Validation of rapid oral fluid test (ROFT) devices for on-spot screening of drug users

驗證快速口腔液測試工具供即場識別吸毒者

Project reference no.: BDF140027

Final Report

Author: Hospital Authority Toxicology Reference Laboratory

Collaborators: Department of General Adult Psychiatry (Castle Peak Hospital), Department of Psychiatry (Pamela Youde Nethersole Eastern Hospital), Substance Abuse Assessment Unit (Kwai Chung Hospital), the Society of Rehabilitation and Crime Prevention (SRACP)

Date: 5 May 2017

Table of Contents

1. Introduction

Oral fluid is becoming a popular matrix for rapid screening of drugs of abuse. In contrast to blood and urine, collection of oral fluid is easy and non-invasive with minimal intrusion into personal privacy. Oral fluid can also be collected under direct observation, thus eliminating the possibility of sample substitution or adulteration as with urine. As such, oral fluid can be useful in various settings that require drug testing, for example workplace, corrections, probation or for treatment. Importantly, it is by far the most convenient biological matrix that facilitates roadside testing for driving under the influence of drugs (drugged driving) [1]. Compared with urine, oral fluid is a better reflection of blood concentrations of a drug. It indicates recent drug use and provides better correlation with pharmacological effects such as impaired driving performance [2].

Drugged driving is a major concern worldwide. In the large-scale European Union (EU) study, Driving under the Influence of Drugs, Alcohol and Medicines (DRUID), it has been reported that the detection rate of illicit drugs in the general driving population was 1.9%. This detection rate was higher in seriously injured drivers (2.3%-12.6%) [3]. In Hong Kong, a study on the prevalence of illicit drug use in non-fatal traffic accident casualties showed that 10% of the injured drivers tested positive for drugs. Ketamine was the most commonly detected substance found in 45% of the subjects [4].

Currently, many countries including Germany, France, Belgium, Italy, Finland and Australia routinely conduct roadside rapid oral fluid testing (ROFT) to tackle drugged driving. In Hong Kong, similar to other countries, driving with any measurable amount of the specified illicit substance in the biological matrix constitutes an offence, i.e. so-called “zero tolerance” limit [5, 6].

Prior to usage, ROFT devices must undergo rigorous scientific evaluation to ensure acceptable performance in terms of their sensitivity, specificity and overall accuracy. In the early EU studies, ROSITA-1 and -2, the proposed acceptance criteria of sensitivity and specificity were >90% and accuracy >95% [7, 8]. These criteria were later lowered to 80% in the subsequent DRUID study [9]. During the past two decades, ROFT devices have been extensively evaluated and the results widely

published [10-15]. However, whilst the performance of ROFT devices for detecting amphetamines, opiates, cocaine and cannabis (THC) has been comprehensively investigated, there is currently minimal data for ketamine.

Although the abuse of ketamine is widespread in Hong Kong and Asia, it has not traditionally been a popular drug of abuse in Europe and North America [16]. As a result, detailed investigations of ROFT device performance on screening for ketamine have been scarce thus far. One study evaluated the performance of OratectXP solely on the detection of ketamine [14]. On the other hand, recent publications have reported an increase in the use of ketamine in Europe [10, 17]. More importantly in the local setting, the Road Traffic Ordinance in Hong Kong includes ketamine as one of the specified illicit drugs (in addition to heroin, methamphetamine, cannabis, cocaine and MDMA) [6]. In view of this, the current study was conducted to evaluate ROFT devices suitable for screening all of the above six illicit substances simultaneously. Three ROFT devices (DrugWipe® 6S, Ora-Check® and SalivaScreen®) were chosen for evaluation of their sensitivity, specificity and accuracy in detecting ketamine, opiates, methamphetamine, cannabis, cocaine and MDMA. Prior to conducting the ROFT field test, a liquid chromatography tandem mass spectrometry (LCMS) assay had to be established for confirmation analysis, the results of which will be used to assess the performance of the ROFT devices.

2. Methods

2.1 Materials

Reference standards of 6-monoacetylmorphine (6-MAM), amphetamine (AMP), ketamine (KET), methamphetamine (MET), morphine (MOR), norketamine (NORKET) and cannabis (THC) were obtained from Cerilliant (Round Rock, TX); benzoylecgonine (BEG), cocaine (COC), codeine (COD), methylenedioxyamphetamine (MDA) and methylenedioxymethamphetamine (MDMA) were obtained from Lipomed (Arlesheim, Switzerland). Deuterium internal standards (I.S.) 6-MAM-D3, AMP-D5, BEG-D8, COC-D3, COD-D6, KET-D4,

MDA-D5, MDMA-D5, MET-D5, MOR-D6, NORKET-D4 and THC-D3 were purchased from Cerilliant.

Acetonitrile (ACN) and methanol were obtained from J.T. Baker (Center Valley, PA); whilst ammonium formate and formic acid were from Fluka (Seelze, Germany). Dichloromethane and isopropanol were purchased from Sigma-Aldrich (Darmstadt, Germany).

Isolute® SLE+ supported-liquid extraction (SLE) 400 µL columns were obtained from Biotage (Uppsala, Sweden). Quantisal® synthetic negative oral fluid (pre-diluted in extraction buffer) and Quantisal® oral fluid collection devices were purchased from Alere (Waltham, MA).

The ROFT device DrugWipe® 6S was purchased from Securetec (Neubiberg, Germany), Ora-Check® from Safecare Biotech (Hangzhou, China) and SalivaScreen® from Ulti med Products (Ahrensburg, Germany).

2.2 ROFT field test

Subjects were recruited from the Hospital Authority substance abuse clinics at Castle Peak Hospital (CPH), Kwai Chung Hospital (KCH) and Pamela Youde Nethersole Eastern Hospital (PYNEH), as well as the Society of Rehabilitation and Crime Prevention (SRACP). Informed consent was obtained from all subjects, who were at least 18 years of age. Repeated sampling was allowed provided that each collection was at least one week apart. The protocol had been approved by the Hospital Authority Kowloon West Cluster Research Ethics Committee.

For each subject, a confirmation sample was firstly collected using the Quantisal® oral fluid collection device. The sampling sponge was placed in the subject's oral cavity and left there for 10 minutes (or when the indicator turned blue, whichever was earlier). The sponge, which was supposed to have collected 1 mL of oral fluid, was then deposited into the designated tube containing 3 mL of buffer. This sample was subsequently transported back to the laboratory and the weight of the whole tube was recorded for adjusting the volume of oral fluid collected. The sample was then stored

at 4°C for 3 days, after which a plunger separator was used to harvest all the buffered oral fluid inside the tube. The oral fluid sample was then stored in a separate container at -80°C until analysis. Those samples with weight corresponding to less than 0.5 mL oral fluid (unless if the volume indicator turned blue) were not subjected to confirmation analysis.

All three ROFT devices were evaluated sequentially on each subject. Towards the end of the study, however, when four positive cases have already been achieved for each analyte on a device, testing on this device would terminate. Some subjects would not have sufficient oral fluid to complete all three evaluations. DrugWipe® 6S required the least amount of oral fluid (approximately 0.1 mL), thus was tested last of the three. In order to have similar number of completed tests for Ora-Check® and SalivaScreen®, these two devices were tested first on alternate days. The three ROFT devices are shown in Fig 1.

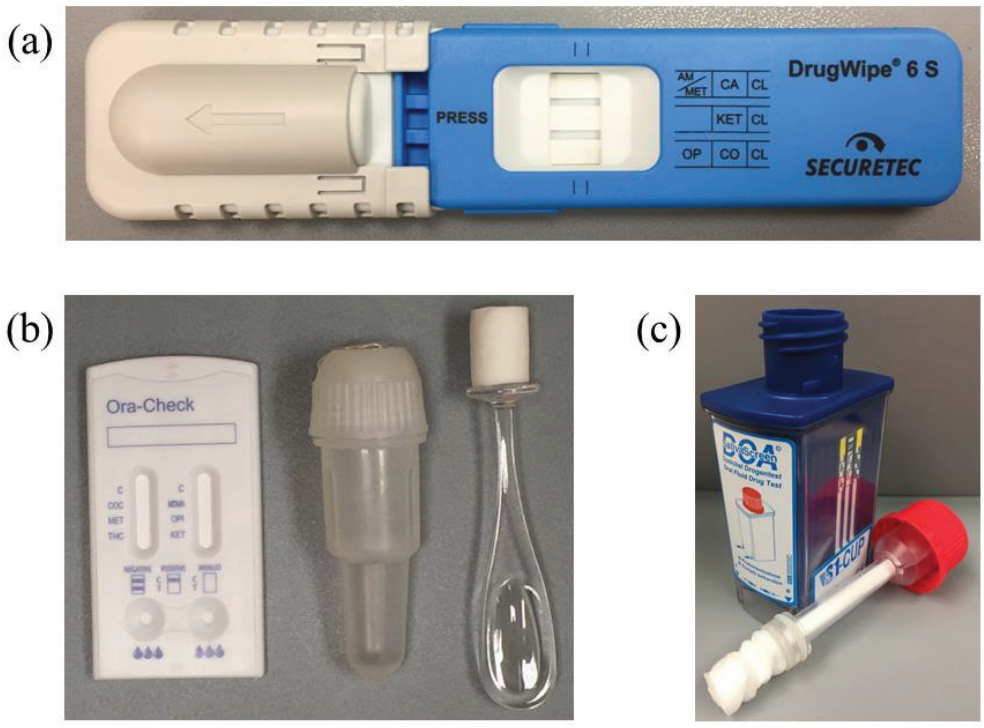

Fig 1. ROFT devices included in the study: (a) DrugWipe® 6S; (b) Ora-Check®; (c) SalivaScreen®

