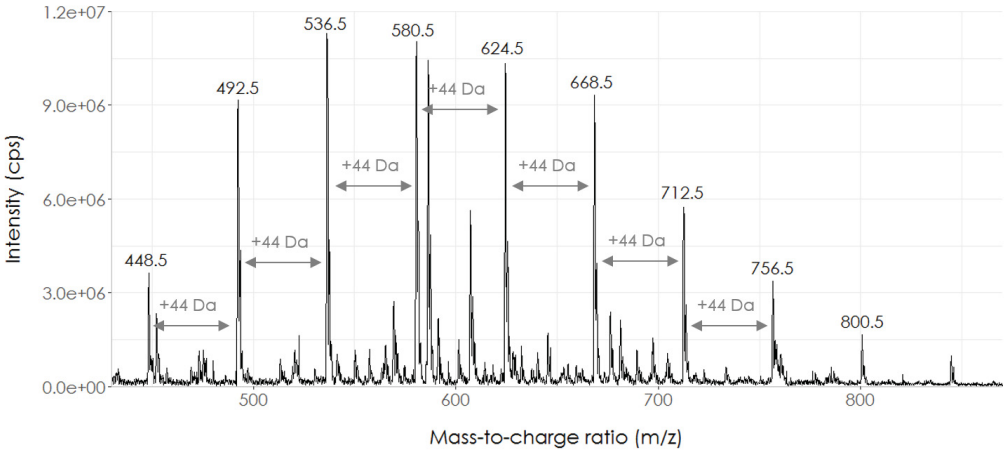

Figure 1: Polyethylene glycol (PEG) mass spectrum identified in the cannabinoids chromatography

fluid sample was not required. Several solutions were tested, including longer incubation times, addition of a sonication step and incubation in the organic solvent. Figure 2 shows analyte recovery density plots for a standard one day preparation, with a 10 minute sonication step, a three day incubation at 4 °C prior to sample preparation, and a three day incubation with the organic solvent. Peak density exceeds the 80% recovery threshold with an incubation period, and incubation with the organic solvent shows an additional recovery gain. Further optimization of the time and volume parameters for the organic solvent incubation was carried out to obtain the final extraction conditions.

3.2. Method validation

The main method validation results are summarized below, and complete results are available in Supplementary Data 2.

3.2.1. Interferences

Interferences from the OF matrix were observed for 7 analytes (aripiprazole, clobazam, cocaine, 7-amino-flunitrazepam, fluoxetine, lorazepam and N-desmethyl-mirtazapine), in 1 to 4 of the OF samples tested. In all cases however, the interferences were below the critical threshold (25% of the average area ratio in the cut-off samples). Interferences were also observed in all OF samples for THC's identification transition, but 0.05 minutes or more after THC's retention time and thus outside of the integration window. No significant interferences from the potentially interfering compounds tested were observed; all interferences observed were either below the set thresholds or chromatographically separated. Despite being below the set threshold and narrowly chromatographically separated, an interference of nitrazepam-D_5 on oxazepam's identification transition was noted. Thus overall, no significant interferences were noted for the analytes and internal standards under study.

