Project No. TREN-05-FP6TR-S07.61320-518404-DRUID

DRUID

Driving under the Influence of Drugs, Alcohol and Medicines

Integrated Project
1.6. Sustainable Development, Global Change and Ecosystem
1.6.2: Sustainable Surface Transport

6th Framework Programme
D 3.2.2

Analytical evaluation of oral fluid screening devices and preceding selection procedures

Due date of deliverable: (14.02.2010)
Actual submission date: (30.03.2010)

Start date of project: 15.10.2006
Duration: 48 months
Organisation name of lead contractor for this deliverable: DTU
Revision 2.0

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

D 3.2.2 Analytical evaluation of oral fluid screening devices and preceding selection procedures

Editors :

Tom Blencowe, Anna Pehrsson and Pirjo Lillsunde

Alcohol and Drug Analytics Unit, National Institute for Health and Welfare (THL), Finland

Work package leader

René Mathijssen

SWOV Institute for Road Safety Research, the Netherlands

Task leader

Inger Marie Bernhoft

Department of Transport, Technical University of Denmark (DTU), Denmark

Participating partners

Anna Pehrsson, Tom Blencowe, Charlotta Engblom, Kaarina Langel and Pirjo Lillsunde

Alcohol and Drug Analytics Unit, National Institute for Health and Welfare (THL), Finland

Sjoerd Houwing[1], Beitske Smink[2] and René Mathijssen[1]

[1] SWOV Institute for Road Safety Research, the Netherlands

[2]Netherlands Forensic Institute (NFI), the Netherlands

Kristof Pil, Trudy Vander Linden, An-Sofie Goessaert, Jolien Veramme, Sylvie Van Stechelman, Sara-Ann Legrand, Alain Verstraete

Department of clinical chemistry, microbiology and immunology, Ghent University, Belgium

Table of contents

Index of tables

Executive Summary

- This analytical evaluation of oral fluid screening devices and preceding selection procedures was carried out as an integral part of the DRUID project (Work package 3, Task 2). The duration of the evaluation was from October 2007 to December 2009.
- The study was carried out by the Faculty of medicine and health sciences, Department of clinical chemistry, microbiology and immunology, Ghent University (UGent) in Belgium, the Alcohol and Drugs Analytics Unit, National Institute for Health and Welfare (THL) in Finland and the SWOV Institute for Road Safety Research in the Netherlands. The Department of Transport, Technical University of Denmark (DTU) was responsible for leading the task due to its connection to the road side survey (Work package 2, Task 2.2a1) for which DTU was Work package leader. THL was responsible for finalising the deliverable.
- Eight on-site tests were evaluated: BIOSENS Dynamic (Biosensor Applications Sweden AB), Cozart DDS (Cozart Bioscience Ltd.), DrugWipe 5^{+} (Securetec Detections-Systeme AG), Dräger DrugTest 5000 (Dräger Safety), OraLab6 (Varian), OrAlert (Innovacon), Oratect III (Branan Medical Corporation) and Rapid STAT (Mavand Solutions GmBH).
- Rapid STAT was tested in all three countries and DrugTest 5000 in Belgium and the Netherlands. All other devices were tested in only one country.
- Tested substance classes were amphetamine(s), methamphetamine, MDMA, cannabis, cocaine, opiates, benzodiazepines and PCP.
- A checklist for clinical signs of impairment (CSI) was also evaluated in order to see if visible signs of impairment can be used as preceding selection criteria for performing an on-site test.
- The checklist was based on several existing checklists, e.g. one developed for the German police and previously used in the European IMMORTAL (Impaired Motorists, Methods Of Roadside Testing and Assessment for Licensing) project.
- Study populations consisted of randomly selected drivers from the roadside survey for DRUID (Work package 2, Task 2.2a1), drivers suspected of driving under the influence of drugs, patients of treatment centres and rehabilitation clinics and customers of coffeeshops.
- Oral fluid was collected as the reference sample. For some cases, in the Netherlands, whole blood samples were also collected.
- The performance of the tests was assessed based on sensitivity, specificity, accuracy, positive predictive value and negative predictive value for the individual substance tests of the device. These were assessed based on both DRUID and manufacturer cut-offs.
- Sensitivity, specificity and accuracy performance values of 80% or more were set as a desirable target value.
- The analytical evaluation of the amphetamine test showed sensitivity varying from 0% to 87 %. Specificity values were from 91% to 100% and accuracy values from 84% to 98%.
- For cannabis tests, sensitivities ranged from 11% to 59%. Specificities were between 90% and 100% and accuracies from 41% to 82%.
- Cocaine tests scored sensitivities of between 13% and 50%, specificities of 99% to 100% and accuracies from 86% to 100%.
- Sensitivities of opiate tests ranged from 69% to 90%. Specificities were between 81% and 100% and accuracies between 75% and 99%.
- Benzodiazepine tests had sensitivities from 48% to 67%. Specificities were from 94% to 100% and accuracies from 77% to 100%.
- Not enough positive cases were gathered to successfully evaluate any of the methamphetamine, MDMA or PCP tests for the devices in which these were included.

- None of the tests reached the target value of 80% for sensitivity, specificity and accuracy for all the separate tests they comprised.
- An overall evaluation, wherein any positive drug screening result was viewed as valid providing that the confirmation sample contained one of the DRUID substances analysed, was performed as a measure of the usefulness of the devices in police controls.
- Three of the devices performed at >80% for sensitivity, specificity and accuracy in the overall evaluation.
- Prevalence of drugs in the study population needs to be considered when assessing the evaluation results. In addition, the type and prevalence of drugs within the population for which the device is intended to be used needs to be taken into account when considering the suitability of the device based on the results presented in this report.
- Some device failures were noted in the study. For one of the tests, 15 individual tests (12%) failed. For other tests, 5 or less tests failed. In the Netherlands the evaluation of Oratect III was stopped because the devices frequently failed to collect oral fluid in a sufficiently short time..
- All countries took their own approach to the evaluation of the checklist for clinical signs of impairment. The results of the evaluations were not very promising. The checklist scored a low sensitivity value (Dutch study), low correlation of symptoms and actual presence of drugs (Belgian study) or there were difficulties in correlating the symptoms to actual drug use due to the insufficient data collection (Finnish study).

List of abbreviations

ACN	acetonitrile
ATS	amphetamine-type stimulant drugs
AMP	amphetamine
B	blood
BuAc	butyl acetate
c	concentration
CSI	clinical signs of impairment
DCM	dichloromethane
DUI	driving under the influence
EI	electron impact ionisation
EtAc	ethyl acetate
FN	false negative
FP	false positive
GC-MS	gas chromatography mass spectrometry
LLE	liquid-liquid extraction
LLOQ	lower limit of quantification
LOQ	limit of quantification
6-MAM	6-monoacetylmorphine
MAMP	methamphetamine
MDA	methylenedioxyamphetamine
MDEA	methylenedioxyethylamphetamine
MDMA	methylenedioxymethamphetamine
MRM	multiple-reaction monitoring
MSTFA	N-methyl-N-trifluoroacetamide
MTBSTFA	N-methyl-N-(tert-butyldimethylsilyl)trifluoroacetamide
N_2	nitrogen
NH_3	ammonia
NICI	negative chemical ionisation
NPV	negative predictive value
OF	oral fluid
PCP	phencyclidine
PPV	positive predictive value
i-Pr	isopropanol
QCM	quartz crystal microbalance
SPE	solid phase extraction
Δ^9-THC	Δ^9-tetrahydrocannabinol

THCC	11-nor-Δ^9-tetrahydrocannabinol-9-carboxylic acid
ULOQ	upper limit of quantification
WB	whole blood
TN	true negative
TP	true positive
UPLC-MS/MS	ultra pressure liquid chromatograph mass spectrometry

1. Introduction

1.1. DRUID Project

The European Integrated Project DRUID (Driving Under the Influence of Drugs, Alcohol and Medicines) is a part of the 6th Framework Programme, the European Community Framework Programme for Research, Technological Development and Demonstration. The DRUID project focuses on the improvement of road safety related to the problem of alcohol, drugs and medicines used or abused by drivers of vehicles in the road transport system. The objective of DRUID is to give scientific support to the EU transport policy by providing a solid basis to generate harmonised, EU-wide regulations for driving under the influence of alcohol, drugs and medicine. This study is a part of Work Package 3, Enforcement, of the DRUID project.

1.2. Work package 3 - Enforcement

The objective of this work package is to conduct a large scale evaluation of on-site screening devices for impairing psychoactive substances other than alcohol in drivers. The preceding selection criteria, based on signs of impairment, for performing on-site screening devices will also be evaluated. The selection criteria should allow the police to check for suspicious signs leading to a conclusion of possible drug useage. The results of work package 3 should improve the possibilities of detecting drug driving in Europe, providing a good grounding for harmonising the European police requirements for on-site drug screening. Work package 3 comprises a practical evaluation of a number of on-site drug screening devices, an analytical evaluation of devices that were deemed promising in the practical evaluation and a cost-benefit analysis of the use of such devices.

1.3. Task 3.2

Task 3.2 of work package 3 is the analytical evaluation of oral fluid screening devices and preceding selection procedures. The study was carried out in the Netherlands, Belgium and Finland. The analytical evaluation assesses the reliability and accuracy of results from the screening devices against the results of a confirmation analysis. For this purpose, oral fluid (OF) samples, and in the Netherlands in some cases blood samples, were collected from the participants. The devices were to be used to screen both non-suspect and suspect drivers, thus allowing sufficient collection of both positive and negative samples. However, initial screenings with quick tester devices, during epidemiological roadside survey sessions (Task 2.2a of DRUID) revealed that, as might be expected, the number of suspected positive cases from this study population was too low to enable evaluation of the devices with a high proportion of positive cases. Therefore each partner country decided to use study populations with an expected higher prevalence of illicit drug or medication use, as well as the roadside survey sampling. In the Netherlands this was achieved by additional testing in a coffeeshop. In Belgium the majority of samples were collected in centres for treatment of drug addiction. In Finland the police asked apprehended suspected driving under the influence (DUI) cases to participate in the study and, in addition, a small number of patients from a cooperating rehabilitation clinic for drug addicts were recruited to the study. The Finnish study also presents data from Finnish police use of on-site screening in traffic, using blood as the confirmation sample. For evaluation of preceding selection criteria for performing on-site drug screening a number of observable signs and symptoms, as well as self-reportage of drug or medicine use, were chosen. The expected outcome of the task is to detect the best practices for police controls and provide an evaluation of on-site drug screening devices leading to recommendations for roadside procedures for drivers suspected of driving while impaired by psychoactive substances. The selection of devices for Task 3.2 was based on the outcomes of the preceding study Task 3.1, Practical evaluation of on-site screening devices, by the Dutch National Police Agency (KLPD).

1.4. Matrices for drug screening and confirmation analysis

In the past, selective police enforcement on suspected drugged drivers has been performed in a limited number of EU member states, mostly using urine-based screening devices. The experiences of the police have shown that performing on-site urine screening is very complicated at the roadside, often potentially intrusive or vulnerable to adulteration. In addition, a positive urine test may indicate exposure to a drug several days previously since the analytes accumulate in urine. Detection of drugs in oral fluid indicates more recent drug use and this matrix is suitable for inexpensive, non-invasive and easy-to-use diagnostic aids for detection of illicit use of drugs (1). On-site screening for drugs of abuse in oral fluid provides a relatively quick and increasingly effective means of detecting if a motorist has consumed drugs or medicines, which are of concern to traffic safety. Nevertheless a positive OF on-site screening result can only be interpreted as a possible indication of impairment rather than definitive proof.

There are significant correlations between oral fluid concentrations of drugs of abuse and behavioral and physiological effects (2, 3). The state of impairment of a drug user is related to the physiologically active fraction of the substance in question. In oral fluid only the free, physiologically active, component of a drug is found, whereas in whole blood many drugs are highly bound to blood proteins and this fraction of the drug is physiologically inactive (4, 5). For many drugs transfer of the free component from blood to oral fluid occurs via simple passive diffusion (1), although transfer is affected by various factors such as oral fluid pH and the characteristics of the drug (e.g. acidic or basic, lipid solubility). Therefore the presence of a substance in OF corresponds to its presence in blood and therefore OF is a potential alternative matrix as a confirmation sample for toxicological analysis. However, knowledge concerning the OF/blood ratios for drugs remains incomplete and there has been very little comparison of blood and OF analysis. Nonetheless in the Rosita-2 project, which evaluated usability and analytical reliability of on-site OF drug testing devices in 2003-2005, a comparison of results from laboratory analyses on OF and blood concluded that OF is a good screening fluid for the presence or absence of amphetamines, cannabis, cocaine and opiates in blood (6).

Since the concentrations of individual drugs found in OF are directly comparable with the results for the on-site tests, it was used as the primary confirmation matrix in this study. However, it should be remembered that in the majority of countries where legislation exists for DUI of psychoactive drugs the matrix used for confirmation analysis is blood rather than OF. In the Netherlands some blood samples were collected, in addition to OF, to enable further comparison in this study. On-site test and blood confirmation analysis results from police DUI enforcement are discussed further in the Finnish country report, however there are no OF analysis data for these cases so direct comparison is not possible.

1.5. On-site tests in police procedure

Intervention for traffic safety is based on legislation and enforcement of driving under the influence of drugs (DUID) based on detection. The form of detection that may be used is specific to the type of impairment factor and the type of legislation in consideration. With regard to this, several EU countries have introduced *per se* legislation for drug driving, making it sufficient to prove use of a substance rather than actual impairment. An indication of drug use can be achieved with an on-site drug screening test, however the device should be easy to use, rapid and reliable. Such an on-site drug screening test is a requirement that police officers involved in road safety have expressed for a number of years. It should, however, be remembered that even with efficient on-site testing a confirmation analysis of a body fluid sample from the driver remains necessary.

In addition, several studies have suggested that random roadside drug screening can act as an effective deterrent to DUID. Indeed random screening of drivers in the state of Victoria, Australia has already proven to be a useful means for providing a deterrent to drug driving. In December 2004 police in Victoria started a random drug testing programme for drivers modelled on more than 30 years of random testing for alcohol in Victoria and Australia. The observed prevalence in screened drivers for the drugs which are currently included in the legislation of Victoria (MDMA, methamphetamine and cannabis) have all decreased (7). Furthermore, awareness of random oral fluid testing has increased from 78 % to 92 % of drivers in the initial period following the introduction of testing (6). By using two on-site screening devices in combination and a confirmation analysis, a false positive rate of about 1 % over 5 years was also attained. For this purpose a false positive was

defined as both devices giving positive results which are not confirmed in the laboratory toxicological analysis (7).

Previously selective police enforcement on drugged driving has been performed in a limited number of EU member states, mostly with urine based screening devices. The experiences of the police have shown that performing a urine on-site screening procedure is very complicated at the road side. However, it is already apparent that with new national legislations continuously arising, on-site screening devices are being used more and more. Germany was the first European country to introduce zero-tolerance legislation for drugs, prohibiting driving under the influence of the drugs cannabis, cocaine, heroin, morphine, amphetamine and the designer drugs MDMA and MDEA in 1998 (Road Traffic Code, § 24 a StVG). In most of the Federal States of Germany roadside drug tests have been introduced on a routine basis, these can take the form of a urine, sweat or oral fluid screening device (8). In 2006 the French Ministry of the Interior began a tendering phase to test potential on-site oral fluid screening devices. This resulted in the reward of the French national tender in July 2008 and roadside controls were started in August 2008.

The legislation and police controls for enforcement of drugged driving vary for each of the partner countries involved in Task 3.2. In Finland the zero tolerance law was implemented in February 2003, alongside the existing impairment legislation which provides for prescription pharmaceuticals as well as non-prescription medications. In the Finnish legislation on-site testing devices for drugs have the same position as for breath testing for alcohol and since December 2005 the police have been routinely using on-site oral fluid screening devices in traffic control. The decision to use the devices came at the same time that the Rosita-2 evaluation of available on-site oral fluid screening devices was nearing completion. The Finnish traffic police were closely involved in the Rosita-2 project and in the experience of most of the police officers the devices were a valuable tool for helping with identification and confirmation of initial suspicion of drug use (6). Oral fluid tests are usually performed either in the process of a random control for alcohol, where the performing officer suspects impairment but the breath test result is negative, or in cases of suspicious or impaired driving where there is no evidence of alcohol impairment. A screening test for alcohol is normally performed first since this is cheaper and less time consuming.

In Belgium legislation relating to DUID of certain illicit drugs was adopted in 1999, which included analytical legal limits for cannabis, cocaine, opiates and amphetamines. This regulation was a zero tolerance-type law, but blood sampling and analysis was only allowed if signs of impairment were obvious, and if a roadside urine test was positive for one or more of those substances (8). In addition, the high number of false positives from urine screening (15%) was expensive (7). It is stipulated that driving while being impaired is forbidden in Article 35 of the Belgian Traffic Law, however until now very few subjects have been convicted under Article 35 (7). The rationale for change was comprehended both scienticially and by the police and judicial authorities. In Belgium the police were also active participants in the Rosita-2 project and it was recognized that oral fluid offered a non-invasive means of screening under direct supervision. The police were very willing to participate in the development of a suitable on-site test. It was also perceived that the use of oral fluid would facilitate the legal procedure (6). In June 2009 a proposal to modify the Traffic Act related to driving under the influence of psychoactive substances was brought to parliament. In this proposal random oral fluid drug testing and analysis of an oral fluid sample for evidence is described. The new procedure allows the police officer to make the screening test and, if it is positive, collect an oral fluid sample for confirmation (8). This law has been passed and is expected to be implemented by October 2010.

In the Netherlands an impairment-oriented approach to DUID is pursued. Driving under the influence of any substance that effects driving behaviour (i.e. illicit drugs and certain prescription drugs) is punishable. Impairment of driving performance, as well as the presence of significant concentrations of drugs in a driver's blood, must be demonstrated. If the result of an alcohol screening breath test is less than 350 µg/l, but there are indications of impairment, the policy of the public prosecutor is that the police officer can try to prove the abuse of psychoactive substances. Nevertheless, no roadside devices for the detection of the abuse of illegal drugs in road traffic are in use in the Netherlands. The Dutch Traffic Act does not contain any regulation concerning drug screening devices or their applications (8).

2. Description of the devices

Altogether eight different on-site oral fluid screening devices were tested in the study. Six of the devices were evaluated in the Netherlands, Belgium or Finland, one device was evaluated in the Netherlands and Belgium and one device was evaluated in all three countries. The manufacturers of the devices provided the tests for free. More specific information on the procedure for operating the devices can be found from the report of DRUID Task 3.1, Practical evaluation of on-site screening devices, by Cor Kuijiten (8).

2.1. BIOSENS® Dynamic (Biosensor Applications Sweden AB)

The BIOSENS Dynamic oral fluid test consists of a collector and a reader. The detection system is based on a piezoelectric quartz crystal microbalance (QCM) technology using a monoclonal antibody as the specific element. At the time of writing the cut-off values for this device were undetermined. The detected drugs are: methamphetamine and MDMA, cocaine, opiates, benzodiazepines and cannabis (Δ^9-THC).

Figure 1. BIOSENS Dynamic reader and collection device.

Samples for testing are collected with a collection device by wiping the tongue of the subject. The sample is then processed in the BIOSENS oral screening system. A complete analysis takes approximately two to three minutes including sample acquisition.

The operation temperature range of the device is +15°C - +40°C and storage/transport temperature 0°C - +50°C.

2.2. Cozart® DDS 806 (Cozart Bioscience Ltd.)

The Cozart DDS system comprises a collector swab, a buffer bottle, a disposable test cartridge, a handheld instrument for result interpretation and a printer for permanent recording of test results.

Figure 2. Cozart DDS system.

The DDS 806 device is a six-drug test kit for cannabis, cocaine, opiates, methamphetamine, amphetamine and benzodiazepines. The collector stick has a sample presence indicator to ensure collection of adequate volume of oral fluid. The cut-offs for detected drug groups are listed in Table 1.

Table 1. Detected analytes and their cut-off values for Cozart DDS 806.

Detected drug group (target compound)	Cut-off / ng/ml
Amphetamine	50
Methamphetamine	50
Cocaine (benzoylecgonine)	30
Opiates (morphine)	30
Cannabis (Δ^9-THC)	31*
Benzodiazepines (temazepam)	20

* Validation for cannabis cut-off performed using real patient samples.

When testing, the collector swab is swabbed around the gums, tongue and inside cheek until the sample presence indicator turns completely blue. The swab is placed into the buffer bottle and the contents of the bottle are mixed. Four drops of fluid from the dropper are applied across the sample well of the test cartridge. As soon as fluid appears on each of the four white cartridge membrane strips (between 2 to 30 seconds), the cartridge is inserted into the DDS instrument and a new test is initiated. Once the test has been completed, the results are displayed on the DDS instrument. The results are also is printed if the printer is connected. The manufacturer states that 2 drugs can be tested in 90 seconds, and 5/6 drug classes in 5 minutes.

The device has a shelf life of 9 months when stored at +15°C - +25°C.

2.3. DrugWipe® 5+ (Securetec Detections-Systeme AG)

The DrugWipe 5+ is an enhanced version of the DrugWipe 5 test. The test detects multiple substances within one oral fluid sample, the drug groups detected are opiates, cocaine, amphetamines including methamphetamine and MDMA (ecstasy), and cannabis. Cut-off values for different drug groups are presented in Table 2.