Ora-Check® and SalivaScreen® were capable of separately testing all 6 drugs: ketamine, methamphetamine, cannabis, cocaine, MDMA and opiates (none of the devices could differentiate among heroin metabolite 6-MAM, codeine or morphine). DrugWipe® 6S only detected 5 types of drugs: ketamine, cannabis, cocaine, opiates and the amphetamines. This device was unable to differentiate among amphetamine-type drugs; this class of drugs was tested collectively by one "AMP/MET" test.

The DrugWipe® 6S device consists of a sample collector containing 3 small sampling pads, the test cassette and an integrated liquid ampoule. Oral fluid is collected by wiping the sampling pads on the tongue several times until the pads change colour. The collector is then placed back onto the test cassette, with the pads in contact with the test strips. The device is held vertically; the liquid ampoule is broken by compression and the buffer flows along the test strips. After 10 seconds, the device is placed on a horizontal surface and the results may be read after 8 minutes. Result interpretation was performed according to the manufacturer's instructions (i.e. a visible band indicates a positive result. Faint bands were regarded as positive).

The Ora-Check® device comprises a sampling sponge, a collection chamber and the test cassette. The sponge is placed in the subject's mouth for 3 minutes (with occasional sweeping motion), during which supposedly 0.5 mL oral fluid will have been collected. The sponge is then firmly pushed into the collection chamber to harvest the oral fluid. The chamber is inverted and the oral fluid is transferred through the dropper onto the sampling area of the test cassette. After 10 minutes, results are interpreted according to the manufacturer's instructions (i.e. a visible band indicates a negative result. Faint bands were regarded as negative).

The SalivaScreen® device consists of a sampling sponge with volume indicator and a test cassette that extracts the oral fluid and houses the test strips. The subject is first instructed to sweep the sampling sponge inside the oral cavity several times and leave the sponge inside for 7 minutes (or when the 1mL volume indicator turns red, whichever is earlier). The sponge is then pushed into the test cassette to release the oral fluid. The device is left on a flat surface for 10 minutes, after which results may be read according to the manufacturer's instructions (i.e. a visible band indicates a negative result. Faint bands were regarded as negative).

For all devices, absence of the quality control (QC) band indicates a failed test, i.e. QC failure, and the results were regarded as invalid.

2.3 Standards, calibrators and quality controls

All standards and I.S. were supplied in ampoule form at concentrations of 0.1 or 1 mg/mL in methanol. Calibrators and QC were prepared by spiking synthetic negative oral fluid (pre-diluted in extraction buffer) with the standards. Since any oral fluid collected from participants (presumably 1 mL) was immediately diluted 4-fold once it was deposited into the plastic tube containing 3 mL buffer, this was taken into account when spiking the calibrators and QC, i.e. the concentration in the spiked calibrator/QC was 4-fold lower than the actual concentration in the original undiluted oral fluid sample.

Calibrators were spiked at the following concentrations (in undiluted oral fluid): THC (0.5-200 ng/mL); 6-MAM, COC and BEG (1-200 ng/mL); AMP, MET, MOR, COD, MDMA and MDA (5-500 ng/mL); KET and NORKET (5-1500 ng/mL). Three levels of QC were prepared by spiking at the low and high ends as well as near the middle of the calibration range of each analyte.

I.S. mix was prepared in 50% methanol at the following concentrations: AMP-D5 at 5 ng/mL; MET-D5, MDMA-D5, KET-D4, MOR-D6, THC-D3, BEG-D8, 6-MAM-D3 and NORKET-D4 at 50 ng/mL; MDA-D5 at 100 ng/mL; COC-D3 at 250 ng/mL; COD-D6 at 1000 ng/mL.

2.4 Oral fluid analysis

All calibrators, QC and participant samples were equilibrated to room temperature and mixed thoroughly. Samples that required dilution was first diluted 50-fold using blank oral fluid. To 400 µL of oral fluid sample was added 50 µL of I.S. mix. After vortex mixing, the sample was subjected to supported liquid extraction (SLE). A 400 µL aliquot of sample was loaded onto the SLE column. Vacuum was applied briefly to initiate flow, and the sample was allowed to flow into the column bedding for 5

minutes. The sample was then eluted by repeating the following step twice: 1 mL of elution solution (dichloromethane:isopropanol 70:30 $^{v}/_{v}$) was loaded and allowed to flow by gravity for 5 min. A strong vacuum was applied at the end to ensure completion elution. The eluate was collected and dried under nitrogen at 40°C. The sample was then reconstituted with 100 µL reconstitution solution (mobile phase A:methanol 50:50 $^{v}/_{v}$). For THC, MOR, 6-MAM and AMP, the reconstituted fraction was injected directly for LCMS analysis. For the other analytes, the fraction was diluted 25-fold with reconstitution solution prior to LCMS analysis.

2.5 LCMS analysis

LCMS analysis was performed on the Sciex 5500 QTrap triple-quadrupole mass spectrometer (Framingham, MA, USA) equipped with turbo ion spray source and Waters Acquity UPLC consisting of a binary solvent manager, a column manager and a sample manager (Milford, MA, USA). The temperature of the thermostat column compartment was 40°C; whilst the autosampler remained at ambient temperature. Chromatographic separation was performed with a Waters Acquity HSS C18 SB column (1.8 µm, 2.1x100 mm) and gradient elution comprising 1 mM ammonium formate, 0.1% formic acid in water (mobile phase A, MPA) and 1 mM ammonium formate, 0.1% formic acid in ACN (mobile phase B, MPB). The gradient program started at 0.25 mL/min flow with 2% MPB at 0 - 0.5 min, increasing to 20% at 5 min and held until 7 min. The MPB content was further increased to 85% by 10.5 min. At 10.51 min, the flow was increased to 0.3 mL/min and held until 11.5 min. The MPB content was subsequently reverted to 2% by 12 min, with further equilibration at starting conditions until the run stops at 13 min. In between injections, the auto-injector was washed sequentially with 1 mL of 50% ACN and 1.8 mL of 100% ACN.

Analytes were detected by mass spectrometry using scheduled multiple reaction monitoring (MRM) in positive electrospray ionization mode. Analytes were monitored within a ±15 seconds retention time (RT) window. The dwell time was automatically calculated by the software under the scheduled MRM mode with a total cycle time of 0.4 second. For each analyte, the following MS parameters were applied: dwell weight 1, entrance potential (EP) 10 V and cell exit potential (CXP) 13 V. The

source parameters were: curtain gas 30 psi, high collision gas, IonSpray Voltage 5500 V, temperature 600°C, ion source gas 1 (nebulizer gas) 50 psi and ion source gas 2 (turbo gas) 50 psi. Table 1 shows the analytes detected by the method and the LCMS parameters used. The declustering potential (DP), the most abundant product ions and their respective collision energies (CE) were first optimized for each compound by infusion using a 100 ng/mL standard solution.