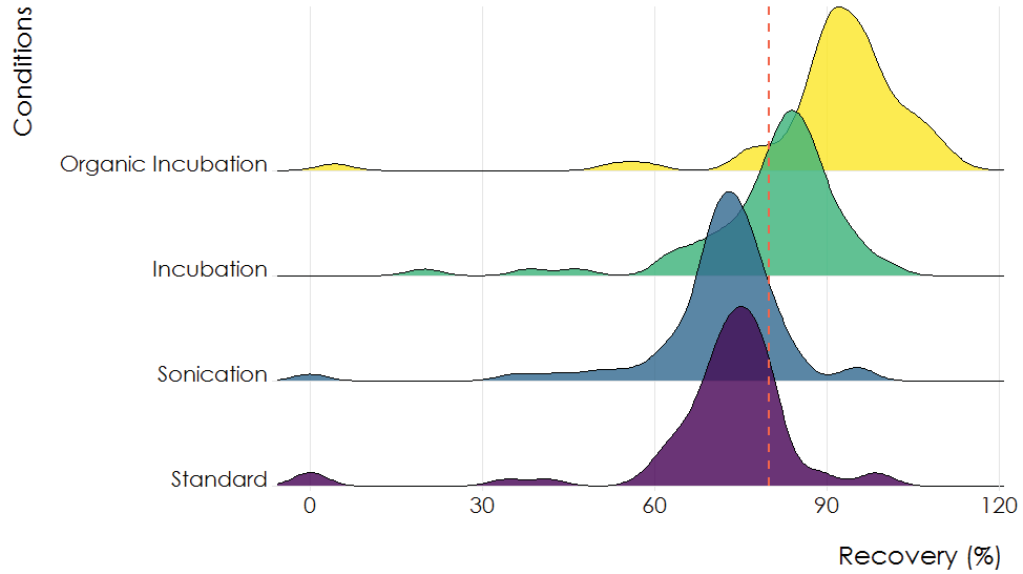

Figure 2: Density plots of the measured recoveries for all analytes under different preparation conditions

3.2.2. Carryover

Over the six blank samples analyzed, only one sample had analytes which did not satisfy the criteria for carryover from the high concentration sample (50 times the cut-off concentration). The analytes where carryover exceeded the threshold of 25% of the average area ratio in cut-off were: 7-amino-clonazepam (30%), benzoylecgonine (36%), 7-amino-flunitrazepam (33%) and O-desmethylvenlafaxine (35%). Given that the average carryover for each of these analytes was below the set threshold, none of the analytes were considered to have significant carryover. It is however noteworthy that 7-amino-flunitrazepam and oxycodone produce below threshold, but constant, carryover after injection of a high concentration sample. Samples immediately following those with a high concentration of these two analytes should be examined for potential carryover.

3.2.3. Ion ratios

The ion ratio of threshold and positive samples (100% and 150% of the cut-off concentration) meet the identification criteria for 99% of analyzed samples on average for all analytes except one: oxazepam. In this case, only 63% of the 60 samples analyzed meet the ion ratio criteria. This poor identification rate was attributable to the nitrazepam-D_5 interference on the identification transition, influencing the ion ratio for this analyte. Nitrazepam-D_5 was kept as an internal standard (IS) despite this interference because it is part of an IS mix used for other methods in the laboratory.

3.2.4. Stability

Stability of the OF samples, stored at 4°C, was evaluated over a period of 4 weeks. This time frame was longer than strictly necessary, since it is anticipated that samples will be fully analyzed within 3 weeks of collection.

Several species exhibit an instability pattern in the Quantisal® stabilizing buffer alone, starting after 1 to 3 weeks of storage: α-hydroxyalprazolam, chlordiazepoxide, clonazepam, 7-aminoclonazepam, flunitrazepam, 7-aminoflunitrazepam, N-desmethyl-flunitrazepam, nitrazepam, olanzapine, N-desmethylolanzapine, zopiclone and N-desmethylzopiclone.

The good news is, addition of the organic solvent addresses this instability problem for most analytes. After addition of the organic solvent, the only analytes exhibiting an instability pattern are chlordiazepoxide (starting at 3 weeks), olanzapine (3 weeks) and N-desmethyl-olanzapine (2 weeks), while 7-aminoflunitrazepam and N-desmethylflunitrazepam exhibited some small instability pattern which did not cross the -20% threshold. These results are not surprising, since olanzapine and its metabolite [34] and chlordiazepoxide [35] are known to suffer from instability issues in most biological matrices and extraction solvents. In all cases, treating samples with the organic solvent less than one week after collection, and analyzing them within the following week should suffice to avoid potential instability.