Figure 3. DrugWipe 5+.

The device consists of an oral fluid collector, a detection element and an integrated liquid ampoule. Results are indicated with red lines. A red line indicates a positive result. Red control lines indicate a successful test.

Table 2. Detected analytes and their cut-off values for DrugWipe 5+.

Detected drug group (target compound)	Cut-off / ng/ml
Amphetamines (D-amphetamine)	50
Methamphetamine (D-methamphetamine)	25
MDMA	25
Cocaine (benzoylecgonine)	30
Opiates (codeine)	10
Cannabis (Δ^9-THC)	30

To carry out the test, the oral fluid collector is separated from the test body. The tongue or the cheek of the tested person is wiped and the collector is then attached to the test body. The ampoule is pressed so that it opens and the buffer solution flows to the test strips. The results can be read when

control lines appear on both strips. The total time needed for testing is 3-10 minutes according to the manufacturer.

The device is stored at +15°C - +25°C.

2.4. Dräger DrugTest® 5000 (Dräger Safety)

The Dräger DrugTest 5000 test system comprises the Dräger DrugTest 5000 Analyzer and a test kit. The test kit consists of a test cassette with an oral fluid collector. The Dräger DrugTest system is designed for qualitative measurement of specific substances and their metabolites. Substances detected are amphetamines, methamphetamines, opiates, cocaine, benzodiazepines and cannabis. The cut-off values for detected drug classes are listed in Table 3.

Figure 4. Dräger DrugTest 5000 test system: analyzer and the cassettes.

Table 3. Detected analytes and their cut-off values for Dräger DrugTest 5000.

Detected drug group (target compound)	Cut-off / ng/ml
Amphetamine	50
Methamphetamine	35
Cocaine (cocaine)	20
Opiates (morphine)	20
Cannabis (Δ^9-THC)	5 (25)*
Benzodiazepines (diazepam)	15

*Cut-off for device evaluated in the Netherlands, discontinued in April 2009

Oral fluid is collected by using the oral fluid collector on the test cassette. The built-in indicator turns blue when collection is ready. Then the test cassette and the cartridge are placed into the analyzer. Results are shown within a few minutes.

The test kit must be stored at +4°C - +30°C in its original foil pouch. The test kit must be used immediately after opening the pouch. The operation temperature range of the analyzer is +5°C - +40°C and storage/transport temperature -20°C - +60°C.

2.5. OraLab6 (Varian)

OraLab6 is an oral fluid test that detects six drug classes simultaneously. Drug classes tested are amphetamines, methamphetamines, cocaine, opiates, phencyclidine and cannabis. The cut-offs for tested drugs are shown in Table 4.

The OraLab6 test consists of a collection stick and an expresser vial. With this device, an oral fluid sample is collected simultaneously during the testing. A red line adjacent the given drug name indicates a negative result. A red 'test valid' line indicates a successful test.

Figure 5. OraLab6 collection stick and expresser vial.

Table 4. Detected drug groups and their cut-off values for the OraLab6.

Detected drug group (target compound)	Cut-off / ng/ml
Amphetamines	50
Methamphetamines	50
Cocaine*	20
Opiates (morphine)	40
Phencyclidine	10
Cannabis (Δ^9-THC)	50

*Target compound not specified

Testing starts with oral fluid collection. The collector stick is kept in the mouth of the tested person for 3 minutes until the foam is thoroughly soaked. The collector is then placed foam-first into the expresser. The cap of the expresser is used to push the collector all the way down into the expresser and the cap is twisted tightly into place. The oral fluid collected will flow to the vial and results can be read from between 10 to 15 minutes.

The device should be stored at room temperature (+15°C - +30°C) and used immediately after opening.

2.6. OrAlert (Innovacon)

The OrAlert test detects six drugs simultaneously. The drugs tested with the device used in this study are amphetamines, methamphetamines (including MDMA), cocaine, opiates, cannabis and phencyclidine (PCP). The test consists of a sucrose and citric acid coated collector stick and of a test device. An oral fluid sample is collected during the testing and can be sent to laboratory for confirmation. Red lines indicate a negative result for the substance in question. Red control lines indicate a successful test. The cut-offs for different substances are listed in Table 5.

Figure 6. OrAlert test device and collector stick.

Table 5. Detected drug groups and respective cut-off values for the OrAlert device.

Detected drug group (target compound)	Cut-off / ng/ml
Amphetamines	50
Methamphetamines	50
MDMA	50
Cocaine (benzoylecgonine)	20
Opiates (morphine)	40
Phencyclidine	10
Cannabis (Δ^9-THC)	100

Oral fluid is collected with the sponge of the collector stick. When the sponge softens slightly, the tested person is instructed to gently press the sponge between the tongue and teeth to ensure saturation of the collector. The sample is collected for 3 minutes. The collector stick is inserted to the test device. After one minute the collection chamber is turned counterclockwise. The results can be read after a following 9 minutes. The total test time is approximately 10 minutes.

The OrAlert device should be stored at +2°C - +30°C.

2.7. Oratect® III (Branan Medical Corporation)

The Oratect III is a one-step lateral flow immunoassay device for qualitative detection of drugs in oral fluid. The device evaluated in this study detects six drug groups: amphetamines, cocaine, cannabis, methamphetamines including MDMA, opiates and benzodiazepines. The cut-offs for the substances detected are shown in Table 6.

Figure 7. Oratect III on-site screening device

In this on-site test, sample collection and on-site testing are integrated in one device. Red lines indicate a negative result. The collection pad can be sent to laboratory for confirmation analysis. For this, the pad is detached from the collector and put to the buffer vial included in the kit.

Table 6. Detected drug groups and their cut-off values for the Oratect III.

Detected drug group (target compound)	Cut-off / ng/ml
Amphetamines	25
Methamphetamines	25
MDMA	25
Cocaine (Cocaine)	20
Opiates (Morphine)	10
Cannabis (Δ^9-THC)	40
Benzodiazepines	5

Prior to carrying out the test, the cap is removed from the device. The collection pad is rubbed inside the mouth and then placed underneath the tongue to collect oral fluid. Once a blue line indicator for the sufficient collection of oral fluid appears the collection pad is removed from the mouth and the device is recapped. The test is layed on a flat surface and results can be read within 5 to 30 minutes, as red lines on the test strip. The red control lines indicate a successful test.

The Oratect III should be stored at room temperature.

2.8. Rapid STAT® (MAVAND Solutions GmBH)

Rapid STAT is an on-site test for the simultaneous detection of two to six drugs of abuse or medicinal drugs in oral fluid specimens. It is a lateral flow immunoassay in which each analyte is represented by a separate line in the test window of the device. For the device evaluated in this study, the compounds detected and the cut-offs are listed in Table 7.

Figure 8. Rapid STAT mobile reader and test device: collection stick, buffer bottle and test panel.

The device consists of a collector stick with an aroma field, a buffer bottle and the test panel. Results are indicated with red lines. A line indicates a negative result. The appearance of the control line indicates a successful test. The test results can be read either straight from the test device or by inserting the device into a mobile reader (Reader 600). The mobile reader determines the device results by camera measurement and determines whether the test lines are positive or negative (Figure 9).

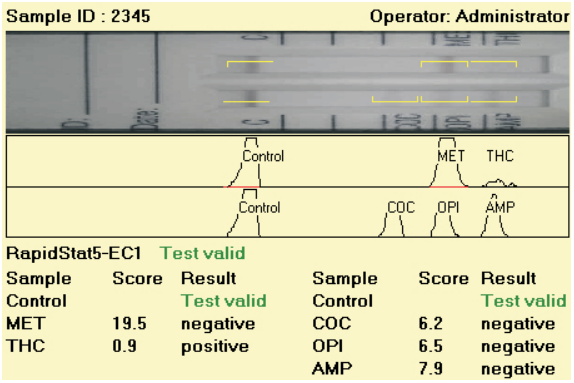

Figure 9. Example of mobile reader results from Reader 600.

Table 7. Detected drug groups and their cut-off values for the Rapid STAT device.

Detected drug group (target compound)	Cut-off / ng/ml
Amphetamine (D-amphetamine)	25
Benzodiazepines (oxazepam)	25
Cocaine (benzoylecgonine)	12
Methamphetamine	25
MDMA	50
Opiates (morphine)	25
Cannabis (Δ^9-THC)	15

The aroma field is used in order to increase salivation. Oral fluid is then collected by rotary movements of the microfiber collector stick. The collector stick is then put into the buffer bottle and agitated before removal. The buffer fluid mixture is then pipetted to each well of the test device. The lid is closed to the first position and left for 4 minutes. The test is then started by pressing down the lid completely so that the buffer flows to the test strips. As soon as all lines have formed the test results may be interpreted. If all lines have formed the test is interpreted as negative. The results should be read within 8 minutes of the buffer flowing. The total time needed for testing is 7-12 minutes according to the manufacturer.

The device should be stored at room temperature (+2°C - +30°C). The test must always be allowed to warm up to room temperature before any testing is conducted.

3. Description of the checklist for clinical signs of impairment

Together with the analytical evaluation of on-site oral fluid screening devices, an integral part of Task 3.2 is to assess the preceding selection criteria for performing a test. The police will normally only perform on-site analytical drug screening when a driver is suspected of being drug impaired, since the test procedure is time consuming and the screening devices are relatively expensive. Therefore an observational method, or checklist for clinical signs of impairment (CSI), for detecting drug-impaired drivers is necessary. Identifying external symptoms of drug use and documenting them to a field sobriety sheet is already a common practice among police forces.

For this study the method for identifying likely impairment consisted of a checklist for probable indicators of drug use supplemented with questions regarding self-reported use of medicines and illicit drugs. All the symptoms could be observed without performing complicated tests. The person supervising the screening test (e.g. a police officer in the case of a suspect driver) was also asked to give a final opinion on whether the driver had used impairing drugs other than alcohol. Based on the reliability of the method, the intended outcome of the CSI checklist study will be recommendations for roadside selection procedures of drivers suspected of driving while impaired by psychoactive substances.

The checklist was based on several existing checklists, e.g. one developed for the German police and previously used in the European IMMORTAL (Impaired Motorists, Methods Of Roadside Testing and Assessment for Licensing) project. The substance groups for which each of the symptoms is expected to serve as indicators are identified in Table 8.

Table 8. Drug categories and related symptoms.

	Substance group			
Symptom	**Opiates**	**Amphetamines**	**Cannabis**	**Cocaine**
unsteady on one's feet, swaggering	X			
uncoordinated movements	X			
drowsy, sleepy	X			
euphoria	X	X	X	X
not understanding instructions	X			
incoherent speech	X			
chattering		X	X	X
slurred speech			X	
low, rasping voice	X			
scratching one's face	X			
trembling		X	X	X
shaking leg		X	X	X
excited, aggressive behaviour		X		X
bloodshot eyes		X	X	X
red nostrils				X
trembling eyelids		X		X
sniffing				X
undue perspiring	X			
swallowing	X	X	X	X
smell of hash			X	
pinpoint pupils (<3.0mm)	X			
dilated pupils (>6.5mm)		X	(sometimes)	X
slowed pupil reaction to light		X	X	X
Nystagymus test – jerking pupil movement	X		X	

4. Description of the Analytical evaluation

Oral fluid samples were analysed in each country either with LC-MS/MS or GC-MS. The analysis methods are presented in more detail in the country reports.

4.1. Substances analysed an cut-offs

All oral fluid samples collected for the study were analysed for a list of substances. The substances analysed are presented in Table 9 and Table 10. Table 9 consists of the relevant core substances that every country must analyse. In addition to these, countries were to pick a number of additional substances in order to adapt to the national situation in drug use. These additional substances are listed in Table 10. It should be noted that the countries involved had their analytical cut-offs for some substances even lower than the ones listed in the tables. The cut-offs used were set for a separate epidemiological study within the DRUID project.

Table 9. List of relevant core substances and their DRUID cut-offs in oral fluid and whole blood.

Substance	Oral fluid / ng/ml	Whole blood / ng/ml
6-Acetylmorphine	5	10
Alprazolam	1	10
Amphetamine	25	20
Benzoylecgonine	10	50
Clonazepam	1	10
Cocaine	10	10
Codeine	20	10
Diazepam	5	20
Flunitrazepam	1	2
Lorazepam	1	10
MDA	25	20
MDEA	25	20
MDMA	25	20
Methamphetamine	25	20
Morphine	20	10
Nordiazepam	1	20
Oxazepam	5	50
THC	1	1
THCC	-*	5

* Substance not measured in OF

Table 10. List of relevant extra substances and their DRUID cut-offs in oral fluid and whole blood.

Substance	Oral fluid / ng/ml	Whole blood / ng/ml	Country
α-OH-alprazolam	1	1	FI
Aminoclonazepam	1	10	BE
Aminoflunitrazepam	1	2	BE and NL
Bromazepam	5	20	FI, BE and NL
Chlordiazepoxide	10	20	FI and NL
Clobazam	5	5	NL
Desalkylflurazepam	2	2	NL
Flurazepam	1	2	NL
Lormetazepam	1	1	NL
Midazolam	2	10	FI and NL
Nitrazepam	2	1	FI and NL
Temazepam	10	20	FI and NL
Triazolam	1	1	NL

4.2. Oral fluid collection

For oral fluid collection, Belgium and Finland used the Saliva•Sampler (StatSure Diagnostic Systems, Inc., Framingham, MA, USA) device. The device was chosen for sample collection based on a study made by THL (9). The device consists of an absorptive cellulose pad with a volume adequacy indicator and a plastic tube containing buffer solution. The window of the stem turns blue when 1 ml of OF is collected.

Figure 10. The Saliva•Sampler oral fluid collection device.

When collecting the sample, oral fluid is first allowed to gather in the mouth. The collection pad is placed under the tongue and removed when the indicator window has turned completely blue. The pad is then placed into the collection tube. In the laboratory, the collection pad is disconnected from the stem and dropped to the bottom of the tube, and a filter is inserted into the tube to recover the oral fluid-buffer solution.

The Netherlands collected oral fluid into polypropylene containers (Deltalab, S.L.U., Barcelona, Spain), which did not contain any buffer solution.

Figure 11. Plastic spit cup used in the Netherlands.

4.3. Adjustment of analyte concentration for volume of oral fluid sample

The StatSure Saliva Sampler collection device is designed to collect 1 ml of oral fluid device and the device also includes 1 ml of a buffer solution. Therefore the concentrations of substances found in the oral fluid sample are determined using a calibration curve, which is based on the assumption that the oral fluid to buffer ratio is 1:1. The standards used for the calibration curve are adjusted accordingly.

In reality, the amount of oral fluid collected in the device will vary to some extent. This can be because of intrinsic variation in the collection volume when the indicator turns blue or, for example, if the subject is unable to provide a full 1 ml of oral fluid due to a dry mouth.

Thus, it is necessary to adjust the concentration of analyte determined in analysis according to the actual volume of oral fluid collected. This is calculated using the following formula:

$$C_{corrected} = \frac{C_{uncorrected} \times (1 + w - \overline{w})}{2 \times (w - \overline{w})}$$

where:

$\overline{w}$ = average weight of empty StatSure device

w = weight of sample and StatSure device

$C_{uncorrected}$ = uncorrected concentration of analyte

$C_{corrected}$ = concentration of analyte corrected for volume of oral fluid collected

The density of the buffer – oral fluid mixture is assumed to be 1 g/ml for this adjustment.

4.4. Interpretation of the results

4.4.1. Classification

The evaluation of the results is based on classification into the following categories

- True positive (TP): number of cases with a positive test result and a positive confirmation analysis result in OF
- True negative (TN): number of cases with a negative test result and a negative confirmation analysis result in OF
- False positive (FP): number of cases with a positive test result and a negative confirmation analysis result in OF
- False negative (FN): number of cases with a negative test result and a positive confirmation analysis result in OF

Using these classifications, several parameters for the evaluation can be calculated.

4.4.2. Sensitivity, specificity, accuracy and prevalence

Sensitivity is the proportion of positive cases that are correctly identified by the test (equation 1).

$$Sensitivity = \frac{TP}{TP + FN} \quad (1)$$

Specificity is the proportion of negative cases that are correctly identified by the test (equation 2).

$$Specificity = \frac{TN}{TN + FP} \quad (2)$$

Accuracy is the proportion of correctly identified positive and negative results from all of the test results (equation 3).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

Although sensitivity and specificity are independent of prevalence, the study population can still influence them, because the concentrations in one study population can be higher than in another. For instance, if the concentrations found in one study population are much higher than those found in another study population, it is easier for the screening test to detect the positive cases in former study population than the latter. This means that sensitivity will appear higher for this study population. To detect differences in concentration ranges between study populations, box and whisker plots are used.

The prevalence of a substance within the study group is derived as the proportion of cases in which the substance is detected in the confirmation sample from all the subjects participating in the study (equation 4).

$$Prevalence = (TP + FN) / (number\ of\ subjects) \quad (4)$$

4.4.3. Positive Predicitive Value and Negative Predictive Value

The positive predictive value (PPV) and negative predictive value (NPV) are important to calculate for determining the usefulness of a diagnostic test in a particular study population. It is well known that both positive and negative predictive values are directly dependent on the prevalence of the substance within the population to be investigated. The theorema of Bayes describes this relationship (equations 5 and 6).

$$PPV = \frac{sens \cdot prev}{sens \cdot prev + (1 - spec)(1 - prev)} \quad (5)$$

$$NPV = \frac{spec(1 - prev)}{spec(1 - prev) + prev(1 - sens)} \quad (6)$$

The theorema of Bayes shows that PPV and NPV are a combination of sensitivity, specificity and prevalence. The predictive values can therefore be used to evaluate if it is useful to use a certain test in a certain population. On this basis the prevalences used to determine PPV and NPV for the devices evaluated are those for the relevant substances of abuse in suspect DUI drivers in each country.

4.4.4. Evaluation of results

The devices included in this study were evaluated according to two sets of criteria.

The first criteria were the DRUID cut-offs for the relevant substances (Table 9 and Table 10 above). For this evaluation each test result was interpreted as TP, TN, FP or FN according to the concentration of the individual substances detected in the confirmation sample. So, for example, with an amphetamine test detecting amphetamine, methamphetamine, MDA and MDMA, the investigated case is interpreted as TP if any of the detected substances is found at a concentration above the relevant DRUID cut-off value (25 ng/ml in OF for each substance). However, a positive test case for which concentrations of, for example, amphetamine and methamphetamine are found, both at a lower level than the DRUID cut-off (e.g. 15 ng/ml and 20 ng/ml respectively) would be interpreted as FP. Therefore no cross reactivity was taken into account. The DRUID cut-offs used were set for a separate epidemiological study within the project.

The second set of criteria were the cut-offs and cross reactivities of the devices claimed by the manufacturers. For most of the devices the manufacturers provided information regarding the ‘cut-off’ concentration levels for substances at which the device could be expected to give a positive result (Tables 1-7). In reality, the presence of two or more substances, both relevant to a specific test (e.g. amphetamines) can give rise to cross-reactivity, wherein the ‘combined’ concentration of the two substances is sufficient to be expected to give a positive result for the test. In addition, the individual concentrations at which each substance alone is expected to give a positive result can differ; therefore it is necessary to take these individual concentrations into account when calculating the combined concentration. This can be done by means of a cross-reactivity equation.

A hypothetical example for an amphetamines test is shown in Table 11.

Table 11. Substance cut-offs and cross-reactivities for a hypothetical amphetamines test.

Substance	Cut-off in OF / ng/ml	Cross reactivity / %
Amphetamine	50	100
Methamphetamine	50	100
MDA	25	200
MDMA	65	77

The combined concentration of substances for this example may be calculated by the following equation:

$$c(\text{combined}) = 100\%\ c(\text{AMP}) + 100\%\ c(\text{MAMP}) + 200\%\ c(\text{MDA}) + 77\%\ c(\text{MDMA})$$

This combined concentration is directly comparable to the expected cut-off of 50 ng/ml shown for amphetamine. Using the manufacturer cut-offs and derived cross-reactivity equations the on-site test results can be directly compared to the concentrations of substances found in the confirmation sample to ascertain whether the case is classified as TP, TN, FP or FN.

4.4.5. Criteria for performance evaluations

In order to obtain statistically valid sensitivity and specificity calculations, it was determined that at least six positive (in case of sensitivity) or six negative (in case of specificity) cases were needed. If there were fewer cases in the material (for example only four positive cases & sensitivity calculation), calculations were not done and the corresponding calculation was determined “not applicable”. Consequently, PPV and NPV were not calculated if sensitivity or specificity calculations could not be done.

Box and whisker plots were drawn for cases for which there were at least six positive cases both for a negative non-zero concentration test result and for a positive test result. All zero concentration negative results connected to a negative test result were omitted from the plots and are mentioned in the caption of the plots. This way, differences in the concentrations related to positive and negative test results can be clearly compared.

The box in these box and whisker plots represents those cases between the 75^{th} and 25^{th} percentile (Q_3-Q_1), whilst the line that bisects the box is the median concentration of the cases. The whiskers that protrude from the box extend to 1.5 times ‘Q_3-Q_1‘ or, if no case has a value in that range, to the minimum or maximum values. If the data are distributed normally, approximately 95% of the cases are expected to lie between the whiskers. Outliers, denoted by a point, are defined as cases that do not fall within the whiskers. Extreme outliers are denoted by asterisks and represent cases that have values more than three times ‘Q_3-Q_1’ beyond the limits of the box. It should be remembered that the scale for concentration in the box and whisker plots in this report is logarithmic base 10, not linear. This scale was used in order to facilitate viewing cases with very high concentrations whilst not unduly compressing the plot.