Table 1. LCMS parameters

Analyte	RT (min)	DP (V)	Q1 m/z	Q3 m/z	CE (V)
AMP 1	5.01	75	136	91	65
AMP 2				65	23
IS_AMP-D5	5.01	36	141	96	15
6-MAM 1	5.79	180	328	165	33
6-MAM 2				152	62
IS_6-MAM-D3	5.79	120	330.7	165	47
MOR 1	3.73	190	286	165	53
MOR 2				115	85
IS_MOR-D6	3.73	138	292.1	152	67
THC 1	12	134	315	193	34
THC 2				123	43
IS_THC-D3	12	112	318.1	196.1	31
KET 1	6.91	80	238	125	17
KET 2				89	51
IS_ KET -D4	6.91	74	242	129	39
NORKET 1	6.46	80	224	125	17
NORKET 2				89	55
IS_ NORKET -D4	6.46	45	228	129	33
COD 1	5.36	160	300	152	79
COD 2				115	95
IS_ COD -D6	5.36	120	306.1	152	47
MET 1	5.62	75	150	91	11
MET 2				119	8
IS_ MET -D5	5.62	53	155	92	16
MDMA 1	5.97	75	194	163	7
MDMA 2				105	18
IS_MDMA-D5	5.97	68	199	165	12
MDA 1	5.45	60	180	105	29
MDA 2				77	49
IS_MDA-D5	5.45	41	185	110.1	31
COC 1	9.37	155	304	182	18
COC 2				77	49
IS_ COC -D3	9.37	120	306.9	185.1	47
BEG 1	8.89	170	290	168	21
BEG 2				77	54
IS_BEG -D8	8.89	163	298.1	171.1	20

Positive identification of an analyte was based upon the following criteria: (i) retention time (RT) within ±0.5 min of reference standard; (ii) MRM ratio within tolerance limits as defined by the European Communities and the Clinical and

Laboratory Standards Institute (CLSI), as follows: ratio >50% (± 20%); ratio 20-50% (± 25%); ratio 10-20% (± 30%); ratio ≤10% (± 50%). In each analysis batch, the MRM ratios of the calibrators were averaged to establish the reference MRM ratio for the analyte.

Compounds were quantified by comparing the analyte/I.S. peak area ratio against the calibration curve. It should be noted that calibration were based upon the assumption that 1 mL of oral fluid was diluted in 3 mL of buffer. However, for participant samples, the actual volume of oral fluid collected might not be exactly 1 mL. Hence, the calculated concentration had to be adjusted according to the actual volume of oral fluid collected. This adjustment could be made using the following formula:

$$C\ adjusted = \frac{C\ unadjusted\ x\ (3 + w - w')}{4\ x\ (w - w')}$$

where:

C *adjusted* = analyte concentration with adjustment of oral fluid volume collected

C *unadjusted* = unadjusted analyte concentration

w = weight of sample and Quantisal® oral fluid collection tube

w' = average weight of Quantisal® oral fluid collection tubes (n=30) without sample

2.6 Method validation

The analytical method was validated according to international guidelines and published protocols [18-22]. The protocol included evaluation of selectivity, linearity, limit of quantitation (LOQ), accuracy, precision, extraction efficiency, matrix effect, carryover, dilution integrity and stability. Analyte recovery from the Quantisal® oral fluid collection device was also assessed.

Interference from endogenous components was assessed by analysing 10 blank oral fluid matrices. Method selectivity was evaluated by spiking blank oral fluid with high concentrations (500 ng/mL) of possible interfering compounds including caffeine, paracetamol, chlorpheniramine, promethazine, dextromethorphan, phentermine, methadone, famotidine, diclofenac, hyoscine butylbromide, terazosin, pyridium,

ciprofloxacin and alpraxolam. The method was deemed to be selective if no analyte could be detected that fulfilled all identification criteria.

Linearity was determined by least-squares regression with 1/x weighting (n=5). Acceptable linearity was defined as having coefficient of determination (r^2) >0.995 and the calibrators could be quantified within ±20% for LOQ and ±15% for all other levels. The accuracy and precision at LOQ, defined as the lowest calibration level, was verified by analysing five replicates over three days. The accuracy at LOQ should be within ±20% and the imprecision (expressed as the coefficient of variation, CV) <20%.

Accuracy was assessed using external quality assurance (EQA) samples from LGC Standards Proficiency Testing (Lancashire, UK); satisfactory performance was defined as having a z-score of within ±2 (z-score=deviation from assigned value/standard deviation for proficiency assessment). Norketamine, for which EQA was not available, was assessed by analysing blank oral fluid spiked with known concentrations of the analyte (5 replicates across 4 days); the accuracy should be within ±15%.

Evaluation of precision involved spiking analytes into blank oral fluid at three concentrations (low, mid, high). These spiked samples were assayed in five replicates over four days. Precision was deemed to be acceptable if the within-day, between-day and total imprecision were less than 15% CV.

The extraction efficiency and matrix effect were assessed in a single experiment containing 3 sets of samples, as proposed by Matuszewski et al [23]. The same amount of analytes (at low and high ends of the calibration range), plus a constant concentration of I.S., were spiked into matrix-free solvent and 2 sets of blank oral fluid from six different sources. For the latter, the specimens were spiked with analytes either before or after extraction. The peak areas of the analyte in matrix-free solvent (A), standards spiked into different matrices before extraction (B) and after extraction (C) were determined. The extraction efficiency and overall matrix effect (expressed as the matrix factor, MF) were then assessed at each concentration as follows: Extraction efficiency (%) = Peak area from [B]/Peak area from [C] x 100. Matrix factor (MF) = Peak area from [C]/Peak area from [A]. The I.S.-normalised MF

was calculated for each analyte at the respective concentration using the formula: I.S.-normalised MF = analyte MF/I.S. MF. The precision of the I.S.-normalised MF was expressed as the %CV and should ideally be <15%.

Carry-over was assessed by inspecting the blank matrix run following injection of the highest calibrator, and was considered acceptable if the carry-over was below LOQ. Dilution integrity was evaluated by spiking blank oral fluid at high concentrations (1800-14000 ng/mL) and analysing with 50-fold dilution in replicates of five. Three measurements without dilution were averaged to provide the reference concentration. The accuracy (%) was calculated by (average of factor-adjusted concentration/Reference concentration) x100 and was considered acceptable if the deviation was <15%. The precision of the dilution step was expressed as the %CV of the 5 measurements and should be <15%.

Analyte stability was evaluated for oral fluid spiked at low and high concentrations stored at -80°C (for 4, 6 or 8 weeks) or after three freeze/thaw cycles (one cycle includes freezing at -80°C overnight followed by defrosting for 3 hours at room temperature). The post-preparative stability was assessed for samples stored at 4°C for up to 4 days. All assessments were done in triplicates. Stability was considered acceptable if the deviation from reference samples (not subjected to any storage or freeze/thaw cycle) was within ±15%.

The recovery of analytes from the Quantisal™ oral fluid collection device was assessed by adding 1 mL of oral fluid spiked with analytes (at low, mid and high concentrations) to the collection pad. The collection pad was left in the buffer (each device contains 3 mL) and stored at different temperatures for certain time points, including: room temperature for 1 day, 4°C for 1 day, 4°C for 2 days, 4°C for 3 days and 4°C for 4 days. In addition, one set of samples was analysed directly without any storage. The buffered oral fluid was then separated from the collection pad using a plunger and then analysed. To establish the reference value, 1 mL of the neat oral fluid (without adding to device) was diluted with 3 mL buffer and analysed. Triplicate was performed at each concentration. The recovery at each concentration was calculated by: Recovery (%) = average of samples using device/average of reference without using device. A recovery rate of >80% was considered desirable.

2.7 Data interpretation

In order to evaluate the sensitivity, specificity and overall accuracy of the ROFT devices, the analyte concentrations measured by LCMS (and adjusted for volume of oral fluid collected) were used as the "gold standard" result. These results were compared against the DRUID cut-off; for ketamine and norketamine, no DRUID cut-off was available, hence the LCMS cut-off (LOQ of the method) was used. If any drug or its cross-reacting compound is quantitated at or above the respective cut-off, the result is considered to be positive. A summary of the DRUID and LCMS cut-offs, as well as the manufacturer-claimed device cut-offs, is shown in Table 2.

Table 2. Summary of cut-off values

	Cut-off value (ng/mL)				
	DrugWipe® 6S	Ora-Check®	SalivaScreen®	DRUID	LCMS
KET	5	50	25	N/A	5
NORKET	75	50	30	N/A	5
MET	80	50	50	25	5
AMP	80	--	--	25	5
MDMA	25	50	50	25	5
MDA	10	250	250	25	5
6-MAM	5	25	10	5	1
COD	5	10	8	20	5
MOR	10	40	10	20	5
COC	10	20	20	10	1
BEG	75	20	200	10	1
THC	20	50	50	1	0.5

In this way, the ROFT field test data could be classified into the following categories: true positive (TP) where a positive ROFT device result matches a positive LCMS result; true negative (TN) where a negative ROFT device result matches a negative LCMS result; false positive (FP) where the ROFT device result was positive but with a negative gold standard result; and false negative (FN) where the ROFT device result was negative but the gold standard result was positive.

