FORENSIC SCIENCE INTERNATIONAL POST-PRINT

A threshold LC–MS/MS method for 92 analytes in oral fluid collected with the Quantisal® device

Brigitte Desharnais[*,§,1,2] | Marie-Jo Lajoie[§,1] | Julie Laquerre[§,1] | Pascal Mireault[1] | Cameron D. Skinner[2]

[1]Department of Toxicology, Laboratoire de sciences judiciaires et de médecine légale, 1701 Parthenais Street, Montréal, Québec, H2K 3S7, Canada

[2]Department of Chemistry & Biochemistry, Concordia University, 7141 Sherbrooke Street West, Montréal, Québec, H4B 1R6, Canada

Correspondence

[*]Brigitte Desharnais, 1701 Parthenais Street, Montréal, Québec, H2K 3S7, Canada.
Email: brigitte.desharnais@msp.gouv.qc.ca

[§]B. Desharnais, M.-J. Lajoie and J. Laquerre contributed equally to this study and are named in alphabetical order.

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 developed and validated.

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 testing 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.

KEYWORDS:

qualitative decision point methods, oral fluid, saliva, Quantisal, liquid chromatography tandem mass spectrometry, roadside survey

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]. In Canada, the lower invasiveness of OF collection reduces the legal burden to obtain this biological sample: OF drug screening results, following observations by the police officer (e.g., erratic driving) can actually be grounds for the more invasive blood collection[12]. 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. A few important considerations should be kept in mind with regards to OF use. First, drug concentrations in OF are not as well documented as they are in blood, complicating the interpretation process. Moreover, correlation of OF concentration to blood concentration can be hindered by contamination of the oral cavity with substances in direct contact with the mouth, for example THC from smoked cannabis[7,13]. Furthermore, low saliva production in certain drug users might complicate an otherwise easy collection process and result in artificially decreased concentrations due to incomplete volume collected[10]. In any event, impairment evaluation should not be based solely on blood or OF concentration, but should include, for example, observations by the arresting officer.

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[4,5,6,7,7,10,11] have been developed since 1990[8,14]. 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[15] has been accompanied by several modifications to the impaired driving legislation[16]. Of particular interest here is the introduction of OF drug testing using point-of-collection screening devices[17] 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[18]. 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[5,8].

In the midst of these legislative modifications, a provincial study in Québec (Canada) 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 for use in Canada[17,19]. 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)[20,21,22,23].

This paper presents the development and validation of a qualitative decision point method fit for the aforementioned study. A high-throughput method was required, 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[24]. Samples were prepared using dilution with an organic solvent, and analyzed by liquid chromatography coupled to tandem mass spectrometry (LC–MS/MS).

Targeted qualitative methods, or "qualitative confirmation/identification" methods as per the AAFS Standards Board[25], can be classified as identification methods, yielding a "present" or "not present" result, or decision point (threshold) methods, yielding a "above cut-off" or "below cut-off" outcome. In the current case, a qualitative decision point method was developed rather than a quantitative or qualitative identification 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 Camirand Lemyre et al.[26] and ISO 17025:2005[27] 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 allows classification of samples as being below cut-off, likely above cut-off or above cut-off. This classification takes into account the uncertainty of measurement, which is required by several accreditation organizations for transition under the new version of the norm, ISO 17025:2017[28]. Although uncertainty of measurement is typically associated with quantitative methods, qualitative methods results are also affected by it. Classifying samples as below/likely below/likely above/above cut-off (or a variation on this theme), allows the laboratory to take into account uncertainty of measurement and clearly indicate which samples fall in an unreliability zone where repeated measurement would not always yield the same result (sometimes falling above, sometimes falling below cut-off). The interested reader will find further details on uncertainty of measurement in the context of qualitative methods in Camirand Lemyre et al.[26].

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[29] 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 (see Section 2.2.3 below) in full scan mode rather than targeted multiple reaction monitoring (MRM) mode. The resulting total ion chromatograms (TICs) 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)[30].

2.1.2 | Maximizing recovery from the oral fluid collection device

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:

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

Recoveries > 80% are preferred for any given analyte. This threshold generally allows for a method sensitive enough that different options (e.g., dilution) can be used in the course of method development. Nonetheless, lower recoveries are acceptble if the other validation criteria are met. Additionally, recovery variability < 20% is preferred (%RSD over $n = 3$ replicates).