In order to allow a uniform description of the sensitivity, specificity, and accuracy of devices in each country report the designations presented in Table 12 were used.

Table 12. Designations used for performance evaluation.

Score	Designation
100%	excellent
90-99.9%	very high
80-89.9%	high
70-79.9%	moderate
60-69.9%	low
<60%	very low

Application of these terms to the specific values is not a standard procedure and this was only done to facilitate comparison of the different devices.

5. Country report - Belgium

Kristof Pil, Trudy Van der Linden, Sara-Ann Legrand, An-Sofie Goessaert, Jolien Veramme, Sylvie Vanstechelman, Alain Verstraete.
Faculty of medicine and health sciences,Department of clinical chemistry, microbiology and immunology, Ghent University (UGent)

5.1. Introduction

The Belgian part of the study was carried out in centres for drug addiction and a small part during the roadside survey. The aim of the study was to gain information on the analytical performance of the on-site tests.

In the Belgian study, five devices were tested: Varian OraLab6, Dräger DrugTest 5000 (cut-off 5ng/ml for Δ^9-THC), Cozart DDS, Mavand Rapid STAT and Innovacon OrAlert. Altogether 767 tests were done. The tested persons were tested with one or more on-site oral fluid screening tests and gave one oral fluid StatSure sample. The results of the on-site tests were evaluated based on the ultra-performance liquid chromatography-mass spectrometry (UPLC–MS/MS) analysis result of the oral fluid sample.

5.2. Materials and Methods

5.2.1. Ethical approval

The study protocol was submitted to the ethics committee of Ghent University Hospital. Approval was obtained on 25/06/2009. (Belgian registration number B67020096426). Informed consents were signed by all volunteers.

5.2.2. Sample collection

Initially it was planned to collect samples during roadside sessions in DRUID task 2.2.a. However, since the number of expected positive cases in this setting would be low, a different sampling population was chosen. The majority of samples were collected in centres for treatment of drug addiction.

For evaluation of the Varian OraLab6, 200 samples were collected in an open centre for treatment of drug addiction: the Medical Social Shelter (MSOC) Ghent. Fifty samples were also collected during the roadside survey, creating a total of 250 samples.

For the evaluation of the Dräger DrugTest 5000, Innovacon OrAlert, Mavand Rapid STAT and Cozart DDS the majority of the samples were collected in MSOC Ostend, smaller numbers of samples were collected at the Psychiatric emergency intervention centre, Ghent university hospital (UPSIE Ghent) and the drug crisis intervention centre in Wondelgem.

Each volunteer was asked to first give an oral fluid sample with the StatSure saliva sampler. After this, one or more on-site tests were performed. If more than one on-site test was taken from one volunteer, at least five minutes time had to elapse between the two oral fluid screenings.

All samples were immediately transported to the laboratory for confirmation analysis. Samples were stored at -20°C. Confirmation analyses were performed less than one month after collection.

A checklist of clinical signs of impairment (CSI) was conducted on study participants during the evaluation of the Varian OraLab6. The CSI - sheet contained 28 parameters as described in Table 42 (Annex 1). These parameters were chosen because there is a known or suspected correlation with certain drugs or drug classes. The CSI performers were two medical students in their 4th and 5th year of medical school, not trained police officers.

5.2.3. Analytical method

General confirmation analysis

Oral fluid samples were extracted using liquid-liquid extraction. Internal standard (20 ng/ml) was added to 400 µl oral fluid prior to extraction. 200 µl ammonium bicarbonate (0.2 M, pH 9.3) and 1.25 ml ethyl acetate/heptane (4:1) were added. The mixture was shaken (15 minutes) and centrifuged (3000 rpm, 5 minutes); organic phase was removed and evaporated at room temperature. Analytes were resuspended in 100 µl methanol/water (1:1) and transferred to a vial for analysis with UPLC-MS/MS.

Chromatographic separation was performed on an Acquity^{TM} ultra performance liquid chromatograph (Waters, Zellik, Belgium). The system was equipped with an Acquity UPLC BEH C18 column (1.7µm, 2.1 x 100mm) and a Vanguard BEH C18 pre-column (1.7µm, 2.1 x 5mm). A gradient elution of water with 2mM ammonium bicarbonate pH 9.3 (A) and methanol (B) was used. The column oven was heated to 60°C. Twenty-five µl was injected in a partial loop using needle overfill mode. Both the ratio between aqueous and organic solvent and total flow-rate were changed over time. Total-run-time including re-equilibration was 7 min. (Figure 12)

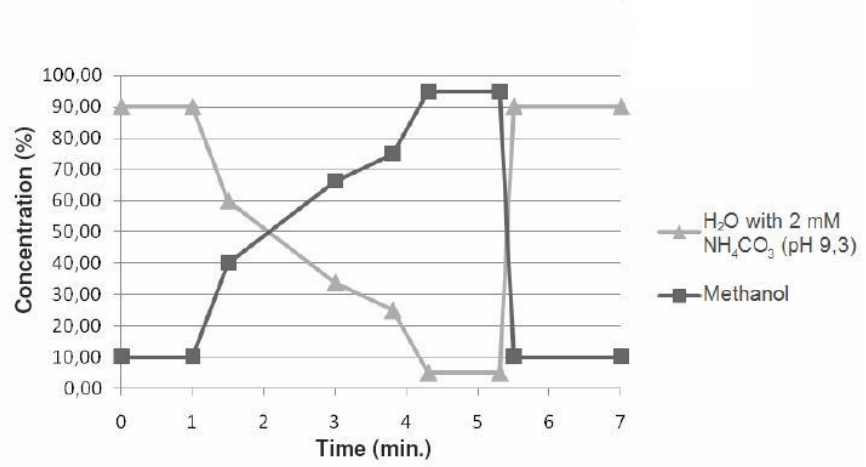

Figure 12. Mobile phase gradient used for UPLC.

All standards and deuterated internal standards were purchased from LGC Standards (Molsheim, France).

Detection was performed using a Waters Quattro Premier XE tandem mass spectrometer. The general source parameters used are given in Table 43 (Annex 2).

To determine component-dependent mass spectrometric parameters, all analytes were infused separately. Retention time windows were defined as the retention time ± 0.15 min for analytes except amphetamines (± 0.35 min). For each analyte, the maximum dwell times were chosen at which for all analytes a minimum of twelve data points per peak was obtained. Multiple reaction monitoring (MRM) transitions, collision energies, cone voltages, retention times and dwell times are shown in Table 44 (Annex 2).

For calibration, oral fluid was collected from drug-free volunteers using the StatSure saliva sampler. Concentrations are expressed as ng/ml undiluted oral fluid. Samples were spiked in oral fluid/StatSure buffer mixture assuming a collection volume of 1 ml oral fluid.

Method validation

Selectivity was tested by injecting ten blank samples from different sources to check for interfering signals. No interfering signals were observed.

Nine calibration points were used: 0, 0.5, 1, 5, 10, 20, 50, 100 and 200 ng/ml. The best fitting regression model and weighting factors were determined by injection of 6 replicates of 7 concentration levels (0, 0.5, 1, 5, 10, 20, 50, 100 and 200 ng/ml) on 2 separate days. All analytes had good linearity (r^2>0.99 except benzoylecgonine: r^2 = 0.983)

Accuracy and precision were determined based on 4 injections of low, medium and high concentrations (5, 20 and 100 ng/ml) relative to the calibration range on five different days.

To measure absolute and relative matrix effects, oral fluid was collected from 10 different volunteers. Five replicates were used for low, medium and high concentrations. Absolute matrix effect was defined as the ratio between peak areas of samples spiked after extraction compared to injections of spiked mobile phase. To determine if the absolute matrix effect has an influence on quantification of analytes, the relative matrix effect was defined as described by Matuszewski (10): calibration lines were prepared in oral fluid from five different volunteers. Slopes from the standard lines were determined. It is recommended that the coefficient of variation (CV) of these slopes does not exceed 3-4% in order to conclude that the method is not affected by relative matrix effects. This value was lower than 2% for all analytes for which a deuterated internal standard was used. The analytes without deuterated internal standards were slightly more affected by relative matrix effects, but values were still within acceptable ranges. (highest CV: 3.1 % for bromazepam).

All validation parameters are described in Table 45 (Annex 2).

Phencyclidine - analysis (Varian OraLab6 samples)

Five positive screening results were found for phencyclidine (PCP) in our evaluation, although the prevalence of PCP use is extremely low in Belgium. Since the anti-depressant venlafaxine and its main metabolite, o-desmethylvenlafaxine, can cause false positive results for PCP, a confirmation method for PCP, venlafaxine and o-desmethylvenlafaxine was developed on UPLC-MS/MS. The same liquid-liquid extraction procedure was used as described in 2.2.1, but only 100 µl of sample was used.

PCP-D5 and trimipramine-D3 were used as internal standards. Detection was performed in multiple reaction monitoring mode using the following transitions: 244.3 > 158.9 and 116.8 for PCP, 249.3 > 164.0 for PCP-D5, 278.3 > 120.9 and 90.9 for venlafaxine, 264.3 > 106.9 and 152.9 for o-desmethylvenlafaxine and 298.3 > 102.93 for trimipramine-D3.

Inaccuracy and imprecision were lower than 15%, and selectivity of the method was proven.

5.2.4. External quality control

Regular participation took place in the external quality control program organized by RTI International (NC, USA). Based on the results, the performance of the analytical method was at an excellent level.

5.2.5. Positive and negative predictive value (PPV and NPV)

Calculation of PPV and NPV in all sections below are based on prevalence of drug use in suspected drivers published by Raes *et al.* (11) These data show that in blood samples taken from May 2000 to February 2005, cannabis was found most often above the legal cut-off (73.5 % of the cases), followed by MDMA (methylene-dioxy-methamphetamine) (20.4 %), amphetamine (19.8 %), benzoylecgonine (17.9 %), cocaine (6,9 %) and morphine (2.7 %).

For calculation of PPV and NPV for opiates and cocaine, the prevalence for benzoylecgonine and morphine were used respectively.

Since benzodiazepines are not included in driving under influence (DUI) legislation in Belgium, no data are available on prevalence in the suspected drivers and hence no calculations of PPV and NPV are included for these substances.

5.3. Results

5.3.1. Varian OraLab6

Table 13. Statistical evaluations for Varian OraLab6 using DRUID and device cut-offs.

	DRUID cut-offs						Device cut-offs					
	COC	**OPI**	**CAN**	**AMP**	**PCP**	**MAMP**	**COC**	**OPI**	**CAN**	**AMP**	**PCP**	**MAMP**
TP	19	84	18	19	0	0	19	83	18	19	0	0
FP	0	3	2	0	5	0	0	2	2	0	5	0
TN	195	125	135	216	244	249	195	128	203	224	244	249
FN	35	37	94	14	0	0	35	34	26	6	0	0
N. of successful tests	249	249	249	249	249	249	249	249	249	249	249	249
Failed device	1	1	1	1	1	1	1	1	1	1	1	1
Missing analysis	0	0	0	0	0	0	0	0	0	0	0	0
Sensitivity	35%	69%	16%	58%	n.a.	n.a.	35%	71%	41%	76%	n.a.	n.a.
Specificity	100%	98%	99%	100%	96%	100%	100%	99%	99%	100%	96%	100%
Accuracy	86%	84%	61%	94%	49%	100%	86%	85%	89%	98%	98%	100%
Prevalence	22%	49%	45%	13%	0%	0%	22%	48%	18%	10%	0%	0%
PPV	100%	46%	97%	100%	n.a.	n.a.	100%	57%	99%	100%	n.a.	n.a.
NPV	88%	99%	30%	91%	n.a.	n.a.	88%	99%	38%	94%	n.a.	n.a.

n.a. - non applicable

Cocaine

54 positive cases for cocaine were found: 19 true positives and 35 false negatives, using either DRUID or device cut-offs. All cases were positive for cocaine (range 10.28-20632.06 ng/ml) and 20 were for positive for both cocaine and benzoylecgonine (range 12.73-944.17 ng/ml). Sensitivity was very low (35%), specificity was excellent (100%)

Based on information provided by the manufacturer following cross-reactivity was used:

COC = cocaine + 6.67 % benzoylecgonine

Comparison of data with or without application for cross-reactivity indicated that the application of cross reactivity does not make a great difference. The box and whisker plot showed that there is a spread of drug concentrations that yielded a negative test card result. There was an overlap between the negative and positive test card results, which means there is no clear distinction in the concentrations that give a negative result with the oral fluid test and those with a positive result.

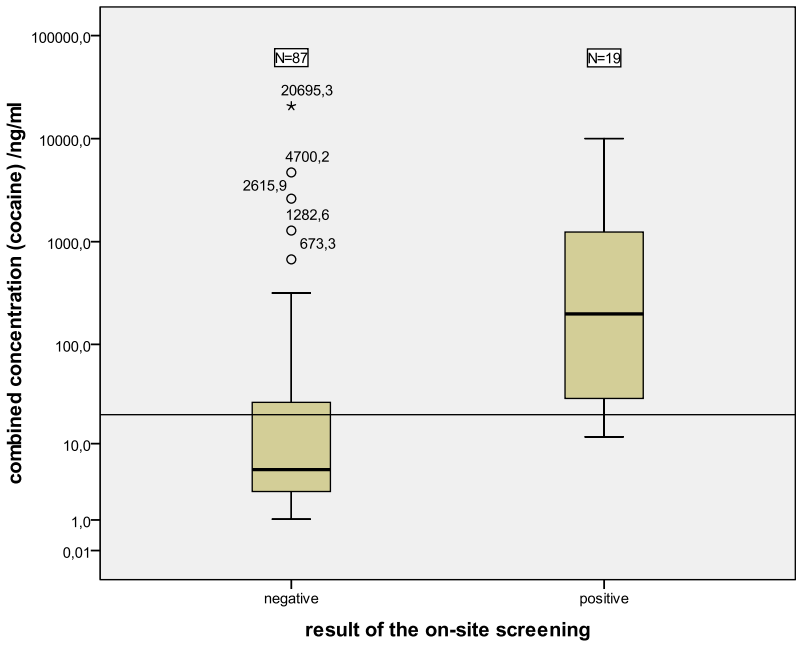

Figure 13. Box and whisker plot for the Varian OraLab6 cocaine test, test cut-off and cross-reactivity stated by the manufacturer used. 143 cases with 0 ng/ml of cocaine in their OF were tested negative for cocaine. These cases are not included in the plot. Horizontal line at 20 ng/ml indicates test cut-off.

Opiates

121 positive cases were found for opiates using DRUID cut-offs. When using device cut-offs 117 positive cases were found. 97 cases were positive for morphine (range 20.59-9159.1 ng/ml), 116 positive for 6-acetylmorphine (range 5.86-15658.77 ng/ml) and 80 positive for codeine (range 20.32-1187.82 ng/ml). From the 121 positive cases 77 were positive for all three analytes.

Sensitivity was low for the DRUID cut-off (69.4%) and moderate for device cut-off (70.9%); specificity was very high for the DRUID cut-off (97.7%) and for device cut-off (98.5%).

Based on the information provided by the manufacturer the following cross-reactivity was used:

OPI = morphine + 160% 6-acetylmorphine + 160% codeine.

Comparison of data with or without application for cross-reactivity indicated that the application does not make a great difference. The box and whisker plot showed that there is a spread of drug concentrations that yielded a negative test card result. There was an overlap between the negative and positive test card results, which means there is no clear distinction in the concentrations that give a negative result with the oral fluid test and those with a positive result.

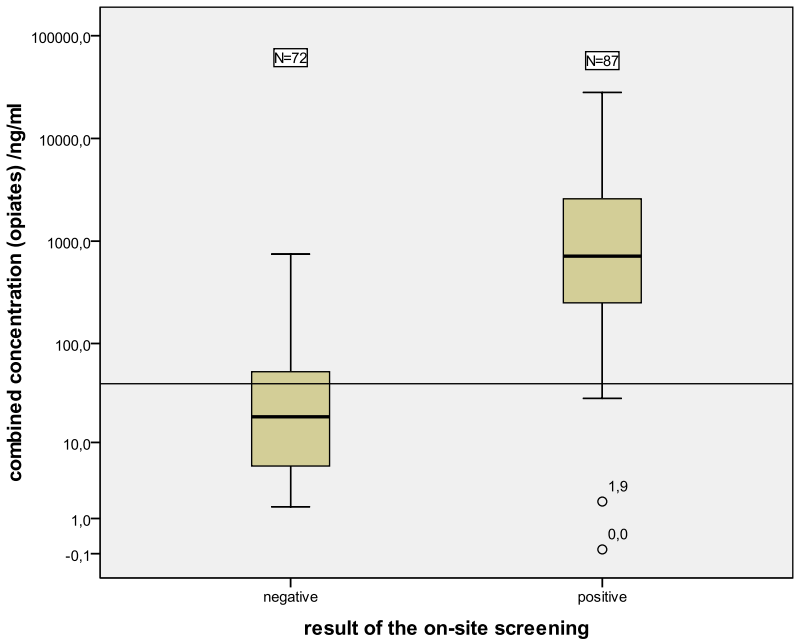

Figure 14. Box and whisker plot for the Varian OraLab6 opiates test, test cut-off and cross-reactivity stated by the manufacturer used. 90 cases with 0 ng/ml of opiates in their OF were tested negative for opiates. These cases are not included in the plot. Horizontal line at 40 ng/ml indicates test cut-off.

Cannabis

112 positive cases for THC (Δ^9-tetra-hydro-cannabinol) were found (18 TP and 94 FP) when using the DRUID cut-off, 44 positive cases were found when using the device cut-off. The device cut-off is higher than the DRUID cut-off (50- fold difference). The concentration range was 1.04-3967.31 ng/ml.

No cross reactivity was taken into account. So both DRUID and device cut-offs are indicated on the same box and whisker plot. The box and whisker plot showed that there is a spread of drug concentrations that yielded a negative test card result. There was an overlap between the negative and positive test card results, which means there is no clear distinction in the concentrations that give a negative result with the oral fluid test and those with a positive result.

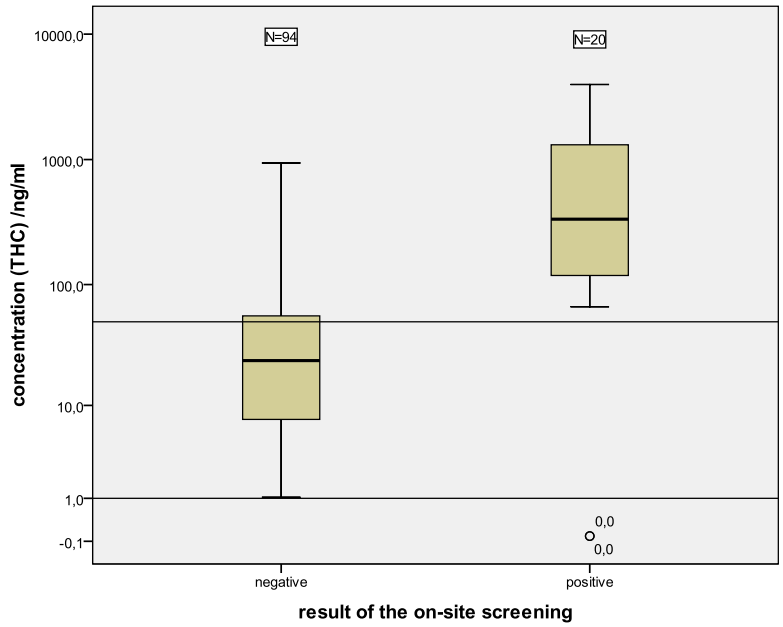

Figure 15. Box and whisker plot for the Varian OraLab6 cannabis test. 135 cases with 0 ng/ml of THC in their OF were tested negative for cannabis. These cases are not included in the plot. Horizontal lines indicate test cut-off (50 ng/ml) and DRUID cut-off (1ng/ml).

Amphetamines

33 positive cases were found for amphetamines using the DRUID cut-offs, 25 were found when using the device cut-off. The concentration range was 25.01-21153.33 ng/ml.

Sensitivity was 57.6% and 76% respectively. Specificity was 100% so PPV was 100% for both cut-offs and NPV 91% and 94.4% respectively.

No cross-reactivity was mentioned by the manufacturer hence none was taken into account. Both DRUID and device cut-offs are indicated on the box and whisker plot. There was a clear distinction at a concentration > 100 ng/ml, except for the outliers.