Taking into consideration the above classification, the following parameters could be calculated:

Sensitivity (%) = TP/(TP+FN)*100

Specificity (%) = TN/(TN+FP)*100

Accuracy (%) = (TP+TN)/(TP+TN+FP+FN)*100

Prevalence (%) = (TP+FN)/total no. of results*100

Positive predictive value (PPV) (%) = TP/(TP+FP)*100

Negative predictive value (NPV) (%) = TN/(TN+FN)*100

Evaluation of the above parameters was only conducted for analytes with at least four positive cases.

3. Results

3.1 Validation results

Ten oral fluid matrices collected from drug-free volunteers and blank oral fluid spiked with high concentrations of common drugs/co-medications were evaluated for interference. None of the analytes was positively identified in any matrix or in the presence of the common drugs/co-medications examined. The method selectivity was found to be satisfactory.

Linear calibration curves were constructed by least-squares regression with 1/x weighting for all analytes with coefficients of determination >0.995. The deviation from nominal value was within ±20% for LOQ and ±15% for all other levels. At the LOQ, the accuracy (within ±20%) and precision (CV <20%) were satisfactory for all analytes. Accuracy evaluation showed acceptable results with z-score within ±2 for all analytes; for norketamine, the accuracy was within ±15% of the spiked values across all concentrations tested. These results are summarized in Table 3.

The within-day, between-day and total imprecision were satisfactory (CV<15%) for all analytes at low, mid and high concentrations. The extraction efficiency ranged from 78%-120%. The matrix factor (MF) ranged between 0.3 and 1.07; the precision of the I.S.-normalised MF was acceptable with CV<15% for all analytes (CV range: 2.6%-10.7%). The results are summarized in Table 4.

Table 3. Linearity, limit of quantitation (LOQ) and accuracy results

Analyte	LOQ (ng/mL)	Linearity (ng/mL)	Intercept ± SD (n=5)	Slope ± SD (n=5)	r^2 ± SD (n=5)	Accuracy	
						z-score[a]	% of spike value[b]
KET	5	5-1500	0.0756±0.1641	0.0047±0.0002	0.9997±0.0002	-0.2 to 0.3 (2)	--
NORKET	5	5-1500	0.0068±0.0038	0.0085±0.0004	0.9997±0.0002	--	99-105%
MET	5	5-500	0.0041±0.0044	0.0112±0.0007	0.9986±0.0008	-1.5 to 0.2 (2)	--
AMP	5	5-500	0.1028±0.0291	0.0446±0.0017	0.9994±0.0004	0.3 (1)	--
6-MAM	1	1-200	0.0029±0.0167	0.2182±0.0575	0.9997±0.0002	0.2 to 0.4 (3)	--
COD	5	5-500	-0.0005±0.0093	0.0264±0.0017	0.9991±0.0007	-0.1 to 0.1 (2)	--
MOR	5	5-500	0.0536±0.0255	0.1188±0.0076	0.9993±0.0006	0.1 to 1.7 (3)	--
COC	1	1-200	0.0025±0.0093	0.0436±0.0050	0.9996±0.0002	0 to 0.2 (2)	--
BEG	1	1-200	-0.0045±0.0062	0.0505±0.0016	0.9996±0.0003	-0.8 to 0.1 (2)	--
THC	0.5	0.5-200	0.0051±0.0024	0.0477±0.0029	0.9995±0.0001	-0.8 to 0.6 (3)	--
MDMA	5	5-500	-0.0020±0.0031	0.0071±0.0003	0.9992±0.0009	-0.1 to 0.3 (2)	--
MDA	5	5-500	0.0037±0.0064	0.0185±0.0004	0.9993±0.0007	-0.1 to 0.3 (2)	--

[a]z-score calculated by EQA body. A range is given except when only one result is available (number of samples given in parenthesis).
[b]The accuracy of norketamine was assessed at low, mid and high concentrations. The accuracy range is shown here.
SD-standard deviation

Table 4. Precision, extraction efficiency and matrix effect results

Analyte	Within-day precision (%CV)			Between-day precision (%CV)			Extraction efficiency (%)		Matrix factor		Precision of I.S.-normalised MF (%CV)	
	Low	Mid	High	Low	Mid	High	Low	High	Low	High	Low	High
KET	3.3	1.7	0.7	3.8	3.1	2.2	99	99	0.99	1.07	6.0	2.8
NORKET	3.5	1.9	2.3	0.5	2.6	4.1	98	96	0.73	0.70	5.7	3.3
MET	6.7	5.8	4.3	8.5	5.7	1.7	99	97	0.99	1.03	10.7	2.9
AMP	2.2	1.9	2.1	2.9	2.4	7.1	98	103	0.75	0.88	3.5	3.1
6-MAM	3.8	2.4	1.9	4.3	3.2	3.3	111	109	0.31	0.46	5.5	2.9
COD	3.7	3.6	7.2	4.5	3.7	6.9	105	101	0.91	0.97	4.0	3.8
MOR	1.0	1.6	0.7	2.4	2.5	2.1	120	115	0.30	0.33	4.7	2.7
COC	2.6	1.7	2.5	5.1	3.3	0.9	103	99	0.90	1.01	5.3	3.4
BEG	3.3	1.4	1.2	2.7	0.7	1.3	101	99	1.00	1.05	3.9	2.6
THC	3.7	3.5	2.7	1.6	10.6	9.2	84	78	0.57	0.53	6.4	5.5
MDMA	3.8	3.8	3.5	8.0	1.5	1.5	100	99	0.93	0.99	5.2	2.6
MDA	6.6	3.5	5.8	7.2	4.7	2.8	100	98	0.95	0.99	7.5	3.9

In the carryover study, no analyte peak was detected in the blank run following injection of the highest calibrator. The 50-fold dilution integrity of the method was assessed and found to have acceptable accuracy (97%-109%) and precision (CV 2.0%-9.8%).

The long term stability of analytes in oral fluid when stored at -80°C for 4, 6 or 8 weeks was evaluated; results showed that across all time points, the deviation was within ±15% for all analytes (range: -11.6% to 10.6%). Similarly, analytes were found to be stable across three freeze/thaw cycles (range: -13.5% to 3%). The stability of processed samples was shown to be 2 days, with the measured concentration of all analytes having -8.1% to 9.4% variation. On day 3, the lowest calibrator of THC was below the integration threshold; since the calibration curve could not be constructed, the results could not be calculated. On day 4, cocaine and THC showed -35.9% and -17.1% variation respectively compared with reference. The results of the stability studies are summarized in Table 5.

Table 5. Stability results

Analyte	Storage of oral fluid at -80°C[a]						3 Freeze/thaw cycles[a]		Processed samples stored at 4°C[a]					
	Week 4		Week 6		Week 8				Day 2		Day 3		Day 4	
	Low	High	Low	High	Low	High	Low	High	Low	High	Low	High	Low	High
KET	-1.6	-3.2	1.5	-2.0	-6.1	-4.8	-7.2	-4.1	2.2	-1.8	-12.7	1.4	2.1	0.7
NORKET	-0.6	1.1	3.2	-2.7	-10.0	-4.1	-5.0	-2.9	7.5	-2.0	-6.8	-2.2	-4.0	-2.6
MET	6.9	0.5	-7.9	0.8	5.5	-9.1	-13.5	0.1	4.2	9.4	0.4	2.2	-7.2	-5.2
AMP	-4.9	9.3	4.8	7.7	-6.3	-4.4	-4.4	-1.4	-5.7	-6.1	-12.0	-6.5	1.2	-3.3
6-MAM	1.2	3.0	4.1	-2.7	5.0	-3.3	-5.3	-8.3	0.0	2.3	-4.9	1.0	4.4	0.5
COD	10.6	-10.2	9.3	-1.8	-4.5	-0.6	3.0	-3.1	-8.1	-2.8	-6.4	2.3	4.7	4.4
MOR	8.0	3.2	6.3	1.3	-2.2	-5.7	-1.2	-2.7	-3.1	-2.4	-5.5	-4.6	0.2	-2.2
COC	-3.3	-1.3	-7.7	0.4	-8.2	-7.6	-11.8	-5.1	3.8	-0.8	-9.2	-4.7	-35.9	-12.4
BEG	7.9	-0.2	-1.2	0.5	-1.5	-1.8	-2.3	0.2	0.3	1.0	-7.9	1.3	6.4	3.9
THC	3.0	1.1	3.0	-6.1	-0.3	-9.7	0.9	-9.9	-4.1	1.4	--[b]	--[b]	-17.1	-5.4
MDMA	2.0	-1.1	-0.1	-0.8	-2.8	-1.0	1.9	-1.7	7.9	6.5	3.7	0.9	-7.7	1.0
MDA	-5.2	-4.1	-5.9	2.2	-11.6	2.2	-0.5	-2.4	-4.3	2.5	10.6	-2.9	12.0	0.1

[a]Stability results are expressed as the deviation (%) from reference values

[b]Results not calculable since the calibration curve could not be constructed (lowest calibrator below integration threshold)

In order to investigate the optimal storage conditions for maximum recovery of analytes from the Quantisal® device, analytes were spiked onto the device and stored at different temperatures and for different durations prior to analysis. Results showed that THC was poorly recovered from the collection device on the first 2 days (recovery: 48.7%-67.5%); other analytes like norketamine, 6-MAM, codeine, BEG and MDMA also had marginal recovery (77.6%-79.7%). Upon storage at 4°C for 3 days, all analytes had >80% recovery; this storage condition was chosen for all subsequent analysis. On day 4, the recovery of THC and cocaine decreased again (53.4% and 76% respectively).