3.2.5. Recovery

The recovery experiment was performed for information purposes only, i.e. a recovery $< 80\%$ would still be considered acceptable if all other validation criteria were met. Indeed, if the rates experiments, characterizing the quality of the output of this method, meet expected criteria despite a low recovery, this indicates internal standards and matrix matched cut-off samples correct sufficiently for the bias introduced by the low recovery. Only four analytes were found to have a recovery below the 80% threshold, albeit close to it: flunitrazepam (73.9%), norfluoxetine (74.3%), nitrazepam (70.3%) and N-desmethylolanzapine (79.4%).

3.2.6. Rates, reproducibility and robustness

False negative (FNR), false positive (FPR), sensitivity (SNR), selectivity (SLR) and reliability (RLR) rates were obtained for 45 different OF samples divided across three experiments, each run on a separate day. In the first experiment, olanzapine was out-of-specification for FNR, SNR and SLR. In the second experiment, out-of-specification results were obtained for olanzapine (FNR, SNR) and oxazepam (FPR, RLR, SLR). In the third experiment, out-of-specification results were obtained for 7-aminoclonazepam (FNR, SNR), flunitrazepam (FNR, SNR) and nitrazepam (FNR, SNR). Combining results for the three batches, olanzapine, flunitrazepam and nitrazepam still did not meet the set criteria.

Olanzapine, 7-aminoclonazepam, flunitrazepam and nitrazepam all failed either the overall rates criteria or the reproducibility/robustness portion, in addition to having sta-

bility issues. A likely hypothesis is that the stability issues generate a higher rate of false negatives, in turn affecting the SNR. Oxazepam failed the reproducibility/robustness portion, in addition to the ion ratio specifications. The likely explanation here is that the interference from nitrazepam-D_5 contributed to increasing the number of false positives, in turn affecting the RLR and SLR.

Considering these results, olanzapine, 7-aminoclonazepam, flunitrazepam, nitrazepam and oxazepam were removed from the qualitative decision point method scope. These analytes are still monitored by LC-MS/MS, but the results are reported as "detected" or "not detected", i.e. positioning the sample with respect to a cut-off concentration cannot be reliably achieved. Rather, samples can only be screened. That being said, these drugs have low prevalence in the geographical region targeted by the roadside study (province of Québec, Canada), thus the impact of shifting those drugs from a qualitative decision point method to a screening only method is negligible.

3.2.7. Internal proficiency testing

Once the OF analysis method was validated for 92 analytes with a decision point and 5 analytes in screening mode, internal proficiency testing was performed. Spiked samples were treated as unknown samples (blind to the analyst) in exactly the same setting and conditions as in a production setting. 48 analytes were spiked at various levels around the cut-off concentrations in six different samples. Five extracted cut-off samples and three samples at 150% of the cut-off concentration (positive control) were analyzed as they would be in a production batch. Unknown samples, were classified as "negative" ($AR_{UNK} < \overline{AR}_{CO}$), "likely positive" ($\overline{AR}_{CO} \leq AR_{UNK} \leq \overline{AR}_{150\%}$) or "positive" ($AR_{UNK} > \overline{AR}_{150\%}$). The classification "likely negative" suggested in the first part of this paper (I – Validation guidelines and statistical foundation) was dropped due to its lower usefulness in a roadside survey study. Precautions against false positives should be taken, but false negatives in this context entails less consequences than false positive results. In samples analyzed (Supplementary Data 2), some below cut-off and > 150% cut-off samples did score as "likely positive", which was to be expected given the uncertainty of measurement and the concentrations used. More importantly, no false negative nor false positive results were detected, thus confirming the validity of the method.

4. Conclusions

We have successfully developed and validated a high throughput OF qualitative decision point analysis method for 92 analytes (and screening for 5 additional analytes). Amongst other possible applications, this atypically wide-scope method is suitable for OF roadside surveys where a large number of samples are collected and sent to the laboratory on the same day. Given the sample preparation and analysis time requirements, a total capacity for 350 to 400 samples per week is possible using one LC-MS/MS platform.