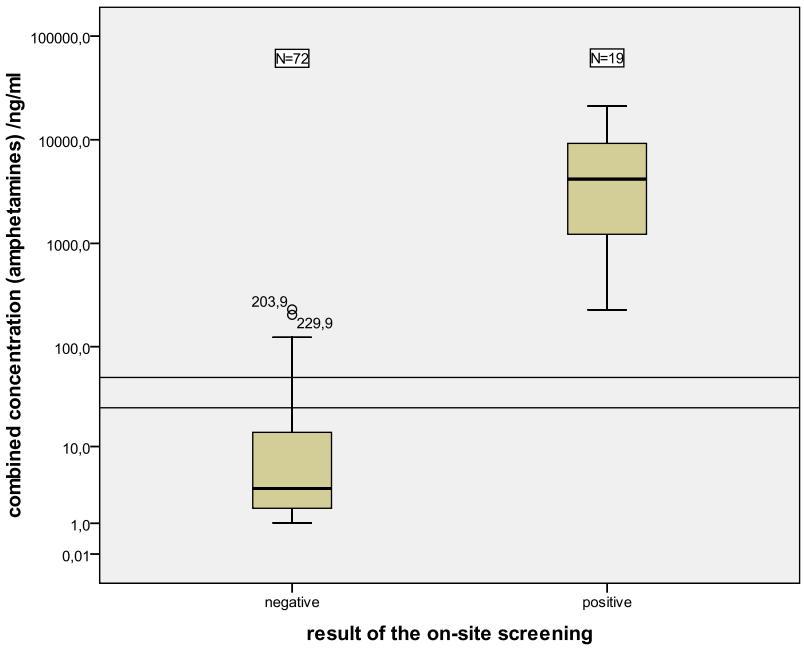

Figure 16. Box and whisker plot for the Varian OraLab6 amphetamine test. 158 cases with 0 ng/ml of amphetamine in their OF were tested negative for amphetamines. These cases are not included in the plot. Horizontal lines indicate test cut-off (50 ng/ml) and DRUID cut-off (25ng/ml).

PCP: False positives

Five cases screened positive for PCP although the prevalence of PCP use is extremely low in Belgium. Use of PCP is very rarely observed in Europe.

A few case reports have demonstrated that the anti-depressant venlafaxine and its main metabolite o-desmethylvenlafaxine could cause false positive results for PCP on qualitative immunoassays for urine testing (12, 13).

The method described in the section 'Phencyclidine - analysis (Varian OraLab6 samples)' above was applied to four samples with a positive screening result; the fifth sample could not be analysed since there was not sufficient oral fluid left after the initial confirmation analysis.

Oral fluid samples were spiked with venlafaxine and o-desmethylvenlafaxine (concentrations ranging between 100 and 2000 ng/ml) and applied to the on-site test to determine the cut-off for a false positive result.

The tests with spiked oral fluid showed that a false positive result is produced with the OraLab6 when concentrations of venlafaxine or o-desmethylvenlafaxine exceeded 400-500 ng/ml. This corresponds to a normal therapeutic concentration. No significant difference in cross-reactivity between the two compounds was observed.

The confirmation analysis showed that there was no PCP present in the oral fluid samples. Venlafaxine and o-desmethylvenlafaxine were present in three out of four samples, in concentrations of 523.8, 426.4, 717.0 and 98.7, 343.6, 474.9 ng/ml respectively. In the fourth false positive case, no venlafaxine or o-desmethylvenlafaxine was found. It can be concluded that three out of four cases with a positive screening result for PCP were false positive due to this cross-reactivity.

Methamphetamine

No positive results were obtained for methamphetamine.

5.3.2. Dräger DrugTest 5000

Table 14. Statistical evaluations for Dräger DrugTest 5000 using DRUID and device cut-offs.

	DRUID cut-offs						Device cut-offs					
	COC	**OPI**	**BZO**	**CAN**	**AMP**	**MAMP**	**COC**	**OPI**	**BZO**	**CAN***	**AMP**	**MAMP**
TP	6	69	32	24	6	0	2	70	31	24	6	0
FP	1	9	0	1	0	0	5	8	1	1	0	0
TN	124	50	87	90	129	137	126	52	104	100	130	137
FN	6	9	18	21	2	0	4	7	1	11	1	0
N of successful tests	137	137	137	137	137	137	137	137	137	137	137	137
Failed devices	2	2	2	2	2	2	2	2	2	2	2	2
Missing analysis	0	0	0	0	0	0	0	0	0	0	0	0
Sensitivity	50%	89%	64%	53%	75%	n.a.	33%	91%	97%	69%	99%	n.a.
Specificity	99%	85%	100%	99%	100%	100%	96%	87%	99%	99%	100%	100%
Accuracy	95%	87%	87%	84%	99%	100%	93%	89%	99%	91%	99%	100%
Prevalence	8.8%	57%	37%	33%	5.8%	0%	4.3%	56%	23%	26%	5.1%	0%
PPV	93%	16%	n.a.	99%	100%	n.a.	66%	16%	n.a.	100%	100%	n.a.
NPV	90%	100%	n.a.	43%	94%	n.a.	87%	100%	n.a.	53%	100%	n.a.

*(cut-off 5ng/ml for Δ^9-THC)

Cocaine

Due to the low number for positive cases for cocaine (12 when using the DRUID cut-off, 6 for the device cut-off), conclusions for this analyte should be interpreted cautiously. Using the DRUID cut-off 12 cases were positive for cocaine (range 10.79-132.83 ng/ml) and 7 positive for benzoylecgonine (range 11.44-138 ng/ml)

Based on the information provided by the manufacturer, the following cross-reactivity was used:

COC = cocaine + 28% benzoylecgonine.

The box and whisker plot for negative and positive cocaine screenings overlap. Concentrations were low in our study population. Since the number of positives for cocaine was very low, little can be concluded based on these results

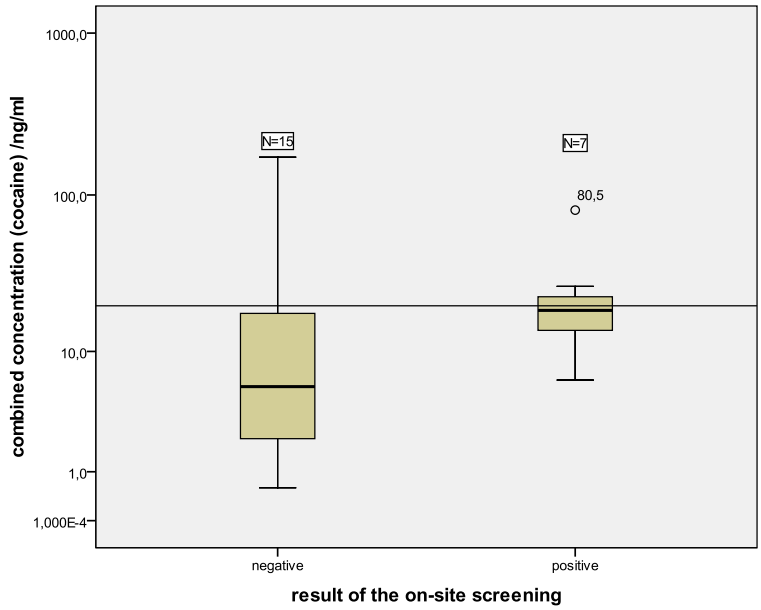

Figure 17. Box and whisker plot for the Dräger DrugTest 5000 cocaine test, test cut-off and cross-reactivity stated by the manufacturer used. 115 cases with 0 ng/ml of cocaine in their OF were tested negative for cocaine. These cases are not included in the plot. Horizontal line at 20 ng/ml indicates test cut-off.

Opiates

78 positives cases were found when using the DRUID cut-off, 77 when using the device cut-off. 67 cases were positive for morphine (range 21.67-13870.21 ng/ml), 68 positive for 6-acetylmorphine (range 5.95-15523.77 ng/ml) and 50 positive for codeine (range 20.39-2783.63 ng/ml). 42 cases were positive for all three analytes.

Application of the DRUID cut-off gave a high sensitivity (88.5%). Sensitivity for opiates detection increased slightly when using the device cut-off (90.9%).

Based on the information provided by the manufacturer, the following cross-reactivity was used:

OPI = morphine + 80% codeine + 57 % 6-acetylmorphine.

Since the DRUID cut-off is the same as the one claimed by the manufacturer, it is clear that the cross-reactivities were the reason why one sample just below the DRUID-cutoffs gave a positive screening result.

Specificity for opiates was mediocre. The cross-reactivities alone were not enough to account for the over-sensitivity (and hence the lower specificity). There were often positive screenings with combined opiate concentrations below the cut-off. One possible explanation for this is cross-reactivity for methadone: it is stated in the user guide of the manufacturer that this compound can cause false positives for opiates at 100000 ng/ml. Although this concentration is very high, this could occur shortly after oral intake of a normal dose. Therefore it is stated in the instructions that a subject should not consume anything 10 minutes prior to taking the sample.

It has to be noted that although the manufacturers only reported false positives at concentrations above 100000 ng/ml, concentrations below this value could also affect the results. For instance, if methadone is present at ‘only’ 10000 ng/ml, this will contribute to the reaction of the OPI test. So 10 ng/ml morphine + 5 ng/ml codeine + 10000 ng/ml methadone would give a positive test result (whilst without methadone present this person would test negative), and we would interpret this as an ‘over-sensitive’ test.

Moreover, presence of other opiates, which are not included in the confirmation analysis could cause this over-sensitivity: e.g. pholcodine, hydromorphone, dihydrocodeine.

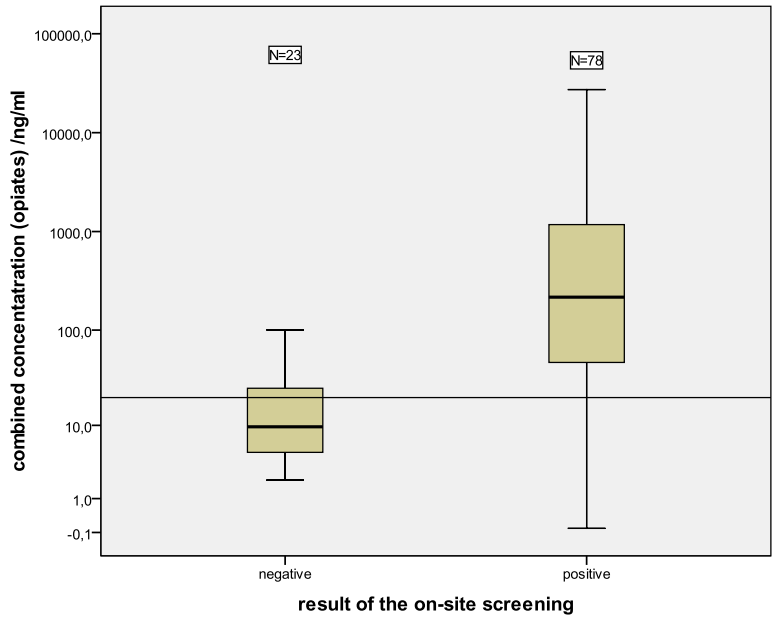

Figure 18. Box and whisker plot for the Dräger DrugTest 5000 opiates test, test cut-off and cross-reactivity stated by the manufacturer used. 37 cases with 0 ng/ml of opiates in their OF were tested negative for opiates. These cases are not included in the plot. Horizontal line at 20 ng/ml indicates test cut-off.

Benzodiazepines

50 positive cases for benzodiazepines were found using the DRUID cut-offs. When using the device cut-offs, 32 positive cases were found. The most common benzodiazepine findings are shown in Table 15.

Table 15. Benzodiazepine findings in the OF samples used for the Dräger evaluations.

Analyte	N	Range / ng/ml	Average / ng/ml	Median / ng/ml
Nordiazepam	38	1.1-448.8	35.1	15.6
Bromazepam	31	1.4-3454.2	285.2	81.2
Diazepam	14	6.1-3903.7	370.8	19.0
Oxazepam	9	1.3-125.07	13.7	4.9
Lorazepam	5	1.6-53.8	20.0	3.9
Alprazolam	1	2.7		

Based on the information provided by the manufacturer, the following cross-reactivity was used:

BZO = diazepam + 66% alprazolam + 30% 7-aminoflunitrazepam + 21% bromazepam + clonazepam + 133% flunitrazepam + 7.5% lorazepam + 33% nordiazepam + 37.5% oxazepam.

Sensitivity and specificity were very high (96.9 and 99% respectively) when the manufacturer cut-offs and cross-reactivities were applied. The box and whisker plot was almost completely separated.

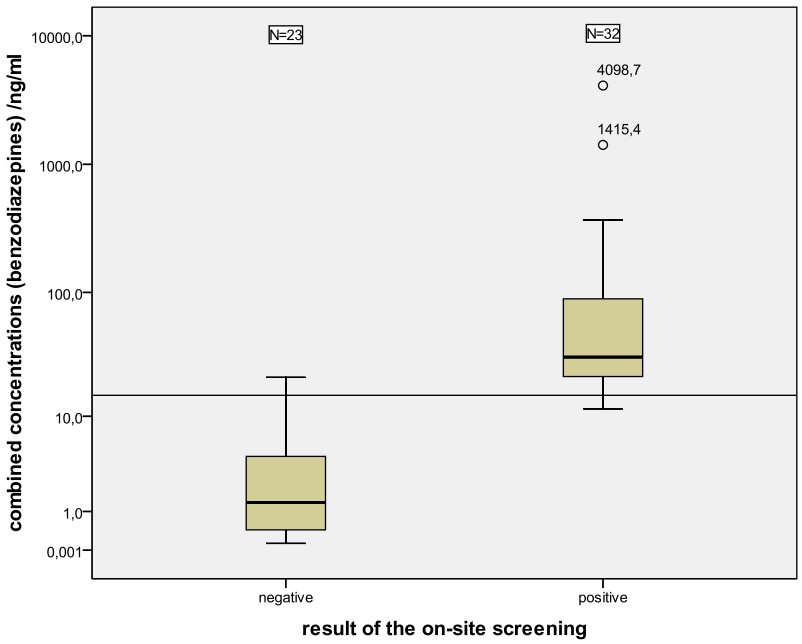

Figure 19. Box and whisker plot for the Dräger DrugTest 5000 Benzodiazepine test, test cut-off and cross reactivity stated by the manufacturer. 82 cases with 0 ng/ml of any benzodiazepines in ther OF were tested negative for benzodiazepines. These cases are not included in the plot. Horizontal line at 15 ng/ml indicates test cut-off.

Cannabis

45 positive cases for THC were found using the DRUID cut-off. When using the device cut-off, 35 positives were found. The concentration range was 10.02-2451.80 ng/ml.

When applying the DRUID cut-off sensitivity was 53%, when applying the manufacturer's cut-off it increased to 68.5%.

The box and whisker plot for THC was well separated: more than 75% of the negative screenings and less than 25 % of the positive screenings were below 10 ng/ml.

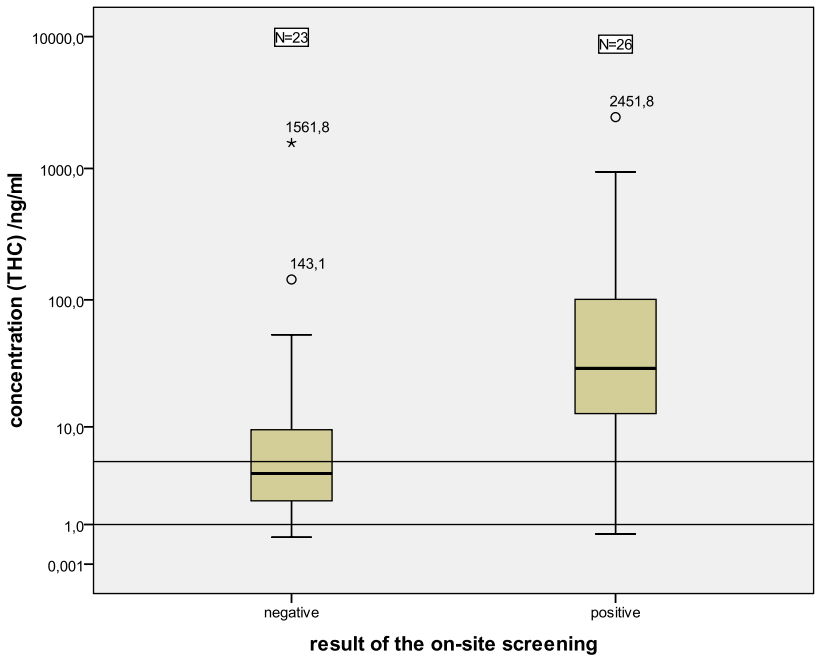

Figure 20. Box and whisker plot for the Dräger DrugTest 5000 cannabis test. 88 cases with 0 ng/ml of THC in their OF were tested negative for cannabis. These cases are not included in the plot. Horizontal lines indicate test cut-off (5 ng/ml) and DRUID cut-off (1ng/ml).

Amphetamines

Due to the low number of positive cases for amphetamines (8 when using the DRUID cut-off, 7 for the device cut-off), results for this analyte should be interpreted cautiously. The concentration range was 40.91-7338.20 ng/ml.

Application of the DRUID cut-off gave a moderate sensitivity (75%). Sensitivity for amphetamine detection increased when using the device cut-off (99%).

No cross-reactivity was mentioned by the manufacturer, hence none was taken into account. Both DRUID and device cut-offs are indicated on the box and whisker plot.

The box and whisker plot for amphetamines was completely separated around a concentration of 1000 ng/ml: all samples with high amphetamine concentrations (and one sample with concentration around 100 ng/ml) gave a positive screening result while all the other samples below 1000 ng/ml had a negative screening result.

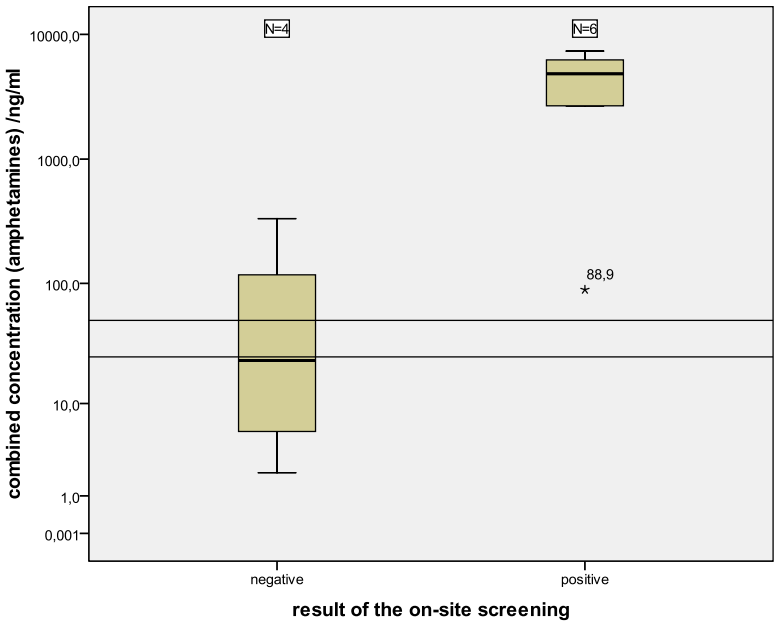

Figure 21. Box and whisker plot for the Dräger DrugTest 5000 amphetamine test. 127 cases with 0 ng/ml of amphetamine in their OF were tested negative for amphetamines. These cases are not included in the plot. Horizontal lines indicate test cut-off (50 ng/ml) and DRUID cut-off (25ng/ml).

Methamphetamines

No positive results were found for methamphetamine.

5.3.3. Cozart DDS

Table 16. Statistical evaluations for Cozart DDs using DRUID and device cut-offs.

	DRUID cut-offs						Device cut-offs					
	COC	**OPI**	**BZO**	**CAN**	**AMP**	**XTC**	**COC**	**OPI**	**BZO**	**CAN**	**AMP**	**XTC**
TP	1	45	25	11	4	0	1	48	25	11	4	0
FP	1	4	5	0	1	0	1	1	5	0	1	0
TN	129	80	81	87	131	138	132	78	99	110	132	138
FN	7	9	27	40	2	0	4	11	9	17	1	0
N of successful tests	138	138	138	138	138	138	138	138	138	138	138	138
Failed devices	0	0	0	0	0	0	0	0	0	0	0	0
Missing analysis	0	0	0	0	0	0	0	0	0	0	0	0
Sensitivity	13%	83%	48%	22%	67%	n.a.	n.a.	81%	74%	39%	n.a.	n.a.
Specificity	99%	95%	94%	100%	99%	100%	99%	99%	95%	100%	99%	100%
Accuracy	94%	91%	77%	71%	98%	100%	96%	91%	90%	88%	99%	100%
Prevalence	5.8%	39%	38%	37%	4.3%	0%	3.6%	43%	25%	20%	3.6%	0%
PPV	77%	33%	n.a.	100%	95%	n.a.	n.a.	64%	n.a.	100%	n.a.	n.a.
NPV	85%	100%	n.a.	32%	92%	n.a.	n.a.	100%	n.a.	37%	n.a.	n.a.

n.a. not applicable

Cocaine

Due to the low number of positive cases for cocaine (8 when using DRUID cut-off, 5 when using device cut-off), the results for this analyte should be interpreted cautiously.

When using DRUID cut-off 8 positive cases for cocaine (range 10.2-450.2 ng/ml) and 6 positive cases for benzoylecgonine (range 10.8-633.1 ng/ml) were found.

Based on the information provided by the manufacturer, the following cross-reactivity was used:

COC = cocaine + benzoylecgonine.

Opiates

54 postive cases were found for opiates when using the DRUID cut-offs. 54 cases were positive for morphine (range 20.86-9538 ng/ml), 37 cases were positive for the three analytes: morphine, 6-acetylmorphine and codeine (range 6-AM: 5.54-14090.2 ng/ml; range codeine: 25.90-1421.9 ng/ml). Sensitivity and specificity were 83% and 95% respectively. 59 positive cases were found applying the device cut-offs; sensitivity and specificity were 81% and 99% respectively.