Recoveries and recovery variability were evaluated for several experimental conditions: collector stick equilibration times of 24, 48, 72, 96 and 120 h (with the collector stick in the stabilizing buffer); inclusion of a 10 min 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 h.

2.2 | Final analytical method

2.2.1 | Preparation of control and threshold samples

Samples of OF were collected from voluntary laboratory employees. A 50 mL polypropylene conical tube was used to collect the expectorant sample over the course of 4 h, with a minimal volume of 5 mL being requested. No period without food or drink was requested, and samples were 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 × 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 (1 mL ± 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, ≥ 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) was added to each well; the compounds and concentrations are detailed in Supplementary Data 1. Note that the IS solution used was repurposed from another in-house method; other laboratories could choose to use lower IS concentrations matching more closely the cut-off concentration. Following mixing (1 min at 1500 rpm on Thermomixer, Eppendorf, Mississauga, Ontario, Canada) and centrifugation (5 min at 3200 × 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) in water. 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 in water.

2.2.3 | LC-MS/MS analysis

For the general chromatography, 5 µl of the extract was separated in 13 min on an Agilent Zorbax Eclipse Plus C18 (2.1 × 100 mm, 3.5 µm) maintained at 50 °C. Mobile phase A was an aqueous solution of ammonium formate/formic acid (pH 3.0):methanol (98:2). Mobile phase B was methanol (EMD Millipore corporation, MX0486-1, Billerica, MA, USA). A 650 µl 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 µl/min flow rate was used for the 6.5 min 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 Analyst1 1.6.2 build 8489. Data was analyzed using Multiquant1 3.0.1 (Version 3.0.6256.0) software.

2.3 | Validation procedures

Method validation was performed according to the principles established in Camirand Lemyre et al.[26], while considering the AAFS Standards Board's[25] recommendations and conforming to ISO 17025:2005[27] and the Standards Council of Canada[31] requirements.

2.3.1 | Decision point concentrations

Decision point concentrations, detailed in Supplementary Data 1, were chosen to fit OF screening devices thresholds set by Canadian legislation[16] and, as far as possible, recommendations by Logan et al.[32] The analytical limit of detection was not evaluated and rather set to the decision point concentrations[25]; however, for all analytes a lower limit of detection would be achievable analytically.

2.3.2 | 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 20% of the average analyte peak area ratio of the cut-off sample[33]; or if internal standard peak area in the blank was greater than 5% of the average internal standard area in cut-off samples.

2.3.3 | 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 20% of the average area ratio of the cut-off samples[33]; 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.

2.3.4 | 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 $\pm$30% of the average ion ratio measured in the reference samples in $\geq$ 90% of cases. This ion ratio tolerance is suggested by the European Commission Directorate General for Health and Food Safety[34]; other guidelines, using a tolerance varying as a function of the ion ratio level, could also be used.

2.3.5 | 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 Quantisal® supplemented with organic solvent (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:

$$\text{Stability}(\%) = \frac{\overline{\text{Area ratio}}_{t=X} - \overline{\text{Area ratio}}_{t=0}}{\overline{\text{Area ratio}}_{t=0}} \times 100 \tag{3}$$

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.

2.3.6 | Recovery

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

2.3.7 | 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. These method performance estimators (FNR, FPR, SLR, SNR, RLR) take into account all processes that are integral to the method, such as recovery variability, ion suppression and matrix effects. 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 expected performance parameters according to the equations as per Camirand Lemyre et al.[26]. Relying on these results, the acceptance criteria for the rates were as follows: FNR < 7%, FPR = 0%, RLR > 93%, SLR = 100%, SNR > 93%[26]. A method satisfying these acceptance criteria will have adequate performance at the production stage.

2.3.8 | 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.

2.3.9 | 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[26].

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 identity and 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[35] 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 time frame.

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[29,36]. This family of compounds is known to be problematic for mass spectrometers by creating large ionic suppression and contaminating the quadrupoles[37]. 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- $\Delta 9$-tetrahydrocannabinol (THC-COOH) and 11-hydroxy- Δ-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[13]. Although the presence of THC-COOH in particular is relevant, it is present at very low levels (pg/mL)[38] 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[38]. 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.