The extraction procedure was developed to maximize recovery from the collection device while remaining efficient in terms of laboratory labour and time required. Incubation with the precipitation organic solvent mixture achieves a high analyte recovery,

while simultaneously curbing the stability issue of several benzodiazepines.

While solid phase extraction (SPE) is typically used for the analysis of such samples to deal with the polyglycol content of the stabilizing buffer, we have found that a dilution approach can be suitably deployed, enabling higher throughput. A judicious chromatographic method can divert the polyglycols before they reach the mass spectrometer. In the present case, co-elution of polyethylene glycol with THC-COOH and THC-OH prevented their inclusion in the final method but, the more important THC was successfully analyzed.

The validation guidelines presented in the first part of this paper (I – Validation guidelines and statistical foundations) allowed for efficient validation in compliance with ISO 17025 [24] requirements. Only four experiments were required to complete the validation stage. As an added benefit, using this type of validation and production analysis produces information of additional value to the client, introducing a notion of measurement uncertainty in the final result reported (i.e. "likely positive" vs. "positive", in accordance with ISO 17025:2017 [25] requirements).

5. Acknowledgements

The authors wish to thank Cynthia Côté and Marc-André Morel for their work on the LC-MS/MS methods which inspired the one presented here and their thorough review of the manuscript, Marie-Pierre Taillon for her quality assurance work on this method, and Édith Viel for her review and comments. Brigitte Desharnais and Cameron D. Skinner gratefully acknowledge support of the National Sciences and Engineering Research Council of Canada. Brigitte Desharnais also gratefully acknowledges the support of the Fonds de recherche du Québec - Nature et technologies.

References

[1] Borzelleca, JF and Cherrick, HM, The excretion of drugs in saliva. Antibiotics., Journal of Oral Therapeutics and Pharmacology 2 (1965) 180.

[2] Horning, Marjorie G and Brown, Linda and Nowlin, Jean and Lertratanangkoon, Khing and Kellaway, Peter and Zion, Tomas E, Use of saliva in therapeutic drug monitoring, Clinical Chemistry 23 (1977) 157–164.

[3] Drummer, Olaf H, Drug testing in oral fluid, Clinical Biochemist Reviews 27 (2006) 147.

[4] Quintela, Oscar and Crouch, Dennis J and Andrenyak, David M, Recovery of drugs of abuse from the Immunalysis Quantisal™ oral fluid collection device, Journal of Analytical Toxicology 30 (2006) 614–616.

[5] Tang, Magdalene HY and Ching, CK and Poon, Simon and Chan, Suzanne SS and Ng, WY and Lam, M and Wong, CK and Pao, Ronnie and Lau, Angus and Mak, Tony WL, Evaluation of three rapid oral fluid test devices on the screening of multiple drugs of abuse including ketamine, Forensic Science International 286 (2018) 113–120.

[6] Gröschl, Michael, Saliva: a reliable sample matrix in bioanalytics, Bioanalysis 9 (2017) 655–668.

[7] Vindenes, V and Yttredal, B and Øiestad, EL and Waal, H and Bernard, JP and Mørland, JG and Christophersen, AS, Oral fluid is a viable alternative for monitoring drug abuse: detection of drugs in oral fluid by liquid chromatography-tandem mass spectrometry and comparison to the results from urine samples from patients treated with methadone or buprenorphine, Journal of Analytical Toxicology 35 (2011) 32–39.

[8] Doyon, Alexandra and Paradis-Tanguay, Laurence and Crispino, Frank and Lajeunesse, André, Les analyses médico-légales de salives: expertise vis-à-vis l'analyse des drogues, Canadian Society of Forensic Science Journal 50 (2017) 90–102.