3.2 LCMS analysis

In total, 549 samples were collected in the study – 207 (38%) from SRACP, 173 (32%) from CPH, 100 (18%) from PYNEH and 69 (13%) from KCH. Among the 549 samples, 491 (89%) could be subjected to LCMS analysis whilst the remainder did not have sufficient oral fluid for confirmation analysis.

Opiates were the most commonly encountered drugs with prevalence of 55% (codeine), 49% (morphine) and 40% (heroin). This was followed by methamphetamine (35%). Ketamine, THC and cocaine were detected at relatively low prevalence rates (2%-8%). MDMA was not detected in any samples. The LCMS analysis results of individual analytes are summarized in Table 6.

Table 6. The number of positive samples and concentrations detected by LCMS analysis

	No. of positive samples	Concentration (ng/mL)		
		Mean	Median	Range
KET	18 (4%)	4887	210	6 - 55136
NORKET	18 (4%)	406	165	7.4 - 2270
MET	174 (35%)	1917	602	5.1 - 23612
AMP	157 (32%)	310	96	5.3 - 16713
6-MAM	197 (40%)	587	28	1.1 - 25436
COD	269 (55%)	1515	93	5 - 40776
MOR	239 (49%)	553	132	5 - 16337
COC	9 (2%)	123	10	1.2 - 753
BEG	8 (2%)	24	19	1.4 - 59
THC	39 (8%)	95	6	0.5 - 1958

3.3 General performance of ROFT devices

All 549 samples were tested on the SalivaScreen® device. For Ora-Check®, 547 tests were done since two subjects refused to perform this test. The number of tests done on DrugWipe® was 515 – testing was not performed on all subjects towards the end of the study, since four positive cases have already been achieved for each analyte.

Many problems were encountered while using the Ora-Check® device. Despite strict adherence to the manufacturer's protocol, in nearly half of the cases the volume of oral fluid collected was insufficient for the testing to continue. Specifically, after placing the collection sponge in the subject's mouth for the designated duration (3 min), the sponge was still too hard and no oral fluid could be squeezed out of the sponge; as such, the testing could not proceed further since no oral fluid was available for adding to the test cassette. This problem was communicated to the manufacturer, whose advice was to increase the collection time to 5 min. However, this was to no avail and the success rate was not found to increase.

The general performance of the three ROFT devices is shown in Table 7. As mentioned above, the success rate of the Ora-Check® device was very low (52%); this was due to the large number of cases with insufficient oral fluid (n=255) and 5 cases of QC failure. There was 1 case missing analysis (i.e. ROFT testing was completed but the LCMS confirmation sample was insufficient in volume). The overall number of valid samples (successful ROFT and LCMS testing) was 286 (52%).

Table 7. General performance of the ROFT devices

	No. of tests performed	No. of successful tests[a]	No. of failed tests		Missing analysis[c]	No. of samples with LCMS analysis
			Insufficient oral fluid[b]	QC failure		
DrugWipe® 6S	515	510 (99%)	0	5	55	455 (88%)
Ora-Check®	547	287 (52%)	255	5	1	286 (52%)
SalivaScreen®	549	426 (78%)	0	123	6	420 (77%)

[a]Successful test denotes a test that could be completed with QC passed
[b]In cases with "insufficient oral fluid", the collection sponge failed to yield sufficient oral fluid for the testing to continue.
[c]Cases in which the ROFT testing was successfully completed with QC passed, but the volume of oral fluid collected for LCMS analysis was insufficient, were defined as "missing analysis".

In contrast to Ora-Check®, the success rate of DrugWipe® 6S was very high (99%). Due to 5 cases of QC failure and 55 cases missing analysis, the overall proportion of valid samples was 88%.

The general performance of SalivaScreen® was acceptable and lies between the other two devices. The success rate was 78%. All failed tests (n=123) were attributed to QC failure, whilst 6 samples were missing analysis. The overall proportion of valid samples was 77%.

Result interpretation of all three ROFT devices involves the subjective determination of whether a "band" is visible. In terms of easy-readability of the results, feedback from frontline device operators indicated that the result band was more easily distinguishable on SalivaScreen® compared with the other two devices.

3.3 Evaluation of ROFT device performance

A summary of the ROFT evaluation data is presented in Table 8.

For ketamine, sensitivities of 41% (DrugWipe® 6S), 36% (Ora-Check®) and 76% (SalivaScreen®) were achieved. The specificity ranged from 94% - 99% and accuracy 92% - 98% across the three devices. The PPV was particularly low for DrugWipe® 6S (21%) whilst being moderate for the other two devices (Ora-Check® 50%; SalivaScreen® 72%).

Similar to ketamine, variation in the sensitivity of the devices for cocaine was observed – 43% (DrugWipe® 6S), 60% (Ora-Check®) and 71% (SalivaScreen®). On the other hand, all devices achieved 100% specificity and 99% accuracy.

The sensitivity for methamphetamine was 83% for both DrugWipe® 6S and SalivaScreen® and 63% for Ora-Check®. Conversely, higher specificity was observed with Ora-Check® (93%) and DrugWipe® 6S (89%) compared with SalivaScreen® (82%).

For opiates, the sensitivity of Ora-Check® (53%) was remarkably lower than DrugWipe® 6S and SalivaScreen®, which scored 93% and 100% respectively. Specificities of 83% or above were achieved for all devices.

All devices performed poorly in detecting THC-positive cases. Whilst DrugWipe® 6S successfully picked up 7 out of 32 positive cases (22%), the other two devices failed to identify any of the 20+ positive samples. The specificity, on the other hand, was 100% for all devices.

No MDMA-positive case was observed across the entire study; hence, it was not possible to assess the sensitivity of the devices. For Ora-Check® and SalivaScreen®, specificities and accuracies of 96% were achieved.

Table 8. Summary of ROFT device performance

	Ketamine				Methamphetamine[a]				Opiates		
	DrugWipe 6S	Ora-Check	Saliva Screen		DrugWipe 6S	Ora-Check	Saliva Screen		DrugWipe 6S	Ora-Check	Saliva Screen
TP	7	4	13	TP	122	54	106	TP	218	73	212
TN	411	271	398	TN	274	185	241	TN	200	145	171
FP	27	4	5	FP	34	15	52	FP	20	4	36
FN	10	7	4	FN	25	32	21	FN	17	64	1
Total	455	286	420	Total	455	286	420	Total	455	286	420
Sensitivity	41%	36%	76%	Sensitivity	83%	63%	83%	Sensitivity	93%	53%	100%
Specificity	94%	99%	99%	Specificity	89%	93%	82%	Specificity	91%	97%	83%
Accuracy	92%	96%	98%	Accuracy	87%	84%	83%	Accuracy	92%	76%	91%
Prevalence	3.7%	3.8%	4.0%	Prevalence	32%	30%	30%	Prevalence	52%	48%	51%
PPV	21%	50%	72%	PPV	78%	78%	67%	PPV	92%	95%	85%
NPV	98%	97%	99%	NPV	92%	85%	92%	NPV	92%	69%	99%
	Cocaine				THC				MDMA[b]		
	DrugWipe 6S	Ora-Check	Saliva Screen		DrugWipe 6S	Ora-Check	Saliva Screen		DrugWipe 6S	Ora-Check	Saliva Screen
TP	3	3	5	TP	7	0	0	TP	--	0	0
TN	422	280	404	TN	397	265	384	TN	--	275	405
FP	0	0	1	FP	0	0	0	FP	--	11	15
FN	4	2	2	FN	25	20	28	FN	--	0	0
Total	429	285	412	Total	429	285	412	Total	--	286	420
Sensitivity	43%	60%	71%	Sensitivity	22%	0%	0%	Sensitivity	--	n.a.[c]	n.a.[c]
Specificity	100%	100%	100%	Specificity	100%	100%	100%	Specificity	--	96%	96%
Accuracy	99%	99%	99%	Accuracy	94%	93%	93%	Accuracy	--	96%	96%
Prevalence	1.6%	1.8%	1.7%	Prevalence	7.5%	7.0%	6.8%	Prevalence	--	0%	0%
PPV	100%	100%	83%	PPV	100%	n.a.	n.a.	PPV	--	n.a.[c]	n.a.[c]
NPV	99%	99%	100%	NPV	94%	93%	93%	NPV	--	n.a.[c]	n.a.[c]

n.a. – not applicable

[a]Includes the whole amphetamine group for DrugWipe 6S

[b]MDMA included in the "Methamphetamine" test for DrugWipe 6S

[c]These parameters were not calculated since the number of positive samples was <4