FIGURE 1 Polyethylene glycol (PEG) mass spectrum identified in the cannabinoids chromatography, at 3.740 min.
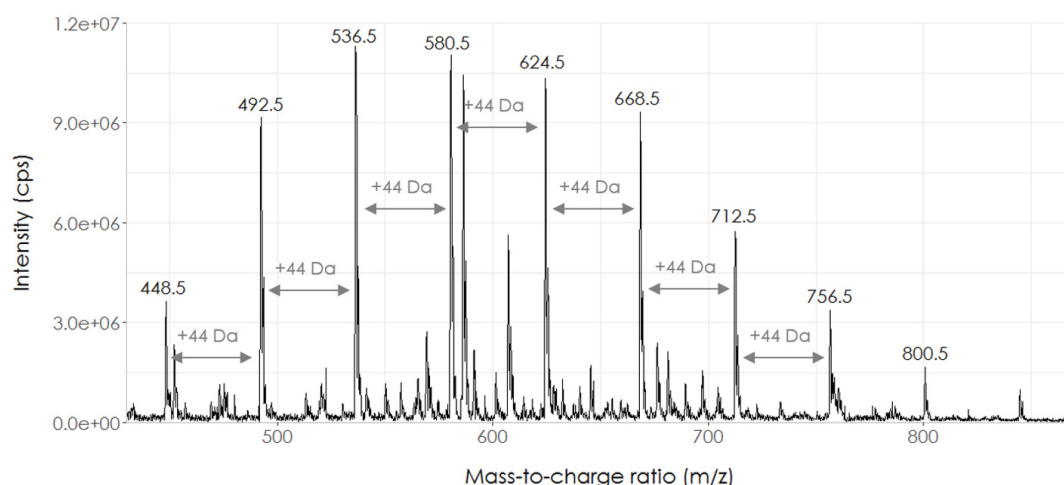
3.1.2 | Maximizing recovery from the oral fluid collection device

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 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 min sonication step, a three day incubation at 4 °C prior to sample preparation, and a three day incubation with the organic solvent (also at 4 °C). Peak density exceeds the 80% recovery threshold with an incubation period, and incubation with the organic solvent shows an additional recovery gain. Recovery variability was similar in all conditions, with the overwhelming majority of analytes (> 94%) satisfying the < 20% criterion. 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, with complete results 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 1–4 of the OF samples tested. In all cases however, the interferences were

FIGURE 2 Density plots of the measured recoveries for all analytes under different preparation conditions, including the addition of a 10 min sonication step, a three day incubation at 4 °C, and a three day incubation with organic solvent at 4 °C.
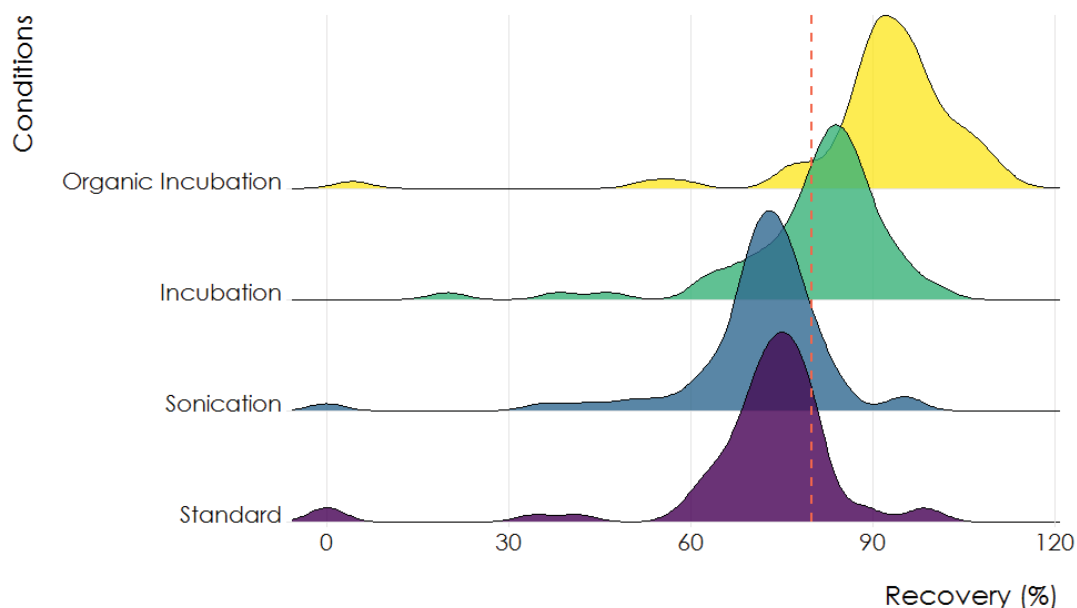
below the critical threshold (20% of the average area ratio in the cut-off samples). Interferences were also observed in all OF samples for THC's identification transition, but 0.05 min or more after THC's retention time and thus outside of the integration window. No significant interferences from the 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.