[9] Wille, Sarah MR and Di Fazio, Vincent and Toennes, Stefan W and van Wel, Janelle HP and Ramaekers, Johannes G and Samyn, Nele, Evaluation of Δ9-tetrahydrocannabinol detection using DrugWipe5S® screening and oral fluid quantification after Quantisal™ collection for roadside drug detection via a controlled study with chronic cannabis users, Drug Testing and Analysis 7 (2015) 178–186.

[10] Cohier, Camille and Mégarbane, Bruno and Roussel, Olivier, Illicit drugs in oral fluid: Evaluation of two collection devices, Journal of Analytical Toxicology 41 (2017) 71–76.

[11] Krotulski, Alex J and Mohr, Amanda LA and Friscia, Melissa and Logan, Barry K, Field detection of drugs of abuse in oral fluid using the Alere™ DDS® 2 mobile test system with confirmation by liquid chromatography tandem mass spectrometry (LC–MS/MS), Journal of Analytical Toxicology 42 (2017) 170–176.

[12] S. M. Wille, V. Di Fazio, N. Samyn, La salive dans les investigations toxicologiques: considérations pratiques et analytiques, in: P. Kintz (Ed.), Traité De Toxicologie Médico-judiciaire, Elsevier Masson, Paris, 2012, pp. 219–255.

[13] Verstraete, Alain G, Oral fluid testing for driving under the influence of drugs: history, recent progress and remaining challenges, Forensic Science International 150 (2005) 143–150.

[14] Justice Canada, S.C. 2018, c. 16: Cannabis Act, Standard, https://laws-lois.justice.gc.ca/eng/acts/C-24.5/, Ottawa, Canada, 2018.

[15] Parliament of Canada, S.C. 2018, c. 21: Bill C-46 – An Act to amend the Criminal Code (offences relating to conveyances) and to make consequential amendments to other Acts, Standard, http://www.parl.ca/DocumentViewer/en/42-1/bill/C-46/royal-assent, Ottawa, Canada, 2018.

[16] Justice Canada, SOR/2018-179: Approved Drug Screening Equipment Order, Standard, https://laws-lois.justice.gc.ca/PDF/SOR-2018-179.pdf, Ottawa, Canada, 2018.

[17] Canadian Society of Forensic Science Drugs and Driving Committee, Drug Screening Equipment – Oral Fluid Standards and Evaluation Procedures, Standard, https://www.csfs.ca/wp-content/uploads/2017/11/Approval-Standards-for-Drug-Screening-Equipment.pdf, Ottawa, Canada, 2017.

[18] Justice Canada, Order Amending the Approved Drug Screening Equipment Order, Standard, http://www.gazette.gc.ca/rp-pr/p1/2019/2019-04-20/html/reg4-eng.html, Ottawa, Canada, 2019.

[19] Edwards, Lorrine D and Smith, Katherine L and Savage, Theodore, Drugged driving in Wisconsin: oral fluid versus blood, Journal of Analytical Toxicology 41 (2017) 523–529.

[20] Strano-Rossi, Sabina and Castrignanò, Erika and Anzillotti, Luca and Serpelloni, Giovanni and Mollica, Roberto and Tagliaro, Franco and Pascali, Jennifer P and Di Stefano, Delfina and Sgalla, Roberto and Chiarotti, Marcello, Evaluation of four oral fluid devices (DDS®, Drugtest 5000®, Drugwipe 5+® and RapidSTAT®) for on-site monitoring drugged driving in comparison with UHPLC–MS/MS analysis, Forensic Science International 221 (2012) 70–76.

[21] Veitenheimer, Allison M and Wagner, Jarrad R, Evaluation of Oral Fluid as a Specimen for DUID, Journal of Analytical Toxicology 41 (2017) 517–522.

[22] Furuhaugen, Håvard and Jamt, Ragnhild EG and Nilsson, Galina and Vindenes, Vigdis and Gjerde, Hallvard, Roadside survey of alcohol and drug use among Norwegian drivers in 2016–2017: A follow-up of the 2008–2009 survey, Traffic Injury Prevention 19 (2018) 555–562.