Graphical representations comparing the sensitivity, specificity and accuracy across the three devices are shown in Fig. 2. The 95% confidence intervals (CI) were calculated using the modified Wald method [24].

As shown in Fig. 2, SalivaScreen® had the highest sensitivity on average across the analytes. The sensitivities for methamphetamine (MET) and opiates (OPI) were >80% for both DrugWipe® 6S and SalivaScreen®; however, the sensitivities for ketamine (KET) and cocaine (COC) were generally higher with SalivaScreen®. On the other hand, the sensitivity for THC was better with DrugWipe® 6S (Ora-Check® and SalivaScreen® had sensitivities of 0%). None of the tests achieved ≥80% sensitivity with the Ora-Check® device.

In contrast to the wide variability in sensitivities, the specificity and accuracy were in general more consistent and satisfactory across the devices. All parameters were >80% except for the accuracy of Ora-Check® in detecting OPI.

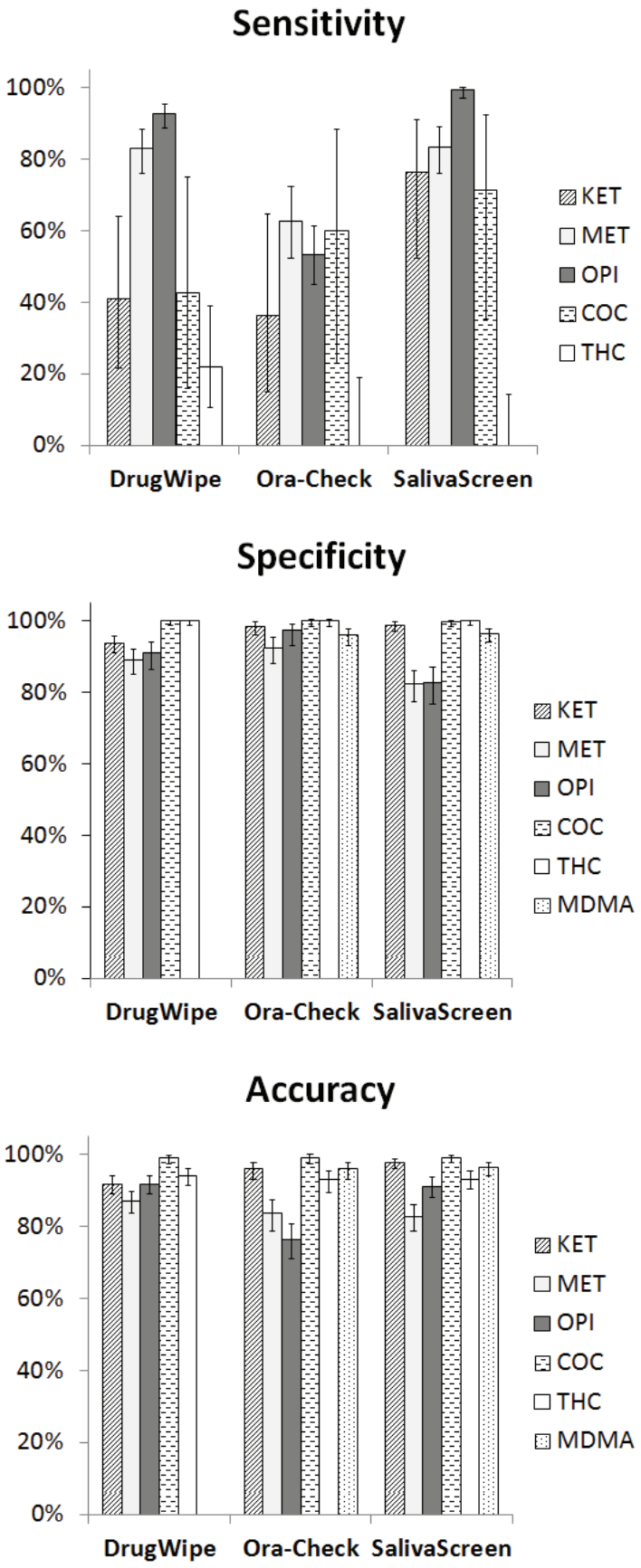

Fig. 2. Sensitivity, specificity and accuracy (±95% CI) of the ROFT devices

Among the analytes, a larger variability in the sensitivities of the KET, COC and THC tests was observed across the three devices. The concentrations of these analytes and/or their cross-reacting metabolite in the false-negative samples are shown in Table 9. For KET, the mean and max concentrations in DrugWipe® 6S and Ora-Check® false-negative samples were remarkably higher than those of SalivaScreen®. Similarly for COC, the concentrations in DrugWipe® 6S false-negative samples were considerably higher than the other two devices. Conversely, Ora-Check® and SalivaScreen® failed to identify THC concentrations as high as 1958 ng/mL.

Table 9. Concentrations of KET, COC and THC and/or their cross-reacting metabolite in false-negative samples

Test	Device	No. of false-negatives	Analyte	Concentration (ng/mL)			
				Mean	SD	Min	Max
KET	DW	10	KET	436	845	6	2087
			NORKET	192	332	7	1024
	OC	7	KET	137	220	6	616
			NORKET	165	179	7	432
	SS	4	KET	15	11	6	29
			NORKET	13	5	7	19
COC	DW	4	COC	68	102	8	186
			BEG	16	14	1	35
	OC	2	COC	23	18	10	36
			BEG	13	16	1	24
	SS	2	COC	10	--	10	10
			BEG	7	8	1	13
THC	DW	25	THC	36	58	1	208
	OC	20	THC	123	434	1	1958
	SS	28	THC	97	368	1	1958

DW: DrugWipe® 6S, OC: Ora-Check®, SS: SalivaScreen®

4. Discussion

In the current study, a liquid chromatography tandem mass spectrometry method was established for the simultaneous quantitation of ketamine, opiates, methamphetamine, cannabis, cocaine and MDMA as well as their metabolites in oral fluid. The method was fully validated and deemed to be fit for use according to international standards [18-22].

Supported-liquid extraction (SLE), which is an emerging sample clean-up technique that is fast and reproducible, was employed in the present study with encouraging results (extraction recoveries ≥78%). A previous study reported recoveries of 58-76% (and even lower for THC) using SLE on similar drugs of abuse [25]. In terms of the matrix effect, ion suppression was more apparent with 6-MAM, morphine and THC (43-70% suppression), as has been reported previously [19, 26]. Deuterated internal standards were used in order to compensate for such ion suppression effects.

The LOQ established presently were comparable to published methods [14, 15, 22] and were, without exception, lower than the DRUID cut-offs [27]. The method was selective for the drugs detected and had satisfactory precision and accuracy. In addition, the analytes were found to be stable in oral fluid under practical storage conditions (up to 8 weeks at -80°C or upon 3 freeze/thaw cycles).

The Quantisal® oral fluid collection device has been shown to have good analyte recovery in previous evaluations [28-30] and was chosen for the present study. Satisfactory performance was also observed presently with >80% recovery of the analytes. Since the storage duration and temperature may also affect the extraction of analytes from the collection sponge into the buffer, these parameters were evaluated. The optimal conditions were found to be storage at 4°C for 3 days. Prior to 3 days, certain analytes might not have sufficient time to extract into the buffer; alternatively, after 3 days susceptible analytes (e.g. THC and cocaine are known to be relatively unstable [22]) could be prone to degradation.