Based on the information provided by the manufacturer, the following cross-reactivity was used:

OPI = morphine + 55% codeine + 60% 6-acetylmorphine

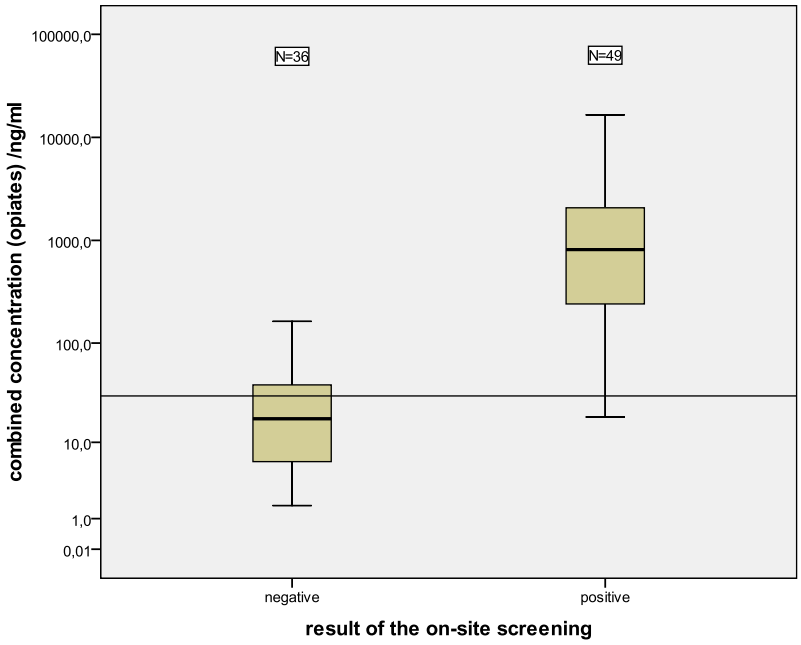

Figure 22. Box and whisker plot for the Cozart DDS opiates test, test cut-off and cross-reactivity stated by the manufacturer used. 53 cases with 0 ng/ml of opiates in their OF were tested negative for opiates. These cases are not included in the plot. Horizontal line at 30 ng/ml indicates test cut-off.

Benzodiazepines

52 positive cases were found when using the DRUID cut-offs. When applying device cut-offs, 34 positive cases were found. The most common benzodiazepine findings are shown in Table 17.

Table 17. Benzodiazepine findings in the OF samples used for the Cozart DDS evaluations.

Analyte	N	Range / ng/ml	Average / ng/ml	Median / ng/ml
Nordiazepam	38	1.1-1307.4	83.9	19.3
Bromazepam	28	1.2-1263.3	140.2	76.3
Oxazepam	21	1.2-163.1	26.1	8.2
Diazepam	17	10.2-1748.9	183.1	26.9
Lorazepam	3	2.7-160.1	56.5	6.0
Alprazolam	3	3.2-14.7	7.1	3.5
Clonazepam	2	2.5-2.6	2.6	2.6
7-amino-clonazepam	2	5.6-8.6	7.1	7.1

Based on information provided by the manufacturer, the following cross-reactivity was used:

BZO = 40% diazepam + 57% alprazolam + 40% clonazepam + 40% flunitrazepam + 10% lorazepam + 27% oxazepam + nordiazepam + bromazepam + 7-amino-clonazepam + 7-amino-flunitrazepam

Note: no cross-reactivities were reported by the manufacturer for four substances included in the confirmation analysis: nordiazepam, bromazepam, 7-amino-clonazepam and 7-amino-flunitrazepam. It can be expected that cross-reactivity exists. However, these were not tested by the manufacturer since there is a multitude of benzodiazepines available on the European market and not all could be tested. Therefore 100% cross-reactivity was assumed in the calculations.

Although sensitivity for detection of benzodiazepines was moderate (device cut-off) to low (DRUID cut-off), the box and whisker plot was very well separated indicating that the test had discriminatory power between low and high concentrations. Sensitivity for benzodiazepines increased from 48.1 % (using DRUID cut-offs) to 73.5 % when applying cross-reactivities and manufacturer cut-offs. This increase can probably mainly be attributed to the increase in the cut-off - from 10 ng/ml for the DRUID cut-off to 20 ng/ml for the device cut-off (target compound temazepam).

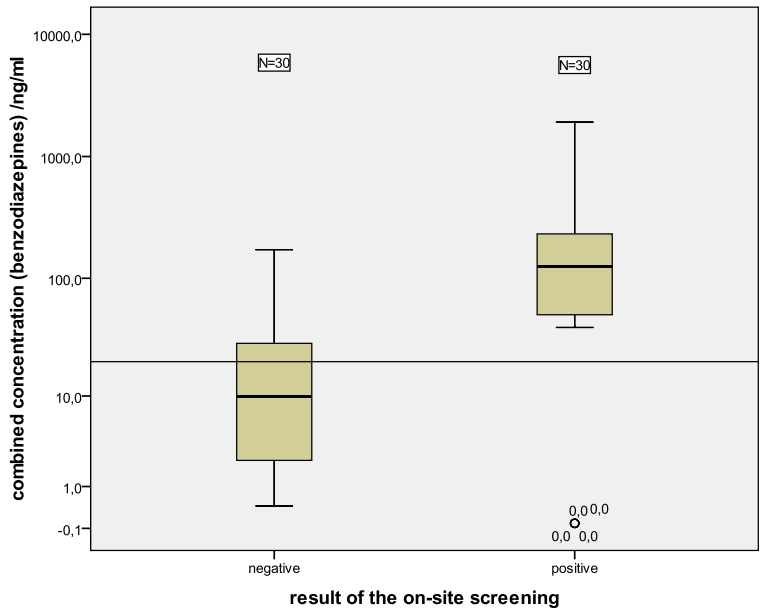

Figure 23. Box and whisker plot for the Cozart DDS Benzodiazepine test, test cut-off and cross reactivity stated by the manufacturer. 78 cases with 0 ng/ml of any benzodiazepines in their OF were tested negative for benzodiazepines. These cases are not included in the plot. Horizontal line at 20 ng/ml indicates test cut-off.

Cannabis

51 positive cases for THC were found when applying the DRUID cut-off (range 1.10-3310.5 ng/ml) and 28 positive cases when using the device cut-off. Sensitivity increased from 21.6% (using the DRUID cut-off) to 39.3% when applying the device cut-off.

Although sensitivity for detection of THC was low, the box and whisker plot was very well separated indicating that the test had discriminatory power between low and high concentrations, but that the actual cut-off of the test is a lot higher than the DRUID cut-off. When the device cut-offs are applied, sensitivity increased from 21.6% (using DRUID cut-off) to 39.3%.

No cross reactivity was taken into account for. So both DRUID and device cut-offs are indicated on the same box and whisker plot.

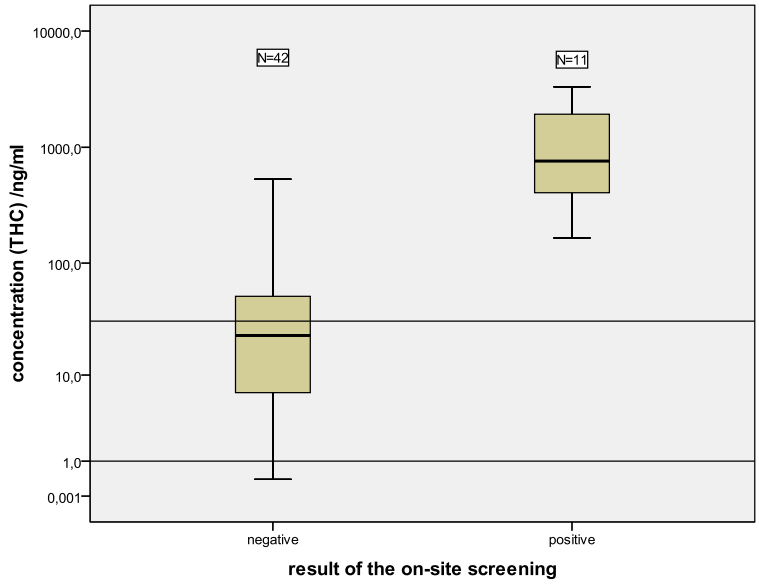

Figure 24. Box and whisker plot for the Cozart DDS cannabis test. 85 cases with 0 ng/ml of THC in their OF were tested negative for cannabis. These cases are not included in the plot. Horizontal lines indicate test cut-off (31 ng/ml) and DRUID cut-off (1ng/ml).

Amphetamines

Due to the low number of positive cases for amphetamines (6 when using the DRUID cut-offs, 5 when using the device cut-off), the results for this analyte should be interpreted cautiously.

When using DRUID cut-off following concentration range was found: 32.41-2472.2 ng/ml.

Sensitivity was 66.7% when applying the DRUID cut-off, for the device cut-off no sensitivity was calculated because the number of positive cases was lower than 6. Specificity was in both cases 99.2%.

MDMA

No positive results were found for MDMA.

5.3.4. Mavand Rapid STAT

Table 18. Statistical evaluations for Mavand Rapid STAT using DRUID and device cut-offs.

	DRUID cut-offs						Device cut-offs					
	COC	**OPI**	**BZO**	**CAN**	**AMP**	**MAMP**	**COC**	**OPI**	**BZO**	**CAN**	**AMP**	**MAMP**
TP	3	57	41	16	1	0	3	62	36	12	1	0
FP	3	7	0	9	4	0	3	2	5	13	4	0
TN	120	63	73	74	123	133	124	57	90	94	123	133
FN	7	6	19	34	5	0	3	12	2	14	5	0
N of successful tests	133	133	133	133	133	133	133	133	133	133	133	133
Failed devices	5	5	5	5	5	5	5	5	5	5	5	5
Missing analysis	0	0	0	0	0	0	0	0	0	0	0	0
Sensitivity	30%	91%	68%	32%	17%	n.a.	50%	84%	95%	46%	17%	n.a.
Specificity	98%	90%	100%	89%	97%	100%	98%	97%	95%	88%	97%	100%
Accuracy	92%	90%	86%	68%	93%	100%	95%	89%	95%	80%	93%	100%
Prevalence	7.5%	47%	45%	38%	4.5%	0%	4.5%	57%	30%	20%	4.5%	0%
PPV	73%	20%	n.a.	89%	57%	n.a.	82%	19%	n.a.	91%	57%	n.a.
NPV	87%	100%	n.a.	32%	83%	n.a.	90%	100%	n.a.	37%	83%	n.a.

n.a. not applicable

Cocaine

Due to the low number of positive cases for cocaine (10 when using DRUID cut-off, 6 when applying device cut-off), the results for these analytes should be interpreted cautiously.

Using DRUID cut-off, 10 positive cases for cocaine (range 10.61-909.34 ng/ml) and 8 positive cases for benzoylecgonine (10.39-456.63 ng/ml) were found.

Based on information provided by the manufacturer, the following cross-reactivity was used:

COC = benzoylecgonine + 6.25% cocaine.

Sensitivity for cocaine was very low (30%) when using the DRUID cut-offs, it increased to 50% when applying the device cut-off. Specificity was 98% for both cut-offs.

Opiates

63 positive cases for opiates were found applying the DRUID cut-offs, 74 when using the device cut-off.

When using the DRUID cut-offs all 63 cases were positive for morphine and 6-acetylmorphine and 48 for codeine. The following concentration ranges were found: morphine: 20.29-7878.5 ng/ml; 6-acetylmorphine: 5.54-14090.2 ng/ml and codeine: 22.89-1421.94 ng/ml.

Sensitivity decreased from 91% when using the DRUID cut-offs, to 84% when applying manufacturer's cut-off and cross-reactivity. Specificity increased from 90% to 97%.

Based on information provided by the manufacturer, the following cross-reactivity was used:

OPI = morphine + 50% 6-acetylmorphine + 333% codeine

The box and whisker plot was separated.

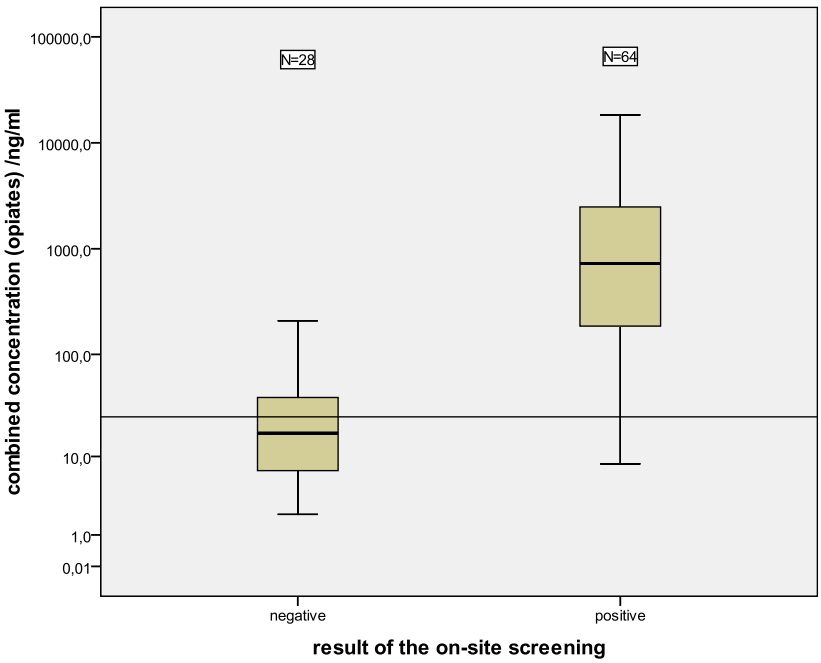

Figure 25. Box and whisker plot for the Rapid STAT opiates test, test cut-off and cross-reactivity stated by the manufacturer used. 36 cases with 0 ng/ml of opiates in their OF tested negative for opiates. These cases are not included in the plot. Horizontal line at 25 ng/ml indicates test cut-off.

Benzodiazepines

60 positive cases were found using the DRUID cut-offs and when applying device cut-off, 38 cases were found. Sensitivity using DRUID cut-offs was 68.3%. The benzodiazepine findings are shown in Table 19.

Table 19. Benzodiazepine findings in the OF samples used for the Rapid STAT evaluations.

Analyte	N	Range / ng/ml	Average / ng/ml	Median / ng/ml
Nordiazepam	40	1.24-209.7	30.7	16.2
Bromazepam	39	1.2-2616.3	231.3	112.8
Oxazepam	14	6.13-192.5	36.0	14.5
Diazepam	14	5.5-843.8	148.5	16.9
Lorazepam	5	2.69-39.1	13.4	6.0
Alprazolam	1	3.5		
Clonazepam	1	2.5		
7-amino-clonazepam	1	5.63		
7-amino-flunitrazepam	1	11.4		

When applying the manufacturer's cut-off and cross-reactivity an increased sensitivity to 95% was obtained. Specificity also increased form 90 to 95 %.

Based on information provided by the manufacturer, the following cross-reactivity was used:

BZO = 400% diazepam + alprazolam + 200% flunitrazepam + 4% lorazepam + 20% bromazepam + nordiazepam + oxazepam + 7-aminoclonazepam + 7-aminoflunitrazepam + 0.02% clonazepam.

Note: no cross-reactivities were reported by the manufacturer for three substances which are included in the confirmation analysis: bromazepam, 7-amino-clonazepam and 7-amino-flunitrazepam. But it was found that bromazepam will give a positive result at 50 ng/ml and a negative result at 20 ng/ml, giving a cross-reactivity of 20% (parent benzodiazepine = oxazepam; cut-off 10 ng/ml). For the latter two substances it can be expected that cross-reactivity exists. However, these were not tested by the

manufacturer since there is a multitude of benzodiazepines available on the European market and not all could be tested. Therefore 100% cross-reactivity was assumed in the calculations.

When applying device cut-off and cross-reactivity, the box and whisker plot was separated indicating that part of the increase of sensitivity with device cut-offs could be explained by cross-reactivity and part by the higher cut-off.

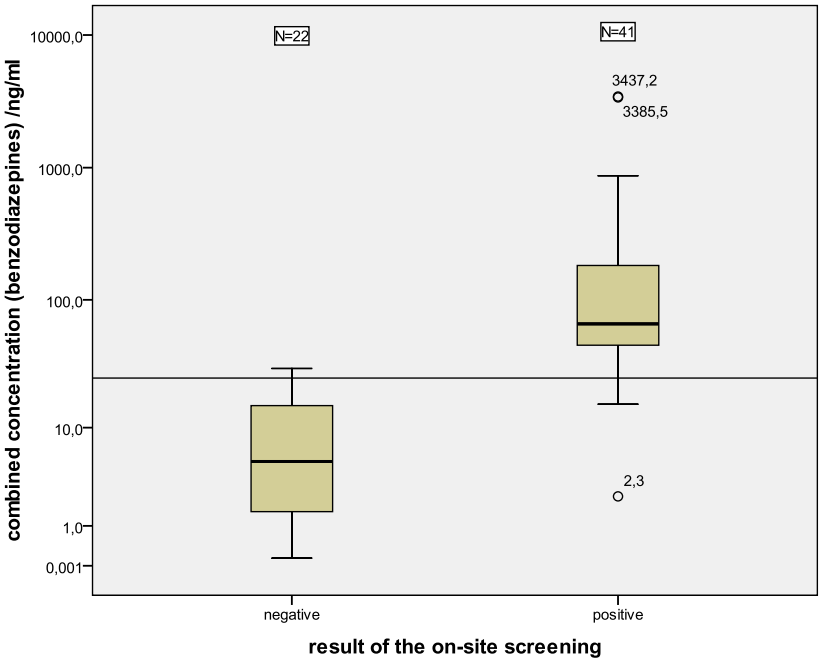

Figure 26. Box and whisker plot for the Rapid STAT Benzodiazepine test, test cut-off and cross reactivity stated by the manufacturer used. 65 cases with 0 ng/ml of any benzodiazepines in ther OF were tested negative for benzodiazepines. These cases are not included in the plot. Horizontal line at 25 ng/ml indicates test cut-off.

Cannabis

50 positive cases for THC were found using the DRUID cut-off (range 1.44-2044.41 ng/ml). When applying the device cut-off, the number of positive cases decreased to 26.

Sensitivity was very low, 32% when using DRUID cut-off and increased to 46% when applying the device cut-off. Specifity decreased from 89.2% when applying the DRUID cut-off to 87.9% with the device cut-off.

The box and whisker plot for THC overlapped, indicating that the test did not have the power to discriminate between low and high concentrations: there was no clear cut-off, even at levels above the DRUID cut-off.

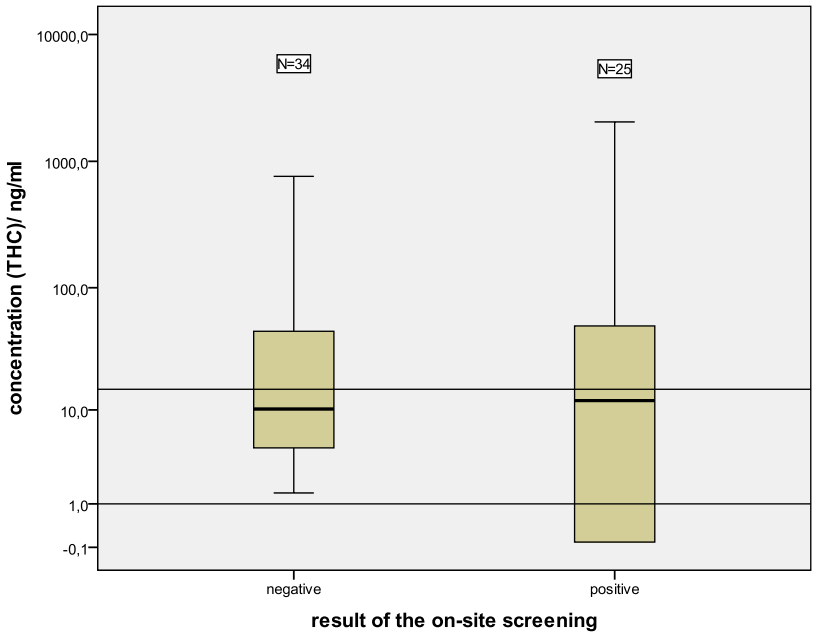

Figure 27. Box and whisker plot for the Rapid STAT cannabis test. 69 cases with 0 ng/ml of THC in their OF were tested negative for cannabis. These cases are not included in the plot. Horizontal lines indicate test cut-off (15 ng/ml) and DRUID cut-off (1ng/ml).

Amphetamines

Due to the low number of positive cases for amphetamine (6 when applying either DRUID or device cut-offs), the result of this analyte should be interpreted cautiously.

When using DRUID cut-offs the following concentration range was found: 25.8-130.9 ng/ml.

Sensitivity was very low (16.7%) and specificity was very high (96.9%) for both DRUID and device cut-offs.

Methamphetamines

No positive results were found for methamphetamine.

5.3.5. Innovacon OrAlert

Table 20. Statistical evaluations for Innovacon OrAlert using DRUID and device cut-offs.