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 20% 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. Furthermore, certain analytes in oral fluid samples have been reported to reach concentrations in excess of thousands and even tens of thousands of ng/mL[39]. These are much more elevated than the concentrations tested here, and absence of carry-over at these levels cannot be considered to be demonstrated. Samples immediately following those with a high concentration (> 50 times the cut-off) of a given analyte 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. As a consequence of these results and the reproducibility/robustness results (see below), oxazepam was removed from the qualitative decision point method scope.

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. Again, this method was intended for analysis of oral fluid samples collected in a roadside survey. If results from this method alone are to be used to bring accusations against a suspect, stability over a longer time frame, and under long-term storage conditions (including freeze/thaw), need to be validated to e.g. accommodate potential counter expertise procedures.

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

This instability problem is solved for most analytes by the addition of the organic solvent. 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[40] and chlordiazepoxide[41] 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.

This stability evaluation was performed using spiked samples. Such a procedure covers the impact of several factors, including the collector itself as well as a range of OF characteristics (pH, composition, etc.). Of course, this does not preclude that a different stability profile would be observed in samples from drug users (authentic samples). Spiking matrices is standard in stability studies, with OF[42,43,44] as well as in other matrices. This is apractical approach especially considering the large spectrum of xenobiotics covered in the method presented here; evaluating the stability for each one using authentic samples would be close to impossible. Nonetheless, once the collection project is started, it would be interesting to evaluate stability for the most prevalent compounds in authentic OF samples.

3.2.5 | Recovery from the oral fluid collection device

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 robstness

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 (i.e., did not satisfy the acceptance criteria) 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 stability 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, identification alone can be performed, and for these analytes, this method is a qualitative identification one. 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 an identification 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. Forty-eight 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 by Camirand Lemyre et al.[26] 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. For additional details on the expected behaviour of results in qualitative decision point methods, the reader is referred to Camirand Lemyre et al.[26]. 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 identification 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 350–400 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 Camirand Lemyre et al.[26] allowed for efficient validation in compliance with ISO 17025[27] requirements. Only four experiments were required to complete the validation stage. 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[26] (i.e. "likely positive" vs. "positive", in accordance with ISO 17025:2017[28] requirements). In this reporting framework, samples falling within the unreliability zone, i.e., samples which would not systematically yield an above or below cut-off result upon re-analysis, are identified as such clearly to the scientists and clients.

5 | FUNDING

6 | CREDIT AUTHORSHIP CONTRIBUTION STATEMENT

Brigitte Desharnais: Conceptualization, Formal analysis, Investigation, Writing - original draft, Visualization. **Marie-Jo Lajoie**: Conceptualization, Methodology, Validation, Data curation, Writing - original draft, Project administration. **Julie Laquerre**: Conceptualization, Methodology, Validation, Data curation, Writing - original draft, Project administration. **Pascal Mireault**: Conceptualization, Writing - review & editing, Supervision. **Cameron D. Skinner**: Writing - review & editing.

7 | DECLARATION OF COMPETING INTERESTS

The authors report no declarations of interest.

8 | ACKNOWLEDGEMENTS

References

1. Borzelleca J, Cherrick H. The excretion of drugs in saliva. Antibiotics.. *Journal of Oral Therapeutics and Pharmacology* 1965; 2(3): 180.

2. Horning MG, Brown L, Nowlin J, Lertratanangkoon K, Kellaway P, Zion TE. Use of saliva in therapeutic drug monitoring. *Clinical Chemistry* 1977; 23(2): 157–164.