The current study population included predominantly patients undergoing drug rehabilitation or persons known to be active drug users, hence a higher prevalence compared with the normal population is expected. This choice is justified since a larger number of positive samples will yield more accurate and precise findings [31]. Indeed, as with previous studies adopting a similar approach, results would not be interpreted for a particular analyte if the number of positive specimens was less than four [13, 27].

In the present study, a total of 549 oral fluid samples were collected from participants, among which confirmation analysis was performed on 491 samples. Analysis revealed that the most prevalent drugs detected were opiates (codeine 55%, morphine

49%, 6-MAM 40%), followed by methamphetamine (35%) and THC (8%). Despite a lack of local prevalence data on drugs of abuse detected in oral fluid, the Central Registry of Drug Abuse in Hong Kong reported similar statistics with heroin and methamphetamine being the most commonly abused substances [32]. A recent local study also reported opiates and methamphetamine as the most prevalent drugs detected in the urine of 964 drug abusers [33]. On the contrary, perhaps owing to the difference in sample matrix and study population, THC was detected at a higher rate (8% versus 3%) and ketamine at a lower rate (20% versus 4%) in the present evaluation. In both studies, cocaine was detected at a relatively low frequency and MDMA was not detected at all.

The oral fluid concentrations of ketamine and norketamine detected in the current evaluation were similar to previously reported studies (6-14431 ng/mL and 7.4-2270 ng/mL respectively) [14, 34], except for the grossly elevated ketamine level in one sample (55136 ng/mL), which might be due to oral contamination by recent drug use. Comparison with a previous study conducted in Belgium [11] of the median drug concentrations of cocaine and BEG showed lower levels in the current study (cocaine 52.2 versus 10 ng/mL; BEG 81.5 versus 19 ng/mL). This could partly be explained by the lower prevalence of cocaine use or variation in the dosage across different regions. Apart from cocaine, considerable (>3-fold) difference in the median drug concentrations of amphetamine (685.1 ng/mL) and THC (31.4 ng/mL) was also observed in the Belgium study. Oral fluid concentrations of other analytes in the current study were broadly similar to those in previous reports [11, 27].

In the present study, three ROFT devices - DrugWipe® 6S, Ora-Check® and SalivaScreen® - were chosen for evaluation. To the investigators' knowledge, these three devices were the only ones commercially available in Hong Kong at the time of the study that included all the six specified illicit drugs. Previous versions of DrugWipe® (mainly for detecting 5 drugs) have been extensively studied [10, 12, 13, 27], whilst Ora-Check® and SalivaScreen® have thus far not been tested on authentic oral fluid samples before.

In addition to analytical accuracy, the success rate of testing (completion of the test with QC pass) is another important factor in determining the usefulness of a ROFT

device. In this regard, DrugWipe® 6S has excellent performance with a success rate of 99%. On the contrary, nearly half of the tests performed on Ora-Check® were unsuccessful. In the majority of unsuccessful cases, the testing could not proceed beyond the collection step since the sponge failed to yield any oral fluid for the testing to continue, despite strict adherence to the manufacturer's protocol. In the investigator's opinion, the sponge was too hard such that even after the designated collection duration, it still could not soften enough to yield any oral fluid. Indeed, the sponge of the SalivaScreen® device is much softer and no unsuccessful tests have been observed due to failure in harvesting oral fluid from the sponge.

On the other hand, compared with DrugWipe® 6S, the QC failure rate of SalivaScreen® was considerably higher (22% versus 1%). This may again possibly be due to not having sufficient oral fluid collected. SalivaScreen® was designed to collect 1 mL of oral fluid, while DrugWipe® 6S only required ~0.1 mL. In drug users who often have reduced salivation [2] (and from whom 1 mL of oral fluid have already been collected for confirmation analysis), the likelihood of having sufficient oral fluid to complete DrugWipe® 6S testing is understandably much higher than that of SalivaScreen®.

Similar to previous reports, the specificity and accuracy of the ROFT devices were in general satisfactory and met the DRUID recommendation of >80% (except for the 76% accuracy of Ora-Check® in detecting opiates). So far, the problem encountered with most ROFT devices has been the sensitivity. In most studies, none of the devices could reach 80% sensitivity for all the detected analytes; in particular, the cocaine and THC tests have always been problematic [11, 13, 27]. In the present study, as shown in Fig. 2, the sensitivity of the methamphetamine and opiates tests reached 80% for DrugWipe® 6S and SalivaScreen®.

Similar to published data [11, 27], the sensitivity of the cocaine and THC tests was considerably lower in comparison and was <80% across all presently studied devices. Indeed, Ora-Check® and SalivaScreen® failed to identify any of the 20+ THC-positive cases; whilst DrugWipe® 6S, albeit far from satisfactory, achieved the highest sensitivity at 22%. In terms of cocaine, SalivaScreen® achieved the highest sensitivity (71%), followed by Ora-Check® (60%) and lastly DrugWipe® 6S (43%). Previous

studies adopting the DRUID cut-off reported THC and cocaine sensitivities of 43-47% and 90%, respectively, for DrugWipe® 5+ [13, 27]. Due to the low prevalence of cocaine and THC in the present study, the 95% CI of the sensitivity were relatively wide for these two analytes (Fig. 2).

For ketamine, the sensitivity was also observed to vary among the three devices. SalivaScreen® achieved the highest sensitivity (76%), while the other two devices achieved only 36-41%. A wide 95% CI was observed due to the relatively low number of ketamine-positive samples in the current study. All three devices had satisfactory specificities (≥94%). In terms of the PPV, however, DrugWipe® 6S had a considerably lower value than SalivaScreen® (21% versus 72%) due to the high number of false-positive results (n=27) in comparison to true-positives (n=7). This indicates that when used in the field, a higher proportion of "positive calls" by DrugWipe® 6S will turn out to be false signals compared with SalivaScreen®.

A previous study evaluated the performance of another device, OratectXP - this device was used for detecting ketamine only and not the other analytes [14]. In this study, the manufacturer's device cut-off (15 ng/mL) was employed in the interpretation of results, with the calculated sensitivity, specificity and accuracy being 88%, 98% and 94% respectively. SalivaScreen® in the present study achieved similar results. When the LCMS cut-off (5 ng/mL) was used, the sensitivity was 76%; however, when the manufacturer's device cut-off of 25 ng/mL was used, the sensitivity was higher at 87% (data not shown).

In the present study, variation across the devices was observed in their sensitivities of the ketamine, cocaine and THC tests. In an attempt to explain in part this variability, the concentrations detected in the false-negative samples were studied in order to investigate whether such cases were due to drug concentrations being close to the device cut-off. As shown in Table 9, for both ketamine and cocaine, the concentrations detected in the SalivaScreen® false-negative samples (max: 29 and 10 ng/mL for KET and COC, respectively) were indeed close to the device cut-offs (25 and 20 ng/mL respectively). On the other hand, KET concentrations as high as 2087 and 616 ng/mL were observed in the DrugWipe® 6S and Ora-Check® false-negative cases respectively; these concentrations were remarkably higher than the device cut-

offs. For cocaine, the concentrations observed in the DrugWipe® 6S false-negative cases (mean 68 ng/mL; max 186 ng/mL) were again considerably higher than the device cut-off (10 ng/mL). In the case of THC, concentrations as high as 1958 ng/mL were missed by both Ora-Check® and SalivaScreen®; the DrugWipe® 6S false-negative cases had comparatively lower concentrations (max 208 ng/mL), despite still being much higher than the device cut-off (20 ng/mL). To conclude, these results indicate that DrugWipe® 6S may be unable to identify ketamine and cocaine even at extremely high concentrations in oral fluid. Conversely, the same is also true for Ora-Check® and SalivaScreen® in detecting THC.

5. Conclusion

In assessing the performance of ROFT devices, several factors may be taken into consideration, including the user-friendliness, test success rate, and most importantly their sensitivity, specificity and accuracy. DrugWipe® 6S was the most user-friendly with the least requirement of oral fluid volume and the shortest analysis time. In terms of the test success rate, it also had the best performance, whilst SalivaScreen® performed moderately and Ora-Check® poorly in this regard.

Overall, the specificity and accuracy were satisfactory and met the DRUID recommendation of >80% for all three devices. The sensitivity, however, was found to vary. All devices performed poorly for THC. Ora-Check® had the poorest sensitivity among the 3 devices and did not achieve 80% in any of the tests. DrugWipe® 6S achieved >80% sensitivity in the methamphetamine and opiates tests but performed relatively poorly for ketamine and cocaine. Among the three devices, SalivaScreen® achieved >80% sensitivity in the methamphetamine and opiates tests, and was found to have the highest sensitivity for ketamine, cocaine and opiates.