	DRUID cut-offs						Device cut-offs					
	COC	**OPI**	**CAN**	**AMP**	**MAMP**	**PCP**	**COC**	**OPI**	**CAN**	**AMP**	**MAMP**	**PCP**
TP	7	61	3	1	0	0	7	60	2	1	0	0
FP	0	5	0	10	0	0	0	6	1	10	0	0
TN	96	21	83	97	110	110	98	40	104	99	110	110
FN	7	23	24	2	0	0	5	4	3	0	0	0
N of successful tests	110	110	110	110	110	110	110	110	110	110	110	110
Failed devices	15	15	15	15	15	15	15	15	15	15	15	15
Missing analysis	0	0	0	0	0	0	0	0	0	0	0	0
Sensitivity	50%	74%	11%	n.a.	n.a.	n.a.	58%	94%	n.a.	n.a.	n.a.	n.a.
Specificity	100%	81%	100%	91%	100%	100%	100%	87%	99%	91%	100%	100%
Accuracy	94%	75%	78%	89%	100%	100%	95%	91%	96%	91%	100%	100%
Prevalence	13%	76%	25%	2.7%	0%	0%	11%	58%	4.5%	0.9%	0%	0%
PPV	100%	9.6%	100%	n.a.	n.a.	n.a.	100%	17%	n.a.	n.a.	n.a.	n.a.
NPV	90%	99%	29%	n.a.	n.a.	n.a.	92%	95%	n.a.	n.a.	n.a.	n.a.

n.a. not applicable

Note: in 15 cases the test was invalid (=12%)

Cocaine

Due to the low number of positive cases for cocaine (14 when using the DRUID cut-offs, 12 when applying the device cut-off), the results for this analyte should be interpreted cautiously.

When using DRUID cut-offs 12 cases were positive for cocaine (range 10.34-159.79 ng/ml) and 10 cases were positive for benzoylecgonine (range 24.61-138 ng/ml).

Sensitivity was 50% when using DRUID cut-offs and it increased to 58.3% when applying manufacturer's cut-off and cross-reactivity. Specificity stayed at 100%

Based on information provided by the manufacturer, the following cross-reactivity was used:

COC = cocaine + benzoylecgonine

Overall the concentrations in this study population were low.

Opiates

84 positive cases were found using the DRUID cut-offs. 58 cases were positive for morphine (range 20.86-6747.8 ng/ml), 58 positive for 6-acetylmorphine (range 5.3-8723.4 ng/ml) and 44 positive for codeine (range 20.7-588.1 ng/ml). 39 cases were positive for all three analytes. When applying the device cut-off, 64 positive opiate cases were found.

Sensitivy was 73.5% when using DRUID cut-offs, it increased to 93.8% applying the device cut-off. Specificity increased from 80.8% to 87%.

Based on the information provided by the manufacturer, the following cross-reactivity was used:

OPI = morphine + 400% codeine + 160% 6-acetylmorphine

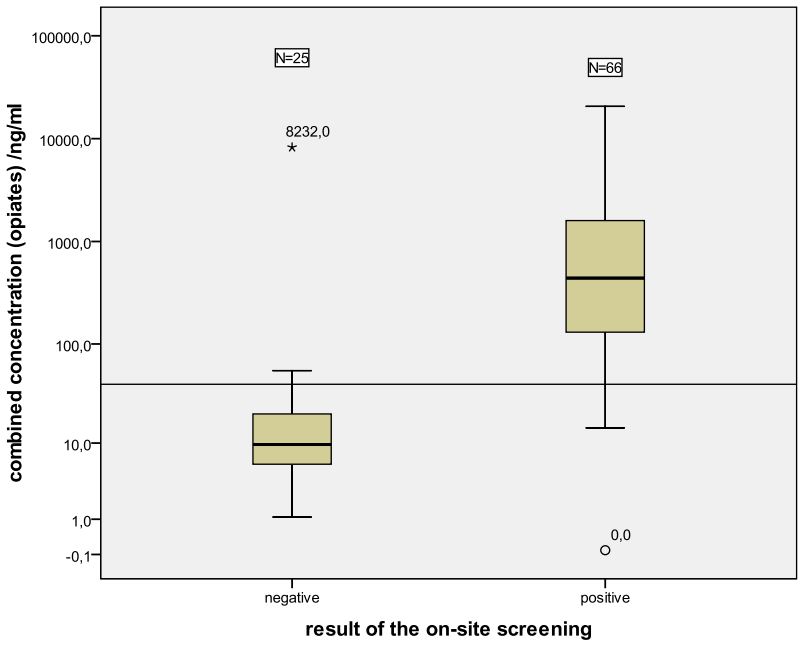

Figure 28. Box and whisker plot for the Innovacon OrAlert opiates test, test cut-off and cross-reactivity stated by the manufacturer used. 19 cases with 0 ng/ml of opiates in their OF were tested negative for opiates. These cases are not included in the plot. Horizontal line at 40 ng/ml indicates test cut-off.

Cannabis

27 positive cases were found using the DRUID cut-off and sensitivity was very low (11.1%). The concentration range was 1.4-539.1ng/ml. Only 5 positive cases were found when applying the device cut-off, so no sensitivity was calculated.

Specificity decreased from 100% using the DRUID cut-off to 99% when using device cut-off.

Amphetamines

Due to the low number of positive cases for amphetamine (3 using DRUID cut-offs and 1 when using the device cut-off), the results for this analyte should be interpreted cautiously. Hence no sensitivity was calculated. Specificity was 91%. The concentration range was 40.6-88.9 ng/ml.

Methamphetamines

No positive results for methamphetamine were found

Phencyclidine

No positive results for phencyclidine were found

5.3.6. Clinical Signs of Impairment checklist evaluation

To check the significance of the CSI tests, the parameters tested for were reduced to those that were positive in at least 3 out of 250 test subjects. This selection led to a reduction to 13 (out of 28) parameters.

A statistical test (Fisher's exact test) was used to test for associations between the CSI checklist parameters and the presence of substances in oral fluid. Results are shown in Table 21. If $p < 0.05$, this means that there was a significant correlation: the presence of these CSI checklist parameters was a good indicator for the presence of a psychoactive substance.

It was clear that most parameters did not correlate significantly with drug intake. The pupil tests seemed to be the best predicting parameters, especially for amphetamine and THC.

Remarkably, some correlations were found between parameters and drugs where no correlation was expected. This can possibly be caused by the presence of combination use of drugs in a lot of subjects.

Table 21. Correlation between CSI parameters and presence of psychoactive drugs.

Parameters	Characteristic for (drug)	P-value for characteristic drug	Remarks
Unstable composure	OPI	0.597	
Dizzy, sleepy	OPI	0.759	Significant for AMP (P 0.009)
Unclear speech	THC	0.695	
Trembling	AMP, MDMA COC	0.590 1.000	
Shaking leg	AMP, MDMA COC	0.357 0.514	
Excited, aggressive behaviour	AMP COC	1.000 0.062	Significant for COD (P 0.044)
Shaking eye lids	AMP, MDMA COC	0.091 0.581	
Sniffing	COC	0.063	
Excessive sweating	OPI	1.000	Significant for COC (P 0.036)
Small pupils (<3mm)	OPI	0.477	
Large Pupils (>6.5mm)	AMP, MDMA COC THC	**0.009** 1.000 1.000	
Nystagmus test	THC OPI	**0.041** **0.028**	Significant for AMP (P 0.047)
Pupil reaction to light	AMP, MDMA COC THC	**0.007** 0.262 **0.002**	Significant for OPI (P 0.001)

It was observed that the concentrations that led to a significant correlation were often a lot higher than the DRUID cut-offs. This meant that the CSI checklist parameters were true positive in subjects who either took drugs very recently or who took drugs in high quantities.

In general, the CSI tests evaluated here correlated very badly with drug presence.

This was most likely caused by the choice of the study population: most volunteers had used a combination of drugs and had developed a tolerance because of long-term use. Moreover, some subjects had not used drugs very recently.

5.3.7. Application to recent Belgian legislation

Recently new legislation for DUID was passed in Belgium using on-site test devices for drugs of abuse in oral fluid. Oral fluid will also used as matrix for confirmation analysis, with a 10 ng/ml cut-off for THC.

Using this cut-off, the test devices gave the following results; the distinction of performance for THC as mentioned above is very clear: sensitivity for DrugTest 5000 was 80.0% while other tests scored between 20 and 43.3% sensitivity (Table 22). PPV and NPV were calculated as described in 4.4.3.

Table 22. Statistical evaluation of THC detection using cut-offs from Belgian legislation.

	DrugTest 5000	Cozart DDS	Rapid STAT	OraLab6	OrAlert
TP	20	11	13	18	3
FP	5	0	12	2	0
TN	106	99	91	159	97
FN	5	28	17	70	10
Sensitivity	80%	28%	43%	21%	23%
Specificity	96%	100%	88%	99%	100%
Accuracy	93%	80%	78%	71%	91%
PPV	98%	100%	91%	98%	100%
NPV	63%	33%	36%	31%	32%

5.4. Discussion

The evaluation of the Varian OraLab6 was performed in a different population than the other evaluations. Especially cocaine and amphetamine were more prevalent in the OraLab6 population. Hence results for cocaine and amphetamine were easier to interpret for OraLab6 then for the other tests.

In general, specificity was high for all drug classes for each test. The most significant specificity issues found were cross-reactivity of the PCP-assay with venlafaxine in the OraLab6 and an over-sensitivity of the opiates assays.

The experiments on Varian OraLab6 for PCP with spiked oral fluid clearly demonstrated that venlafaxine and o-desmethylvenlafaxine could both cause false positive results for PCP, even at therapeutic concentrations. Screening results for PCP in oral fluid should therefore always be confirmed.

The problems with opiates for the various devices could be explained by either cross-reactivity of high levels of methadone (e.g. cross-reactivity of 100,000 ng/ml methadone on Dräger DrugTest 5000) or by the presence of opiates, which were not included in the confirmation analysis (pholcodine, hydromorphone, dihydrocodeine).

As in previous studies, sensitivity problems were encountered for a lot of drugs. However for certain tests sensitivity scores were moderate to high or very high:

- Opiates: Dräger DrugTest 5000 (88.5%). Cozart DDS (83.3 %,). Rapid Stat (90.5%)
- Amphetamines: DrugTest 5000 (75.0%). However, only a very limited number of positive samples was obtained for this test, and box and whisker plot data indicated that the actual cut-off was higher (Figure 21: concentrations (excluding 0 ng/ml, n=127) of amphetamine in negative (0) and positive (1) screenings using Dräger DrugTest 5000.

  Applying the manufacturer's cut-offs and cross-reactivities resulted in most cases in higher sensitivity and specificity, mostly because of the higher cut-offs proposed by manufacturers.

For detection of cannabis use, a clear distinction in performance from the different devices was seen:

- OraLab6 and OrAlert both had very low sensitivity
- Rapid STAT: sensitivity was very low, and the test did not make a clear discrimination between low and high concentrations.
- Cozart DDS: sensitivity was very low, but the test discriminated between low and high concentrations. (discriminating cut-off around 100 ng/ml)
- Dräger DrugTest 5000: sensitivity was higher than any other test and the test discriminated between low and high concentrations (discriminating cut-off around 10 ng/ml).

Further studies with the remaining saliva in the collector devices from the on-site tests could indicate whether the sensitivity problems are caused by the sampling procedure (e.g. no extraction of THC from the oral cavity) or by sensitivity problems of the test itself (e.g. low sensitivity of the immunoassay).

5.5. Acknowledgements

This work could not have been possible without the collaboration of MSOC Ostend, MSOC Ghent, UPSIE Ghent and De Sleutel in Wondelgem.

Thanks to the students who helped to collect the data.

Special thanks to Elke Raes for her contribution to the early phase of the study.

Co-funding was obtained from the National Lottery of Belgium.

The researchers are grateful to all the manufacturers for their technical support and for placing the test devices at the studies disposal free of charge.

6. Country report - Finland

Anna Pehrsson, Tom Blencowe, Kaarina Langel, Charlotta Engblom and Pirjo Lillsunde
Alcohol and Drugs Analytics Unit, National Institute for Health and Welfare

6.1. Introduction

The Finnish part of the study was coordinated by the National Institute for Health and Welfare. The study was carried out in close co-operation with the Police Department of the Ministry of the Interior, the Traffic Police of Helsinki and the Police Surgeon Station (Department of Forensic Medicine, Helsinki University). The aim of the study was to gain information on the analytical performance of the on-site tests.

In the Finnish study, two on-site devices were tested: DrugWipe 5^+ and Rapid STAT. Altogether 268 tests were done. All tested persons gave an oral fluid sample. The results of the on-site tests were evaluated based on the gas chromatography-mass spectrometry (GC-MS) analysis result of the oral fluid sample.

6.2. Materials and Methods

6.2.1. Sample collection

Altogether 221 subjects participated in the study. Participation in the study was voluntary. An ethical approval for the study was obtained from the coordinative ethical committee of Hospital District of Helsinki and Uusimaa. A written consent was obtained from all of the participants. The subjects were tested with the DrugWipe 5^+ and/or Rapid STAT on-site screening devices. Non-suspect subjects were recruited to the study from amongst the participants of the epidemiological roadside study of drivers, DRUID WP2, and also from amongst the personnel of the Alcohol and Drug Analytics Unit of the National Institute for Health and Welfare. These tests were performed under supervision by a researcher. People suspected of driving under the influence of drugs or medicines were asked to take part in the study by the apprehending police officers, who also supervised the on-site drug screening. In addition, 21 patients from a cooperating rehabilitation clinic for drug addicts were recruited to the study. These subjects were recruited and tested by the nurses at the clinic. All testing for DrugWipe 5^+ was done between 5.4.2008-22.8.2009 and for Rapid STAT between 16.9.2008-22.8.2009.

An oral fluid (OF) sample was collected from all the subjects for confirmation analysis. The OF sample was collected with the StatSure Saliva Sampler device. For the people who were also participating in the roadside study or from the personnel of the Drug Analytics Unit the OF samples were collected by researchers for the DRUID project. The participants suspected of DUI were taken to the Police Surgeon Station for blood sampling and clinical evaluation by a physician as part of normal police procedure. The on-site test and collection of the OF sample was performed by the police officer or by the physician. OF samples from the rehabilitation clinic patients were taken by clinic nurses. Informed consent was obtained from all of the subjects recruited to the study. All OF samples were frozen (-20°C) until analysis.

For the CSI evaluation results a Finnish police sobriety test sheet (Annex 3), used as normal police procedure, with some additional observations and questions necessary for the DRUID study, was filled in. For suspected DUI drivers this was performed by the police officer supervising the on-site tests, who are experienced in these procedures as part of their normal duties. The participants recruited from the roadside sessions and from the Alcohol and Drug Analytics Unit were assessed by researchers for the DRUID project. The researchers were not formally trained for the observation of signs of impairment. The CSI observations were not performed for the patients from the rehabilitation clinic.

6.2.2. Analytical method

The sample preparation procedure for OF samples is presented in Figure 29. Liquid-liquid extraction (LLE) and/or solid phase extraction (SPE) were used for analyte extraction. Mixed mode MCX columns (Oasis) were used for the SPE. The samples were then silylised and analysed with gas chromatography-electron impact ionisation/mass spectrometry (GC-EI/MS) or gas chromatography-negative chemical ionisation/mass spectrometry (GC-NICI/MS). OF samples were thawed and 1 ml of sample was pipetted to test tubes. In some cases very little OF was collected and hence it was not possible to obtain 1 ml of sample for analysis. In these cases the volume of pipetted OF was noted and taken into account when calculating the results.

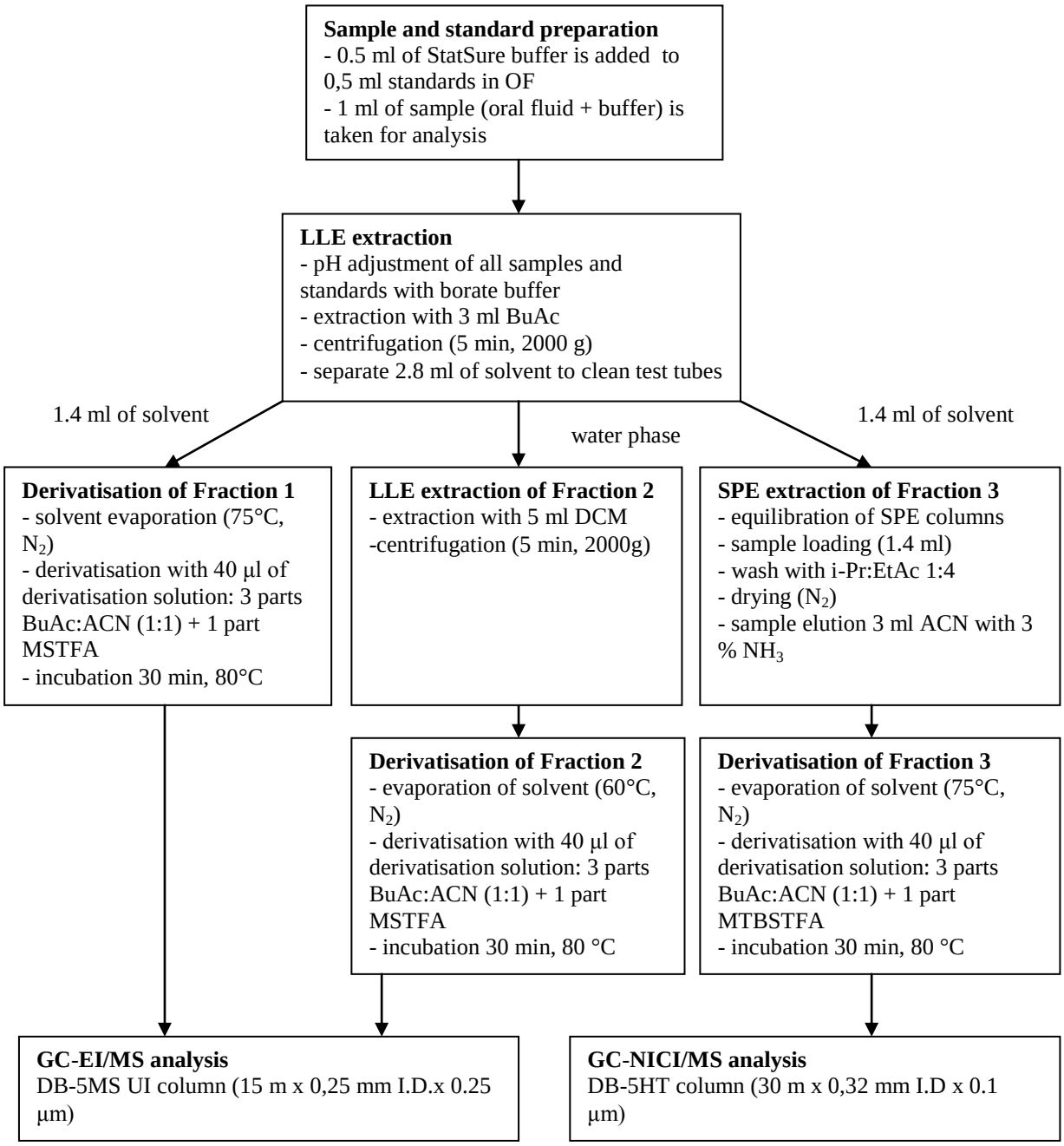

Figure 29. Flow chart of the sample preparation method for OF analysis. Fraction 1: illicit drugs and medicinal drugs other than benzodiazepines, Fracton 2: benzoylecgonine, Fraction 3: benzodiazepines.

Standards for analysis were made by spiking 0.5 ml of blank OF with stock solution containing all analytes. A blank OF sample was included in every run. For the standards preparation, blank OF was collected from laboratory personnel, frozen and thawed. Deuterated analogues of the analytes were used as internal standards. A lower limit of quantitation (LLOQ) sample was run in every analysis as a quality control sample. For samples with concentrations higher than the upper limit of quantitation (ULOQ), it was not possible to make a subsequent analysis with a dilution because of the low sample volume available for analysis. In these cases, the results were extrapolated from the calibration curve.

The OF samples were analysed with Agilent Technologies GC-MS systems. The analysis on EI mode were carried out with 6890/5975 equipment with helium as the carrier gas and the analysis on NICI mode were carried out with a 6890N Network GC System with 5970 Inert Mass Selective Detector with hydrogen as the carrier gas.

Method validation

The OF analysis method has been fully validated. Validation results for all analytes are listed in Table 46-Table 48 in Annex 4. For linearity experiments, six replicates at six concentration points were analysed in order to obtain a linear calibration model. The inverse of the squared concentration was used as a weighting factor. For precision and accuracy calculations, three concentration levels were chosen (LLOQ, medium concentration, ULOQ) for testing. Three replicates were prepared and measured against a calibration curve on five different days. From these results, precision and accuracy were calculated. Extraction recovery was determined by measuring normally prepared calibration standards and blank samples that had methanol/water based stock solution added to them after the extraction ("extraction controls"). The volume of added stock solution was adjusted to equate with full recovery. Extraction recovery was determined at medium concentration level. For selectivity experiments, 10 blank OF samples from different persons were analysed to see if there were any selectivity problems. The selectivity of the methods was found to be very good for all analytes.

6.2.3. External quality control

The laboratory participated in an OF external quality control program organized by RTI International (NC, USA) for all the DRUID project toxicology laboratories. Based on the results the performance of all the OF analytical methods was at an excellent level.

6.2.4. Prevalence of substances among people suspected of DUI

The prevalence of the substances in question among DUI suspected people in year 2008 were used for PPV and NPV calculations. These were taken from the database of the Drug Analytics Unit. Altogether 4419 cases were investigated in that year. It should be noted that these prevalences are calculated from blood sample results and are not based on OF. The most common findings were sleep-inducing and sedative substances 59 % and amphetamines 56 %. Other findings were cannabis 20 %, methamphetamine 8.0 %, morphine 0.9 % and cocaine 0.7 %.