[23] Viel, Édith and Blais, Estelle and Mireault, Pascal, Statistical overview of drug findings in urine sample from the DRE program in the province of Québec, Canada, Standard, 2013 Society of Forensic Toxicologists Meeting (SOFT), Orlando, USA, 2013.

[24] ISO/IEC 17025:2005, General requirements for the competence of testing and calibration laboratories, Standard, International Organization for Standardization, Geneva, Switzerland, 2005.

[25] ISO/IEC 17025:2017, General requirements for the competence of testing and calibration laboratories, Standard, International Organization for Standardization, Geneva, Switzerland, 2017.

[26] D. Menasco, C. Summit, J. Neifeld, S. Marin, L. Williams, E. Gairloch, Practical Considerations using Quantisal Oral Fluid Collection Devices & SPE Method Development by Polymeric Mixed-Mode Cation Exchange, Standard, Annual Congress in Clinical Mass Spectrometry, Palm Springs, USA, 2018.

[27] R. Gudihal, S. Babu CV, N. Tang, S. Palaniswamy, U. S, S. Basingi, Analysis of Polyethylene Glycol (PEG) and a Mono and Di-PEGylated Therapeutic Protein Using HPLC and Q-TOF Mass Spectrometry, Standard, Agilent Technologies, 2012.

[28] Scientific Working Group for Forensic Toxicology, Scientific Working Group for Forensic Toxicology (SWGTOX) Standard Practices for Method Validation in Forensic Toxicology, Journal of Analytical Toxicology 37 (2013) 452–474.

[29] CAN-P-1578, Guidelines for the Accreditation of Forensic Testing Laboratories, Standard, Standards Council of Canada (Conseil canadien des normes), Ottawa, Canada, 2009.

[30] Côté, Cynthia and Desharnais, Brigitte and Morel, Marc-André and Laquerre, Julie and Taillon, Marie-Pierre and Daigneault, Gabrielle and Skinner, Cameron D and Mireault, Pascal, High Throughput Protein Precipitation: Screening and Quantification of 106 Drugs and their Metabolites using LC-MS/MS, Standard, 2017 Society of Forensic Toxicologists Meeting (SOFT) and 55th Annual Meeting of the International Association of Forensic Toxicologists (TIAFT), Boca Raton, USA, 2018.

[31] Fabresse, Nicolas and Aouad, Hassan and Knapp, Adeline and Mayer, Charlotte and Etting, Isabelle and Larabi, Islam Amine and Alvarez, Jean-Claude, Development and validation of a liquid chromatography-tandem mass spectrometry method for simultaneous detection of 10 illicit drugs in oral fluid collected with FLOQSwabs™ and application to real samples, Drug Testing and Analysis 11 (2019) 824–832.

[32] Leverence, Rachael and Avery, Michael J and Kavetskaia, Olga and Bi, Honggang and Hop, Cornelis ECA and Gusev, Arkady I, Signal suppression/enhancement in HPLC-ESI-MS/MS from concomitant medications, Biomedical Chromatography 21 (2007) 1143–1150.

[33] Lee, Dayong and Milman, Garry and Barnes, Allan J and Goodwin, Robert S and Hirvonen, Jussi and Huestis, Marilyn A, Oral fluid cannabinoids in chronic, daily cannabis smokers during sustained, monitored abstinence, Clinical Chemistry 57 (2011) 1127–1136.

[34] Saar, Eva and Gerostamoulos, Dimitri and Drummer, Olaf H and Beyer, Jochen, Assessment of the stability of 30 antipsychotic drugs in stored blood specimens, Forensic Science International 215 (2012) 152–158.

[35] Mata, Dani C, Stability of 26 sedative hypnotics in six toxicological matrices at different storage conditions, Journal of Analytical Toxicology 40 (2016) 663–668.