In conclusion, whilst the specificity and accuracy were satisfactory, none of the devices achieved 80% sensitivity in all the tests. SalivaScreen® had on average the highest sensitivity among the devices.

6. References

[1] M.A. Huestis, A. Verstraete, T.C. Kwong, J. Morland, M.J. Vincent, R. de la Torre, Oral fluid testing: promises and pitfalls, Clin Chem 57 (2011) 805-810.

[2] W.M. Bosker, M.A. Huestis, Oral fluid testing for drugs of abuse, Clin Chem 55 (2009) 1910-1931.

[3] A.G. Verstraete, S. Legrand, Drug use, impaired driving and traffic accidents (2nd Ed), European Monitoring Centre for Drugs and Drug Addiction (EMCDDA), Lisbon, 2014.

[4] O.F. Wong, K.L. Tsui, T.S. Lam, N.N. Sze, S.C. Wong, F.L. Lau, S.H. Liu, Prevalence of drugged drivers among non-fatal driver casualties presenting to a trauma centre in Hong Kong, Hong Kong Med J 16 (2010) 246-251.

[5] A.S. Christophersen, J. Morland, K. Stewart, H. Gjerde, International trends in alcohol and drug use among vehicle drivers, Forensic Sci Rev 28 (2016) 37-66.

[6] Road Traffic Ordinance Chapter 374. https://www.elegislation.gov.hk/ (Accessed 23 March 2017).

[7] Roadside Testing Assessment (ROSITA) Final Report. http://www.transport-research.info/project/roadsite-testing-assessment (Accessed 26 April 2017).

[8] Rosita-2 Final Report. http://rosita.org/ (Accessed 27 April 2017).

[9] Analytical evaluation of oral fluid screening devices and preceding selection procedures. http://www.druid-project.eu/Druid/EN/deliverales-list/downloads/Deliverable_3_2_2.html?nn=613800 (Accessed 27 April 2017).

[10] S. Gentili, R. Solimini, R. Tittarelli, G. Mannocchi, F.P. Busardo, A Study on the Reliability of an On-Site Oral Fluid Drug Test in a Recreational Context, J Anal Methods Chem 2016 (2016) 1234581.

[11] A.S. Goessaert, K. Pil, J. Veramme, A. Verstraete, Analytical evaluation of a rapid on-site oral fluid drug test, Anal Bioanal Chem 396 (2010) 2461-2468.

[12] A. Pehrsson, T. Blencowe, K. Vimpari, A. Impinen, T. Gunnar, P. Lillsunde, Performance evaluation of the DrugWipe(R) 5/5+ on-site oral fluid screening device, Int J Legal Med 125 (2011) 675-683.

[13] S. Strano-Rossi, E. Castrignano, L. Anzillotti, G. Serpelloni, R. Mollica, F. Tagliaro, J.P. Pascali, D. di Stefano, R. Sgalla, M. Chiarotti, Evaluation of four oral fluid devices (DDS(R), Drugtest 5000(R), Drugwipe 5+(R) and RapidSTAT(R)) for on-site monitoring drugged driving in comparison with UHPLC-MS/MS analysis, Forensic Sci Int 221 (2012) 70-76.

[14] T.K. Tsui, A.S. Chan, C.W. Lo, A. Wong, R.C. Wong, C.S. Ho, Performance of a point-of-care device for oral fluid ketamine evaluated by a liquid chromatography-tandem mass spectrometry method, J. Anal. Toxicol. 36 (2012) 210-216.

[15] S.M. Wille, N. Samyn, M. Ramirez-Fernandez Mdel, G. De Boeck, Evaluation of on-site oral fluid screening using Drugwipe-5(+), RapidSTAT and Drug Test 5000 for the detection of drugs of abuse in drivers, Forensic Sci Int 198 (2010) 2-6.

[16] S.S. Rao, D.M. Wood, P.I. Dargan, Ketamine - Epidemiology of misuse and patterns of acute and chronic toxicity, in: D.T. Yew (Ed.), Ketamine: Use and Abuse, CRC Press, Boca Raton, FL, 2015, pp. 104-108.

[17] A.L. van Nuijs, A. Gheorghe, P.G. Jorens, K. Maudens, H. Neels, A. Covaci, Optimization, validation, and the application of liquid chromatography-tandem

mass spectrometry for the analysis of new drugs of abuse in wastewater, Drug Test Anal 6 (2014) 861-867.
[18] Clinical and Laboratory Standards Institute (CLSI), Liquid chromatography-mass spectrometry methods; approved guideline. C62-A, 2014.
[19] M. Concheiro, T.R. Gray, D.M. Shakleya, M.A. Huestis, High-throughput simultaneous analysis of buprenorphine, methadone, cocaine, opiates, nicotine, and metabolites in oral fluid by liquid chromatography tandem mass spectrometry, Anal Bioanal Chem 398 (2010) 915-924.
[20] European Medicines Agency Committee for Medicinal Products for Human Use (CHMP), Guideline on bioanalytical method validation, 2012.
[21] F.T. Peters, O.H. Drummer, F. Musshoff, Validation of new methods, Forensic Sci Int 165 (2007) 216-224.
[22] S. Strano-Rossi, L. Anzillotti, E. Castrignano, M. Felli, G. Serpelloni, R. Mollica, M. Chiarotti, UHPLC-ESI-MS/MS method for direct analysis of drugs of abuse in oral fluid for DUID assessment, Anal Bioanal Chem 401 (2011) 609-624.
[23] B.K. Matuszewski, M.L. Constanzer, C.M. Chavez-Eng, Strategies for the assessment of matrix effect in quantitative bioanalytical methods based on HPLC-MS/MS, Anal Chem 75 (2003) 3019-3030.
[24] A. Agresti, B.A. Coull, Approximate is better than "exact" for interval estimation of binomial proportions, Am Stat 52 (1998) 119-126.
[25] A. Valen, A.M. Leere Oiestad, D.H. Strand, R. Skari, T. Berg, Determination of 21 drugs in oral fluid using fully automated supported liquid extraction and UHPLC-MS/MS, Drug Test Anal (2016) doi: 10.1002/dta.2045.
[26] I. Zancanaro, R.P. Limberger, P.O. Bohel, M.K. dos Santos, R.B. De Boni, F. Pechansky, E.D. Caldas, Prescription and illicit psychoactive drugs in oral fluid--LC-MS/MS method development and analysis of samples from Brazilian drivers, Forensic Sci Int 223 (2012) 208-216.
[27] T. Blencowe, A. Pehrsson, P. Lillsunde, K. Vimpari, S. Houwing, B. Smink, R. Mathijssen, T. Van der Linden, S.A. Legrand, K. Pil, A. Verstraete, An analytical evaluation of eight on-site oral fluid drug screening devices using laboratory confirmation results from oral fluid, Forensic Sci Int 208 (2011) 173-179.
[28] H. Choi, S. Baeck, M. Jang, S. Lee, H. Chung, Simultaneous analysis of psychotropic phenylalkylamines in oral fluid by GC-MS with automated SPE and its application to legal cases, Forensic Sci Int 215 (2012) 81-87.
[29] K. Langel, C. Engblom, A. Pehrsson, T. Gunnar, K. Ariniemi, P. Lillsunde, Drug testing in oral fluid-evaluation of sample collection devices, J Anal Toxicol 32 (2008) 393-401.
[30] O. Quintela, D.J. Crouch, D.M. Andrenyak, Recovery of drugs of abuse from the Immunalysis Quantisal oral fluid collection device, J Anal Toxicol 30 (2006) 614-616.
[31] S. Vanstechelman, C. Isalberti, T. Van der Linden, K. Pil, S.A. Legrand, A.G. Verstraete, Analytical evaluation of four on-site oral fluid drug testing devices, J Anal Toxicol 36 (2012) 136-140.
[32] Central Registry of Drug Abuse, http://www.nd.gov.hk/en/statistics_list.htm (Accessed 2 May 2017).

[33] M. Tang, C.K. Ching, M.L. Tse, C. Ng, C. Lee, Y.K. Chong, W. Wong, T.W. Mak, Surveillance of emerging drugs of abuse in Hong Kong: validation of an analytical tool, Hong Kong Med J 21 (2015) 114-123.
[34] W.C. Cheng, K.M. Ng, K.K. Chan, V.K. Mok, B.K. Cheung, Roadside detection of impairment under the influence of ketamine--evaluation of ketamine impairment symptoms with reference to its concentration in oral fluid and urine, Forensic Sci Int 170 (2007) 51-58.