6.2.5. Study of DrugWipe 5 results in normal police procedure

The various versions of the DrugWipe 5 on-site drug screening device have been routinely used by the Finnish traffic police for several years. For investigation of DUI cases, the Drug Analytics Unit does not require the results of the on-site test, but in many cases the results of the DrugWipe 5 test have been reported to the laboratory. These results were taken from the laboratory database and the performance of the screening devices was evaluated according to the relevant whole blood (WB) analysis results. The WB confirmation samples were routinely analysed at the Drug Analytics Unit for the police.

The data is restricted to the cases that have been shown to be positive for some drug group according to the DrugWipe 5 device, hence there are no data for cases in which all the test strips were negative. In addition, the data for positive cases are not fully comprehensive, being limited to only the cases which are actually reported. For these reasons, as well as the use of a different confirmation sample

matrix, these reults cannot be directly compared to the evaluation of devices in Task 3.2. The cases evaluated here are from the 1st of July 2007 up until the 31st of December 2008. During this period 1942 cases that the police had gathered and that gave a positive result from the device were entered to the database. The WB samples were analysed with the methods described in (14-19). The analytical findings are interpreted as results according to the limit of quantitation (LOQ) for the methods used.

6.3. Results

Sensitivity, specificity, accuracy, positive predictive value (PPV) and negative predictive value (NPV) calculations for both on-site screening devices were made based on the device and the OF analysis results. Where there is a sufficient number (N≥6) of positive and non-zero concentration negative test results for a specific substance (or substance group) box and whisker plots are presented.

6.3.1. Study population

Altogether 221 subjects were recruited to the study. 136 subjects were tested with the DrugWipe 5+ device and 132 with the Rapid STAT. For the Rapid STAT, 21 suspect cases were recruited from a rehabilitation clinic. For these cases, the test was performed by a clinic nurse and the CSI checklist was not performed. The distribution of suspect and non-suspect cases was similar for both devices.

6.3.2. DrugWipe 5+

Altogether 136 cases were tested with the DrugWipe 5+ device. The results of the test evaluation are shown in Table 23.

Table 23. Results of the test evaluation according to the DrugWipe 5+ cut-offs. Note: PPV and NPV calculated using DUI suspect prevalences from 2008.

	According to DRUID cut-off				According to test cut-off			
	AMP	**CAN**	**OPI**	**COC**	**AMP**	**CAN**	**OPI**	**COC**
TP	33	9	0	0	33	5	0	0
FP	5	4	0	0	5	8	0	0
TN	92	109	134	123	92	118	133	123
FN	5	12	1	0	5	3	2	0
N of successful tests	135	134	135	123	135	134	135	123
Failed	0	1	1	1	0	1	1	1
Missing analysis	1	1	0	12	1	1	0	12
Total N of cases	136	136	136	136	136	136	136	136
Sensitivity	87%	43%	n.a.	n.a.	87%	63%	n.a.	n.a.
Specificity	95%	96%	100%	100%	95%	94%	100%	100%
Accuracy	93%	88%	99%	100%	93%	92%	99%	100%
Prevalence	28%	16%	0.7%	0%	28%	6.0%	1.5%	0%
PPV	96%	75%	n.a.	n.a.	96%	71%	n.a.	n.a.
NPV	85%	87%	n.a.	n.a.	85%	91%	n.a.	n.a.

n.a. - calculation not applicable

Amphetamine type stimulant drugs

Altogether 38 cases (28%) had amphetamine-type stimulant (ATS) drugs detected in their OF. Amphetamine was found in all the cases that were positive for ATSs. The concentration range for amphetamine was 231-82500 ng/ml. 11 cases (8.1%) contained methamphetamine, range 42.6-36700 ng/ml. In one case, MDA (281 ng/ml) and MDMA (4850 ng/ml) were found in addition to amphetamine. MDEA was not found in any of the samples.

Cross reactivity was calculated using:
AMP = c(amphetamine) + 200% c(methamphetamine) + 200% c(MDMA) + 200% c(MDA) + 77% c(MDEA)

The sensitivity of the amphetamine test was high and specificity, as well as accuracy, very high. The PPV for the amphetamine test was very high (96 %). NPV was a little lower with 85%.

As well as 92 TN cases with 0 ng/ml of ATSs, there were 5 FN cases for this device, all of which contained only amphetamine from the ATS group of drugs. The concentration range for these 5 cases was 244-3620 ng/ml. This concentration range is clearly a lot higher than both DRUID and manufacturer cut-offs, but still relatively low in comparison to the overall range of concentrations found for amphetamine. The ATS concentration, taking into account cross reactivity, for more than 80% of the TP cases was higher than 3620 ng/ml (the maximum amphetamine concentration found in the FN cases). There were also 5 FP cases observed, each of which contained no ATS drugs.

Cannabis

There were 21 (16%) cases which were Δ^9-THC positive in the OF sample, the concentration range was 1.3-2020 ng/ml. Most of the positive cases (16 cases, 76 % of positive cases) contained concentrations below 40 ng/ml.

Sensitivity of the cannabis test was low when using the device cut-off and very low when the DRUID cut-off was utilised. However, specificity was very high for both cut-offs. The PPV and NPV for the cannabis test at the test cut-off were 71 % and 91 % respectively. Using the DRUID cut-offs PPV and NPV were 75 % and 87% respectively.

Only Δ^9-THC was measured in the study hence no cross reactivity was taken into account. This means both cut-offs can be shown on the box and whisker plot. The box and whisker plot for cannabis (Figure 30) shows that there is no real difference in Δ^9-THC concentrations between the cases that were tested negative with the device and the cases that tested positive. The box and whisker plot for cannabis shows that although the cases which tested negative are largely below the device cut-off (30 ng/ml) for Δ^9-THC – and all above the DRUID cut-off (1 ng/ml) a similar range of Δ^9-THC concentrations can also give a positive test result.

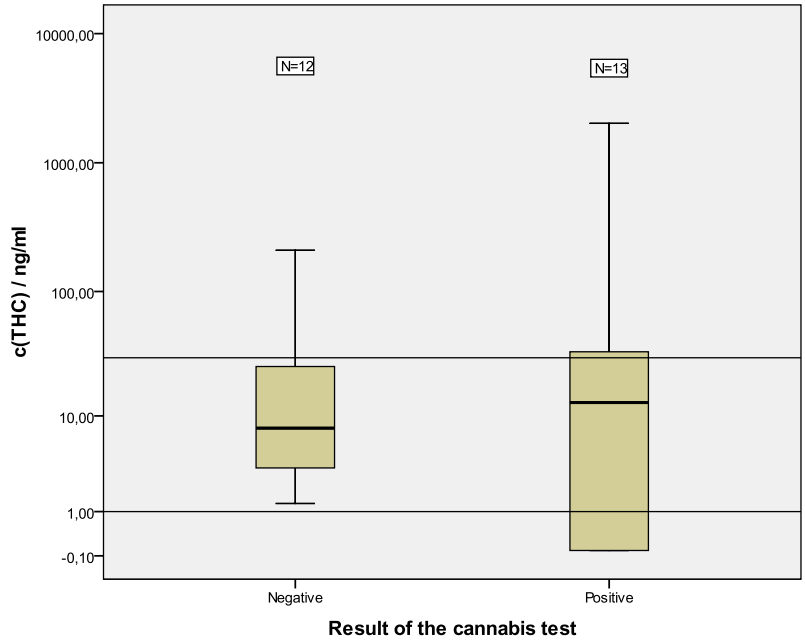

Figure 30. Box and whisker plot for cannabis test. 109 cases with 0 ng/ml of Δ^9-THC in their OF tested negative for cannabis. These cases are not included in the plot. Horizontal lines indicate test cut-off (30 ng/ml) and DRUID cut-off (1 ng/ml).

The manufacturer of the test launched a new version of the DrugWipe 5^+ test with an enhanced test strip for cannabis during the testing period. In this study, 12 tests with the improved cannabis test strip were made. All of the cases tested negative for cannabis. However, one case contained 6.4 ng/ml of Δ^9-THC.

Opiates

Two opiate positive cases were detected. One contained 15.9 ng/ml of morphine and 2.47 ng/ml of the metabolite 6-MAM. The other OF sample contained 105 ng/ml of codeine. 6-MAM by itself was not found in any of the cases. Four cases contained buprenorphine, in one of which norbuprenorphine was also found, however, at a concentrations likely to occur in oral fluid these two substances are not detected by the DrugWipe 5^+ device. In this study, cases containing only buprenorphine and/or norbuprenorphine were not considered as opiate positive cases.

Cross reactivity was calculated using: OPI = c(morphine) + c(codeine) + c(ethylmorphine)

When interpreting the device result according to the test cut-off, both of the cases were FN. When utilising the DRUID cut-off (20 ng/ml, slightly higher that the test cut-off), the morphine positive case was interpreted as TN and only the codeine case was FN. Specificity of the opiate test was excellent. Sensitivity, PPV and NPV calculations were not made due to the low number of positive cases found.

Cocaine

None of the cases contained cocaine or the metabolite benzoylecgonine. All of the cases screened negative with the DrugWipe 5^+ test. PPV and NPV calculations were not made due to the fact that there were no positive cases.

6.3.3. Rapid STAT

Altogether 132 cases were tested with the Rapid STAT device. The results of the test evaluation are shown in Table 24.

Table 24. Results of the test evaluations according to the Rapid STAT and DRUID cut-offs.

	According to DRUID cut-off						According to test cut-off					
	AMP	**MAMP**	**CAN**	**OPI**	**COC**	**BZO**	**AMP**	**MAMP**	**CAN**	**OPI**	**COC**	**BZO**
TP	18	4	13	3	0	23	18	4	11	3	0	12
FP	0	4	12	0	0	3	0	4	14	0	0	14
TN	100	120	98	124	118	91	100	120	102	124	118	100
FN	7	0	6	0	0	11	7	0	2	0	0	2
N of successful tests	125	128	129	127	118	128	125	128	129	127	118	128
Failed	6	3	3	5	5	4	6	3	3	5	5	4
Missing analysis	1	1	0	0	9	0	1	1	0	0	9	0
Total N of cases	132	132	132	132	132	132	132	132	132	132	132	132
Sensitivity	72%	n.a.	68%	n.a.	n.a.	68%	72%	n.a.	85%	n.a.	n.a.	86%
Specificity	100%	97%	89%	100%	100%	97%	100%	97%	88%	100%	100%	88%
Accuracy	94%	97%	86%	100%	100%	89%	94%	97%	88%	100%	100%	88%
Prevalence	20%	3.1%	15%	2.4%	0%	27%	20%	3.1%	10%	2.4%	0%	11%
PPV	100%	n.a.	61%	n.a.	n.a.	97%	100%	n.a.	64%	n.a.	n.a.	91%
NPV	74%	n.a.	92%	n.a.	n.a.	68%	74%	n.a.	96%	n.a.	n.a.	81%

n.a. - calculation not applicable

Amphetamine type stimulant drugs

27 cases (22%) were positive for ATSs. All of these positive cases contained amphetamine. The concentration range for amphetamine was 231-44700 ng/ml. Methamphetamine was also found in four of these cases (3.1 %), with a concentration range of 816-36700 ng/ml. MDA, MDMA or MDEA were not found in any of the samples. The test cut-off for both amphetamine and methamphetamine tests is 25 ng/ml, which is the same as the DRUID cut-off. Hence only one interpretation of the test results needs to be made. The device's amphetamine test failed for two of the amphetamine positive cases.

Cross reactivity for amphetamine and methamphetamine were calculated using:
AMP = c(amphetamine) + 50% c(MDA) + 0.5% c(MDEA) + 1% c(MDMA)
MAMP = c(MAMP) + 50% c(MDMA) + 5% c(MDEA) + 0.05% c(MDA)

Amphetamine test

Sensitivity of the amphetamine test was moderate. Specificity and PPV on the other hand were excellent at 100 %. NPV was not on the same level and only reached 74 %. The box and whisker plot (Figure 31) shows a similar situation to that seen for the DrugWipe 5^+ test: cases with concentrations amounting to several hundreds or even thousands of ng/ml of ATSs still gave a negative screening result with the device. However, the '50 % boxes' for positive and negative test results appear to be distinct.

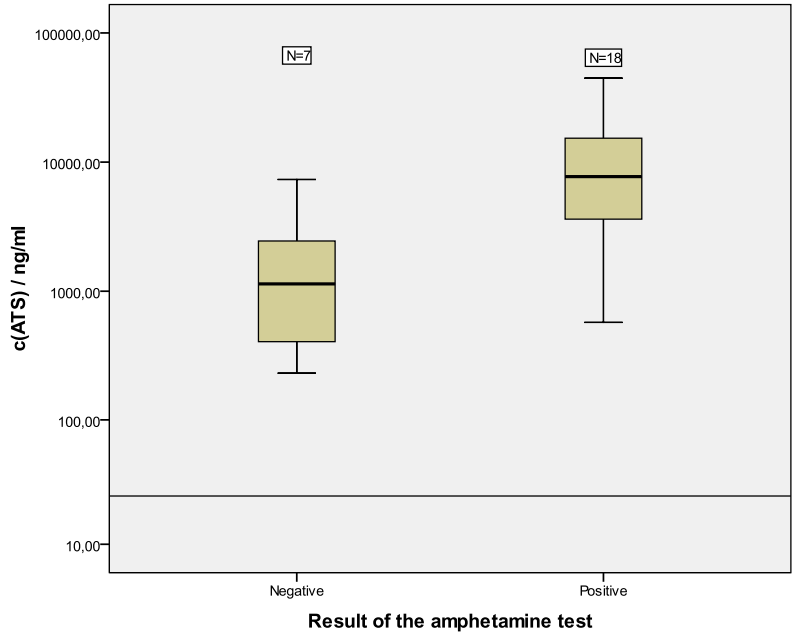

Figure 31. Box and whisker plot for amphetamine test. 100 cases with 0 ng/ml of amphetamine in their OF tested negative with the test. These cases are not shown on the plot. Horizontal line at 25 ng/ml indicates the test and also DRUID cut-off.

Methamphetamine test

There were only four positive cases (3.1%) for methamphetamine. Four negative cases also gave a positive result for methamphetamine with the test. Three of these cases contained very high concentrations of amphetamine (44700 ng/ml, 21000 ng/ml and 15300 ng/ml). The Mavand company have reported 1000 ng/ml of amphetamine to be the highest concentration that has not showed cross reactivity. Hence it can be speculated that these three false positives are a result of high amphetamine concentrations, although the Mavand company does not report amphetamine as a cross reacting compound for the Rapid STAT methamphetamine test. Sensitivity, PPV and NPV were not calculated due to the low number of positive cases. Specificity was very high.

Cannabis

19 cases (14 %) were positive for Δ^9-THC. The concentration range for these positive cases was 1.5-2020 ng/ml. 10 (53 %) of the positive cases were below 40 ng/ml and 12 (63 %) of the positive cases below 50 ng/ml.

Sensitivity was high when the manufacturer's cut-off was used but only low with the DRUID cut-off. Specificity was also high with both cut-offs. The PPV for cannabis was approximately 60% and NPV above 90% with the both of the cut-offs.

The box and whisker plot for cannabis (Figure 32) shows that the cases which have given a negative screening result are frequently of a much higher concentration than the test cut-off (15 ng/ml). Again, similarly to the results observed for the DrugWipe 5^+ test, the range of Δ^9-THC concentrations which can give a positive test result is not distinct from that which gives a negative screening result.

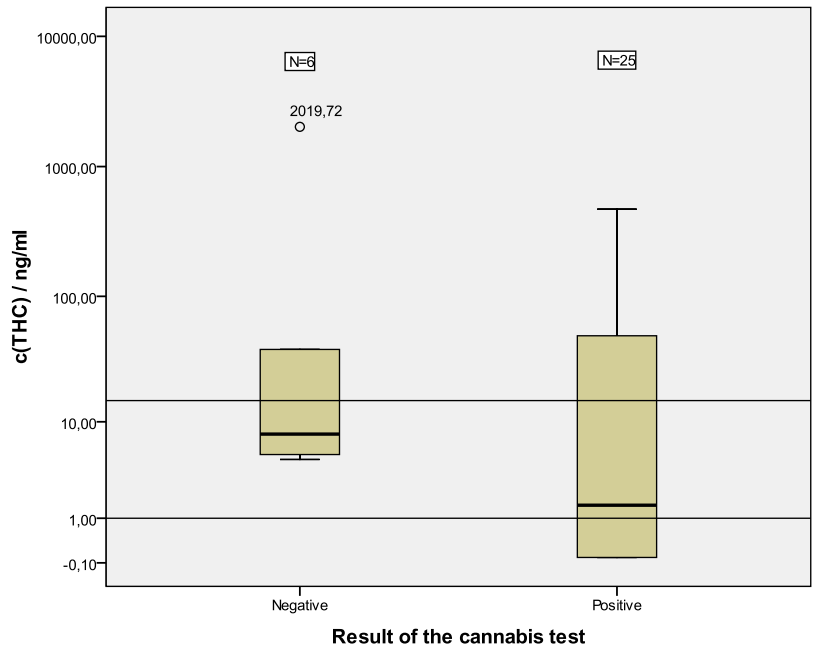

Figure 32. Box and whisker plot for the Rapid STAT cannabis test. 98 cases with 0 ng/ml of THC in their OF tested negative for cannabis. These cases are not included in the plot. Horizontal line at 1 ng/ml indicates DRUID cut-off and at 15 ng/ml indicates the device cut-off.

Opiates

Three positive opiate cases were identified. In two of these cases, codeine was the only opiate found (107 ng/ml and 105 ng/ml respectively). In the other case, morphine (30.4 ng/ml), codeine (1490 ng/ml) and also buprenorphine (8.8 ng/ml) were found. All of these cases tested positive with the Rapid STAT device. Due to the low number of positive opiate cases, only specificity and accuracy were calculated. They both were excellent.

Cross reactivity was calculated using: OPI = c(morphine) + 50% c(MAM) + 333.3% c(codeine).

Cocaine

None of the cases included in the study contained cocaine or benzoylecgonine. All cases tested negative with the Rapid STAT device. Specificity and accuracy of the test were excellent. Sensitivity, PPV and NPV were not calculated.

Benzodiazepines

Altogether there were 37 cases (28 %) which had benzodiazepines in their OF samples (according to DRUID cut-off values). The most common benzodiazepine findings are shown in Table 25. The device's benzodiazepine test failed for three of these positive cases.

Table 25. Benzodiazepine findings in the OF samples used for the Rapid STAT evaluations.

Analyte	N	Range / ng/ml	Average / ng/ml	Median / ng/ml
Nordiazepam	22	1.0-55.8	10.9	4.6
Clonazepam	13	2.3-754	96.2	10.6
Alprazolam	10	1.2.2590	277	14.9
Diazepam	9	6.5-16800	1910	15.0
Oxazepam	6	5.7-560	113	11.6
Temazepam	6	10.4-208	54.7	24.3

Each of the substances lorazepam, midazolam and α-OH-alprazolam were found twice, from six different samples. Concentrations for these cases were: lorazepam 1.6 and 44.7 ng/ml, midazolam 12.8 and 47.8 ng/ml and α-OH-alprazolam 1.0 and 33.3 ng/ml.

Cross reactivity was calculated using:
BZO = 400% c(diazepam) + c(alprazolam) + 200% c(flunitrazepam) + 4% c(lorazepam) + 0.02% c(clonazepam) + c(nordiazepam) + c(oxazepam) + 200% c(nitrazepam) + 200% c(temazepam) + 20% c(triazolam)

There were several cases in which benzodiazepines were found below the DRUID cut-off values. Very low concentrations of oxazepam and diazepam were detected in 12 and 11 cases respectively. Temazepam was detected at very low levels in four cases. Lorazepam and chlordiazepoxide were both detected in one case each. However, in all but one case other benzodiazepines were detected with levels above the DRUID cut-off values.

Sensitivity and specificity of the benzodiazepine test were high when using the manufacturer's cut-off. When utilising the DRUID cut-offs, sensitivity was low but the specificity became very high. This can easily be explained by more cases interpreted as negative when cut-offs are higher. PPV was very high with both DRUID and device cut-offs. NPV was high when the device cut-off was used and cross reactivity taken into account, and moderate with the DRUID cut-offs.

The box and whisker plot (Figure 33) shows, that there is no real separation of low and high concentrations.