3. Drummer OH. Drug testing in oral fluid. *Clinical Biochemist Reviews* 2006; 27(3): 147.

4. Quintela O, Crouch DJ, Andrenyak DM. Recovery of drugs of abuse from the Immunalysis Quantisal™ oral fluid collection device. *Journal of Analytical Toxicology* 2006; 30(8): 614–616. doi: 10.1093/jat/30.8.614

5. Tang MH, Ching C, Poon S, et al. Evaluation of three rapid oral fluid test devices on the screening of multiple drugs of abuse including ketamine. *Forensic Science International* 2018; 286: 113–120. doi: 10.1016/j.forsciint.2018.03.004

6. Gröschl M. Saliva: a reliable sample matrix in bioanalytics. *Bioanalysis* 2017; 9(8): 655–668. doi: 10.4155/bio-2017-0010

7. Vindenes V, Yttredal B, Øiestad E, et al. 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* 2011; 35(1): 32–39. doi: 10.1093/anatox/35.1.32

8. Doyon A, Paradis-Tanguay L, Crispino F, Lajeunesse A. Les analyses médico-légales de salives: expertise vis-à-vis l'analyse des drogues. *Canadian Society of Forensic Science Journal* 2017; 50(2): 90–102. doi: 10.1080/00085030.2017.1303254

9. Wille SM, Di Fazio V, Toennes SW, Wel vJH, Ramaekers JG, Samyn N. 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* 2015; 7(3): 178–186. doi: 10.1002/dta.1660

10. Cohier C, Mégarbane B, Roussel O. Illicit drugs in oral fluid: Evaluation of two collection devices. *Journal of Analytical Toxicology* 2017; 41(1): 71–76. doi: 10.1093/jat/bkw100

11. Krotulski AJ, Mohr AL, Friscia M, Logan BK. 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* 2017; 42(3): 170–176. doi: 10.1093/jat/bkx105

12. Viel E. The year Canada became green - an overview of the legislation surrounding cannabis legalization in Canada. *TIAFT Bulletin* 2018; 48(4): 21–25.

13. Wille SM, Di Fazio V, Samyn N. La salive dans les investigations toxicologiques: considérations pratiques et analytiques. In: Kintz P., ed. *Traité De Toxicologie Médico-judiciaire*Paris, France: Elsevier Masson. 2012 (pp. 219–255).

14. Verstraete AG. Oral fluid testing for driving under the influence of drugs: history, recent progress and remaining challenges. *Forensic Science International* 2005; 150(2-3): 143–150. doi: 10.1016/j.forsciint.2004.11.023

15. Cannabis Act. Legislation S.C. 2018, c. 16, Government of Canada; Ottawa, Canada: 2018.

16. An Act to amend the Criminal Code (offences relating to conveyances) and to make consequential amendments to other Acts. Legislation S.C. 2018, c. 21, Bill C-46, Government of Canada; Ottawa, Canada: 2018.

17. Approved Drug Screening Equipment Order. Legislation SOR/2018-179, Government of Canada; Ottawa, Canada: 2018.

18. Drug Screening Equipment – Oral Fluid Standards and Evaluation Procedures. standard, Canadian Society of Forensic Science Drugs and Driving Committee; Ottawa, Canada: 2017.

19. Order Amending the Approved Drug Screening Equipment Order. legislation, Government of Canada; Ottawa, Canada: 2019.

20. Edwards LD, Smith KL, Savage T. Drugged driving in Wisconsin: oral fluid versus blood. *Journal of Analytical Toxicology* 2017; 41(6): 523–529. doi: 10.1093/jat/bkx051

21. Strano-Rossi S, Castrignanò E, Anzillotti L, et al. 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* 2012; 221(1-3): 70–76. doi: 10.1016/j.forsciint.2012.04.003

22. Veitenheimer AM, Wagner JR. Evaluation of Oral Fluid as a Specimen for DUID. *Journal of Analytical Toxicology* 2017; 41(6): 517–522. doi: 10.1093/jat/bkx036

23. Furuhaugen H, Jamt RE, Nilsson G, Vindenes V, Gjerde H. Roadside survey of alcohol and drug use among Norwegian drivers in 2016–2017: A follow-up of the 2008–2009 survey. *Traffic Injury Prevention* 2018; 19(6): 555–562. doi: 10.1080/15389588.2018.1478087

24. Viel E, Blais E, Mireault P. Statistical Overview of Drug Findings in Urine Samples from the DRE Program in the Province of Québec, Canada. poster presentation, 2013 Society of Forensic Toxicologists Meeting (SOFT); Orlando, United States: 2013.

25. Standard Practices for Method Validation in Forensic Toxicology (Draft). standard, American Academy of Forensic Sciences Standards Board; Colorado Springs, United States: 2018.

26. Camirand Lemyre F, Desharnais B, Laquerre J, et al. Qualitative threshold method validation and uncertainty evaluation: A theoretical framework and application to a 40 analytes liquid chromatography–tandem mass spectrometry method. *Drug Testing and Analysis* 2020; 12(9): 1287–1297. doi: 10.1002/dta.2867

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

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

29. Menasco D, Summit C, Neifeld J, Marin S, Williams L, Gairloch E. Practical Considerations using Quantisal Oral Fluid Collection Devices & SPE Method Development by Polymeric Mixed-Mode Cation Exchange. poster presentation, Annual Congress in Clinical Mass Spectrometry; Palm Springs, United States: 2018.

30. Gudihal R, Babu CV S, Tang N, Palaniswamy S, S U, Basingi S. Analysis of Polyethylene Glycol (PEG) and a Mono and Di-PEGylated Therapeutic Protein Using HPLC and Q-TOF Mass Spectrometry. Application Note 5991-1509EN, Agilent Technologies; United States: 2012.

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

32. Logan BK, D'Orazio AL, Mohr AL, et al. Recommendations for toxicological investigation of drug-impaired driving and motor vehicle fatalities – 2017 update. *Journal of Analytical Toxicology* 2018; 42(2): 63–68. doi: 10.1093/jat/bkx082

33. European Medicines Agency . Guideline on bioanalytical method validation. standard, https://www.ema.europa.eu/documents/scientific-guideline/guideline-bioanalytical-method-validation_en.pdf; London, United Kingdom: 2011.

34. European Commission Directorate General for Health and Food Safety . Guidance document on analytical quality control and method validation procedures for pesticide residues and analysis in food and feed. standard, https://ec.europa.eu/food/sites/food/files/plant/docs/pesticides_mrl_guidelines_wrkdoc_2017-11813.pdf; : 2017.

35. 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.

36. Fabresse N, Aouad H, Knapp A, et al. 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* 2019; 11(6): 824–832. doi: 10.1002/dta.2563

37. Leverence R, Avery MJ, Kavetskaia O, Bi H, Hop CE, Gusev AI. Signal suppression/enhancement in HPLC-ESI-MS/MS from concomitant medications. *Biomedical Chromatography* 2007; 21(11): 1143–1150. doi: 10.1002/bmc.863

38. Lee D, Milman G, Barnes AJ, Goodwin RS, Hirvonen J, Huestis MA. Oral fluid cannabinoids in chronic, daily cannabis smokers during sustained, monitored abstinence. *Clinical Chemistry* 2011; 57(8): 1127–1136. doi: 10.1373/clinchem.2011.164822

39. Drummer OH, Gerostamoulos D, Chu M, Swann P, Boorman M, Cairns I. Drugs in oral fluid in randomly selected drivers. *Forensic Science International* 2007; 170(2-3): 105–110. doi: 10.1016/j.forsciint.2007.03.028

40. Saar E, Gerostamoulos D, Drummer OH, Beyer J. Assessment of the stability of 30 antipsychotic drugs in stored blood specimens. *Forensic Science International* 2012; 215(1-3): 152–158. doi: 10.1016/j.forsciint.2011.02.022

41. Mata DC. Stability of 26 sedative hypnotics in six toxicological matrices at different storage conditions. *Journal of Analytical Toxicology* 2016; 40(8): 663–668. doi: 10.1093/jat/bkw084

42. Crouch DJ. Oral fluid collection: the neglected variable in oral fluid testing. *Forensic Science International* 2005; 150(2-3): 165–173. doi: 10.1016/j.forsciint.2005.02.028

43. Ventura M, Pichini S, Ventura R, et al. Stability of drugs of abuse in oral fluid collection devices with purpose of external quality assessment schemes. *Therapeutic Drug Monitoring* 2009; 31(2): 277–280. doi: 10.1097/FTD.0b013e318198670b

44. Ventura M, Pichini S, Ventura R, Zuccaro P, Pacifici R, Torre d. lR. Stability studies of principal illicit drugs in oral fluid: preparation of reference materials for external quality assessment schemes. *Therapeutic Drug Monitoring* 2007; 29(5): 662–665. doi: 10.1097/FTD.0b013e318156e892