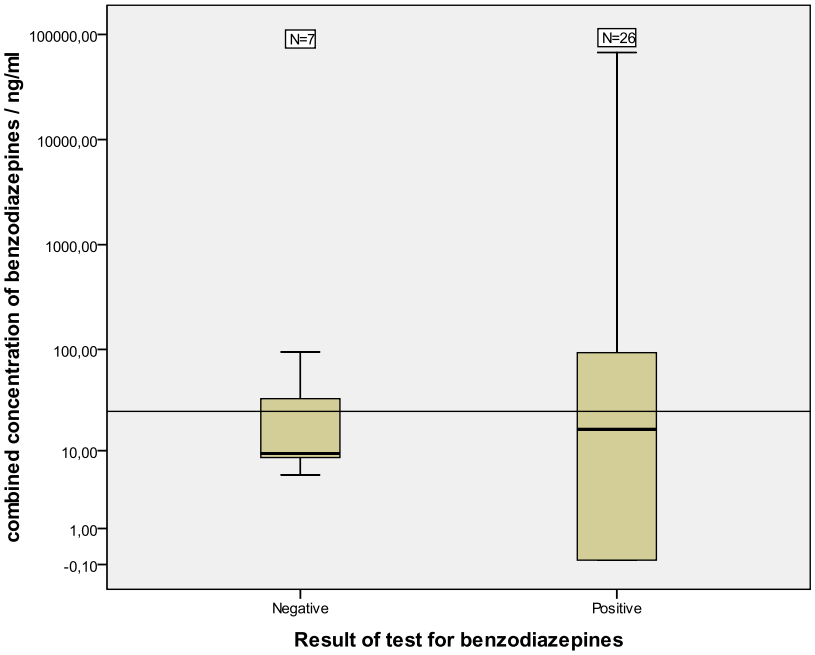

Figure 33. Box and whisker plot for the Rapid STAT benzodiazepine test. Manufacturer cut-off and cross reactivity are used. 95 cases with 0 ng/ml of any benzodiazepines in their OF were tested negative for benzodiazepines. These cases are not included in the plot. Horizontal line at 25 ng/ml indicates test cut-off.

6.3.4. Checklist for signs of impairment evaluation

Results of the CSI evaluation of drivers were collected from 142 non-suspect drivers and 39 drivers who were suspected of DUI. The suspect drivers included to this study were recruited at a point at which suspicion of use was very stong, rather than at the time of initial suspicion. It was impossible to analyze the CSI assessments based on the final judgement on the use of impairing substances since in the majority of suspect cases (n = 30) no final judgement was recorded. The frequency of positive observations for each CSI in the study is shown in Table 26.

Table 26. Frequency of CSI symptoms in study population.

Observation	Negative cases	All cases
Unsteady on one's feet, swaggering	0	13 (1)
Uncordinated movements	0	5
Drowsy, sleepy	1	4 (1)
Euphoria	0	2 (1)
Not understanding instructions	0	0
Incoherent speech	1	9 (1)
Chattering	2	6
Slurred speech	1	11 (1)
Low, rasping voice	0	0
Scratching one's face	0	1
Trembling	1	1 (1)
Shaking leg	0	0
Excited, aggressive behaviour	1	3 (1)
Bloodshot eyes	0	8 (1)
Red Nostrils	0	4
Trembling eyelids	0	1
Sniffing	0	6
Undue perspiring	0	5 (1)
Swallowing	0	1
Smell of hash	0	1
Pinpoint pupils	0 (1)	6 (16)
Dilated pupils	2 (1)	10 (16)
Nystagmus test	0 (142)	2 (172)
Test pupil reaction to light	0 (140)	1 (171)
Total cases	142	181

(x) = no observation made

For the substance positive cases the number of individual symptoms observed was assessed as a possible measure for determining suspicion of drug use. This was done by classifying cases into two groups (Table 27). One group contained those cases which exhibited signs which are nearly always associated with impairing drug use (i.e. smell of hash, pinpoint pupils, slowed reaction of pupils to light) or which exhibited more than 3 of the other symptoms checked for, which might also have been caused by other chronic or acute factors than impairing drug use. The other group was those cases with three or fewer symptoms and no highly suspicious symptoms.

Table 27. Classification of substance positive cases according to symptoms observed.

Positive substance findings	No. of cases with:	
	3 symptoms or less and no suspicious symptoms	4 or more and/or with highly suspicious symptoms
cannabis only	1	0
benzodiazepines only	1	4
amphetamines only	2	1
cannabis + benzodiazepines	1	2
cannabis + amphetamines	1	3
opiates + benzodiazepines	2	0
opiates + amphetamines	2	1
benzodiazepines + amphetamines	5	1
cannabis + opiates + benzodiazepines	0	1
cannabis + benzodizepines + amphetamines	4	4
opiates + benzodiazepines + amphetamines	2	1
Total	21	18

It was notable that 10 of the 27 cases assessed for CSI evaluation which contained amphetamines displayed none of the typical symptoms associated with these drugs (Table 8). Likewise 3 out of 17 cannabis positive cases and 1 of 9 opiate positive cases exhibited no typical symptoms for these substances. There were no cocaine positive cases included to the CSI study.

6.3.5. Performance of the DrugWipe 5 in normal police procedure

Demographics of the study population

The age range of the tested people was 15-62 years (mean 32.4 years, median 31 years). Gender distribution is skewed; the study population consists of 1703 (88%) men and 238 (12%) women. The gender information was missing in one case. (Total study group size: 1942) Most of the subjects were car drivers (91%). Other categories of encountered vehicles were vans, mopeds, motorcycles and lorries but these vehicle groups were each represented in only from approximately 1-3 % of the cases. There were three people apprehended in water-borne traffic and one pedestrian.

Performance of the DrugWipe 5

The results for sensitivity, specificity, accuracy, PPV and NPV calculations for this study population are shown in Table 28. For the PPV and NPV calculations, the same prevalence was used for this study as for the actual DRUID study. Including only those cases in which the person tested was the driver of a motorised vehicle and excluding cases where information was missing or incoherent the number of valid cases was 1807.

Table 28. Results from whole blood for the independent study for DrugWipe 5.

	AMP	CAN	COC	OPI
TP	1460	84	12	4
FP	149	209	26	18
TN	149	1321	1762	1749
FN	59	112	7	36
Total (valid)	1807	1807	1807	1807
Sensitivity	97%	43%	63%	10%
Specificity	50%	87%	99%	99%
Accuracy	89%	82%	98%	97%
Prevalence	84%	11%	1.1%	2.2%
PPV	71%	46%	22%	8.3%
NPV	92%	86%	100%	99%

There are some differences when comparing the results of this study to the results calculated based for OF samples. For amphetamine, specificity is lower when interpreting the test results according to WB analysis results. Also for cannabis, sensitivity is markedly lower when WB results are used for evaluations. Both cocaine and opiates had positive results that made calculations for sensitivity possible. Unfortunately, the sensitivity for cocaine is not very good and is definitely not on an acceptable level for opiates. However, specificity for both drug classes was very good. PPV and NPV values illustrate the same trends as sensitivity and specificity; cannabis, cocaine and opiate tests definitely need improvements. As previously noted, device results with only negative screenings for all drugs are not included here, this will to some extent affect the numbers of TNs for all the individual drug tests and hence the specificity and accuracy can be expected to decrease. The specificity and accuracy results for cocaine and opiates presented here are very high because the prevalences of these substances in Finnish DUI cases are very low.

6.4. Discussion

Regarding the results for sensitivity, specificity and accuracy both tests seem to perform quite well for the ATS drugs, although the sensitivity value for Rapid STAT is only moderate. Nonetheless it is

evident that both screening devices can give negative results for cases that contain concentrations of hundreds or even thousands of ng/ml of ATS drugs. In addition, the DrugWipe 5^+ device also produced a positive test result for some cases that contained no ATSs at all, which is not acceptable.

The performance of the benzodiazepine test of Rapid STAT seems to be at a relative good level already if sensitivity, specificity, accuracy, PPV and NPV are criteria for test performance. However it should be noted that out of all the positive test results only half were TP when test cut-offs were used as a reference point. Unfortunately, the same phenomenon can be seen for Rapid STAT cannabis and methamphetamine tests as well as for the cannabis test in DrugWipe 5^+. This is definitely not acceptable for an on-site device. There would seem to be a lot of development work still ahead on detection of these drug classes for the Rapid STAT test and for the cannabis test of DrugWipe 5^+.

Unfortunately the number of positive cases for cocaine and opiates was very low which made sensitivity calculations not applicable. However, both devices achieved very good results for specificity for these substance classes.

For the CSI checklist evaluation insufficient data was collected concerning final judgement of drug use to draw any meaningful conclusions based on the sensitivity of the checklist. The cases in which most possible symptoms were detected were almost exclusively DUI cases collected by police, relatively few symptoms, or often none at all, were observed in the non suspect CSI checklists performed by researchers. Clearly when using four or more observed symptoms from the list, or signs which are nearly always associated with impairing drug use, as a basis to ascertain suspicion many substance positive cases are missed. On the other hand it is difficult to assign a case as suspicious using very few signs of impairment, which may also be caused by other chronic or acute factors. The results also indicate that many of the amphetamines-positive cases would be missed on the basis of checking for the substance specific signs of impairment; this is also true to a lesser extent for cannabis cases. The prevalence of amphetamine and other ATSs among DUI drivers is relatively high and therefore the Finnish traffic police regard this as a category of drugs of particular concern. Similarly there is a relatively high prevalence of benzodiazepines in Finnish traffic. The number of cases showing 4 or more and/or highly suspicious signs of impairment that contained benzodiazepines, either alone or in combination with other drugs was relatively high. This would appear to indicate that drivers using these sedatives are easier to spot.

6.5. Acknowledgements

We would like to thank police officers Eija-Maija Viirret and Jussi Pohjonen for their help with sample collection and study design planning. We also thank Teemu Gunnar and Kari Ariniemi for the analysis method development. The personnel of the Alcohol and Drug Analytics Unit are thanked for their co-operation. The manufacturers of the device are thanked for providing the team the devices for free and for their co-operation at all stages of the study. We would also like to thank Antti Impinen for his valuable advice on statistical analysis of the results. Finally we acknowledge the cooperation and assistance of Heikki Ihalainen, Lasse Lehtonen and Philippe Lunetta in this task and in DRUID in general.

7. Country report - The Netherlands

Sjoerd Houwing[1], Beitske Smink[2] and René Mathijssen[1]
[1]SWOV Institute for Road Safety Research, the Netherlands
[2]Netherlands Forensic Institute (NFI), the Netherlands

7.1. Introduction

The aim of this study was to analytically evaluate both a number of promising oral fluid (OF) screening devices and a method of identifying recent drug use, based on observable signs and symptoms and self-reported use. In the Netherlands, the use of roadside drug screening devices does not have a legal basis, yet. If there are clear indications of drug-related impairment, a police officer can demand a blood test. However, no standard test battery for drug recognition is in regular use either.

Sensitivity, specificity and accuracy of the selected on-site screening devices were assessed by comparing the on-site screening results with the results of OF and/or blood confirmation analysis in the laboratory of the Netherlands Forensic Institute (NFI).

7.2. Materials and methods

7.2.1. Selection of trial population and sample collection

Subjects were initially selected at random from moving traffic during random breath testing activities by the police, at various days of the week and times of the day. A random sample of stopped drivers were asked by a researcher to participate, on a voluntary basis, in the EU research project DRUID. If they agreed to participate in the study, a trained research nurse would ask them to answer some questions on recent drug use and to deliver a blood sample. If they were not willing or able to provide a blood sample, an oral fluid sample was requested. In case of suspected recent drug use, both sample types were asked for, and additionally an on-site OF screening would be performed if the subject agreed. For each suspected driver that was tested, a non-suspected driver would be tested in the same way.

Venous blood samples were collected in glass tubes containing 20 mg sodium fluoride and 143 IU heparin sodium. OF samples were taken by spitting into a polypropylene container.

During the roadside survey sessions, blood and OF samples were stored in solid carbon dioxide at about -80˚C (dry ice). After transportation to the NFI, blood and OF samples were stored at -20˚C until analysis.

After two years of testing at the roadside, the number of tests with suspected drivers was so low that additional testing in a 'coffeeshop' was required to get a sufficient number of tests for a reliable evaluation of the on-site screening devices. In the Netherlands, a coffeeshop is a place where the purchase and consumption of small quantities of cannabis and jmarijuana is allowed. At the coffeeshop, only oral fluid was collected for confirmation analysis; no blood was taken. Therefore, when comparing the coffeeshop results with the roadside results, only the results of OF confirmation analysis could be used.

7.2.2. Evaluated on-site oral fluid screening devices

The trial lasted 2.5 years, during which period four different devices were evaluated: Rapid STAT from MAVAND Solutions GmbH, DrugTest 5000 from Dräger (cut-off 25ng/ml for Δ^9-THC), Oratect III from Branan, and BIOSENS Dynamic from Biosensor.

All four devices were tested in a coffeeshop. Three devices, Rapid STAT, DrugTest 5000, and Oratect III, were tested at the roadside as well, although the test of the Oratect III was prematurely aborted due to a very high number of failed tests. BIOSENS was added to the trial after the roadside survey had finished. Therefore, this device was only tested in a coffeeshop.

Mavand Rapid STAT was evaluated both at the roadside and in the coffeeshop. It was the first device to be tested in the project, starting early 2007. After the roadside testing, the collection method of the Rapid STAT was improved after feedback from police officers who conducted a practical evaluation as part of DRUID Work Package 3.1. A little later, a second test for Rapid STAT was made, this time in a coffeeshop setting. At this point the "old" version of the device was no longer available, and therefore the improved version was tested instead.

DrugTest 5000 was evaluated during the final phase of the roadside survey. An additional coffeeshop test was conducted in order to be able to compare the results with the Rapid STAT results. The number of tested devices in the coffeeshop was quite low, since at the time of the coffeeshop testing the manufacturer was not able to provide sufficient testing materials.

Oratect III was initially planned to be evaluated at the roadside, but the collection of OF turned out to be very time-consuming, causing subjects to stop cooperation during the collection process. Therefore it was decided to abort the roadside testing of the device. The Oratect III device was later tested successfully in the coffeeshop, where subjects were less hasty. SWOV did not test the Oratect XP, which needs only half the amount of OF the Oratect III needs. The reason was that the panel of the XP version lacked amphetamines and benzodiazepines, thus preventing a fair comparison with the other on-site screening devices.

BIOSENS Dynamic was evaluated only in the coffeeshop since it was included in the final stage of the trial. There are no known cut-off levels from the manufacturer for the BIOSENS. Therefore the device was only evaluated against the DRUID cut-offs. During the first coffeeshop session the sensitivity for THC of the device turned out to be very low. The manufacturer claimed that this was caused by a too high concentration of antibodies due to a production error (see Annex 5). It was decided to conduct a second test session in the coffeeshop and see whether the results would improve drastically.

7.2.3. Checklist for Clinical Signs of Impairment

In addition to the on-site analytical drug screening devices, a checklist of clinical signs of impairment (CSI) was evaluated, supplemented with two questions about recent drug use (see Annex 6). The checklist was based on several existing checklists, e.g. one developed for the German police (20) and previously used in the EU research project IMMORTAL (21). The CSI was completed by the police officers who also performed the breath test for alcohol. In most cases, completing the checklist took no more than one to two minutes. The method was meant to allow a quick scan of potential drug-related impairment. The intended use of the CSI is to preselect drivers for the time-consuming and relatively expensive on-site analytical OF screening. The validity of the method was determined by comparing the police officers' final judgement with the results of OF and blood confirmation analysis. The police officers who used the checklist were not extensively trained and had no experience with drug recognition. They were only briefly instructed just before the start of the testing procedure.

The components of the CSI checklist were:

- Observation of the eyes and the general physical appearance of the driver.
- Breath alcohol test.
- Examination of the eyes (nystagmus, pupil reaction to light) in case of suspicion (dilated or restricted pupils, red eyes).
- Question regarding psychoactive substance use in the past 24 hours and, in case of reported substance use, question regarding the nature of the substance(s).

At the end of the checklist, the police officer was asked to fill in his or her conclusion: did the subject use impairing drugs or not, or was the police officer in doubt. Criteria for suspicion of drug use were one or more of the listed signs and symptoms of impairment and/or self-reported recent drug use.

In total, 4822 drivers were screened for observable signs and symptoms of drug use.

7.2.4. Toxicological analysis

This section contains a brief description of the methods used for the toxicological analysis of blood and OF samples taken from subjects in the framework of the DRUID project. A full description can be found in Annex 7.

Method of sample analysis

Both blood and OF samples were analysed by means of Liquid chromatography-mass spectrometry (LCMS). The analytical method was fully validated for all compounds in blood, which included determination of linearity, accuracy, reproducibility, limit of detection (LOD) and limit of quantification (LOQ), and stability. Additional validation was performed for all compounds in OF.

Sample preparation

Protein precipitation was performed after addition of deuterated analogs of the target compounds, by addition of acetone (0.75 and 0.15 ml for blood and OF, respectively) to the sample (0.25 and 0.05 ml for blood and OF, respectively), followed by centrifugation.

Samples of drivers without suspicion of drug use were pooled before analysis. In case the mixture of 5 x 0.05 ml sample tested negative, all 5 samples were reported to be negative. In case the mixture of 5 samples showed a positive result, the samples were reanalysed one at a time. Samples of drivers with a suspicion of drug use were analysed one by one.

Quality control

The following quality control measures were taken:

- Internal standards (deuterated analogues of most target compounds) were used to correct for analytical variations.
- Calibration of all compounds in blood and OF was performed every 2 months.
- Blank blood samples and control samples (spiked blood samples, prepared at the NFI as well as externally) were analysed daily.
- Shewhart cards were used for a selected number of compounds to monitor the daily performance.
- Regular participation took place in Round Robin tests organized by RTI International (NC, USA) for all the DRUID project toxicology laboratories.

7.3. Results based on confirmation analysis of oral fluid

For confirmation analysis of OF, both DRUID cut-offs and cut-offs of the screening devices themselves have been used. Table 29 gives an overview of the DRUID cut-off levels for OF.

Table 29. DRUID cut-off levels for oral fluid.

Substance	DRUID cut-off levels in oral fluid
Amphetamine	25 ng/ml
Metamphetamine	25 ng/ml
Opiates	20 ng/ml
Cannabis (THC)	1 ng/ml
Cocaine	10 ng/ml
Benzodiazepines	1-5 ng/ml

Cross-reactivity is not reported since in those cases where cross-reactivity occurred, concentrations were so high that the results were positive anyway. Positive predictive value (PPV) and negative predictive value (NPV) were only calculated for THC, and with some devices also for cocaine. For other substances, the numbers of positive samples was too low (N<6) to get statistically significant results. Box and whisker plots are presented for THC only.

7.3.1. Evaluation of Rapid STAT

Rapid STAT 'first version' at the roadside

Table 30. Analytical evaluation of Rapid STAT at the roadside.

	DRUID cut-offs						Device cut-offs					
	AMP	**MAMP**	**CAN**	**OPI**	**COC**	**BZO**	**AMP**	**MAMP**	**CAN**	**OPI**	**COC**	**BZO**
TP	0	0	3	0	3	0	0	0	3	0	3	0
FP	2	2	0	0	0	0	2	2	0	0	0	0
TN	30	33	24	35	29	34	30	33	25	35	31	35
FN	3	0	8	0	3	1	3	0	7	0	0	0
No of tests	35	35	35	35	35	35	35	35	35	35	35	35
Failed tests	0	0	0	0	0	0	0	0	0	0	0	0
Sensitivity	n.a.	n.a.	27%	n.a.	50%	n.a.	n.a.	n.a.	30%	n.a.	n.a.	n.a.
Specificity	94%	94%	100%	100%	100%	100%	94%	94%	100%	100%	100%	100%
Accuracy	86%	94%	77%	100%	91%	97%	86%	94%	80%	100%	97%	100%
Prevalence	8.6%	0%	31%	0%	17%	2.9%	8.6%	0%	29%	0%	11%	0%
PPV	n.a.	n.a.	100%	n.a.	100%	n.a.	n.a.	n.a.	100%	n.a.	n.a.	n.a.
NPV	n.a.	n.a.	33%	n.a.	90%	n.a.	n.a.	n.a.	34%	n.a.	n.a.	n.a.

n.a. = calculation not applicable

Subjects with suspected drug use during the past 12 hours were asked by a member of the research team to volunteer for a test with the Rapid STAT. For each included driver with suspected drug use, a second driver was included who was not suspected of recent drug use. In total 50 tests were carried out with Rapid STAT. 15 cases were excluded from the evaluation because an insufficient amount of OF was collected for confirmation analysis, or because a four-panel test was used instead of the intended 6-panel test. The results of the 35 remaining tests are presented in Table 30.

Cannabis

Eleven OF samples were positive for Δ^9-THC. The concentration range was 3-2770 ng/ml. Three Rapid STAT results were true positive (TP) results. The concentration range for these three cases was 95-2770 ng/ml. Eight results were false negative (FN) results. The concentration range for these eight cases was 3-1144 ng/ml. One of the FN results became a true negative (TN) result when applying the device cut-off instead of the DRUID cut-off. The sensitivity of Rapid STAT for cannabis was very low (27%), even when applying the relatively high device cut-off (30%) of 15 ng/ml.

No false positive (FP) screening results have been observed among the 24 negative cases. As a consequence of the high device cut-off, specificity was 100%.

Cocaine

Six OF samples were positive for cocaine. The concentration range was 10-1464 ng/ml. Three Rapid STAT results were TPs. The concentration range for these three cases was 77-1464 ng/ml. Therefore, sensitivity was 50%. The relatively low sensitivity of Rapid STAT for cocaine (50%) may partly be explained by the fact that for two FN cases the concentration was just above the DRUID cut-off of 10 ng/ml (with 10 ng/ml and 11 ng/ml, respectively). A third FN case had a concentration of 37 ng/ml.

In six out of 29 negative samples, traces of cocaine were found, but below the DRUID cut-offs. No FP results have been observed for cocaine; as a consequence, specificity was 100%.

Amphetamines

Three OF samples were positive for amphetamines. For all three positive cases the screening result was FN. The concentrations of the three samples were: 2128 ng/ml for amphetamines; 37 ng/ml for MDMA in combination with 19 ng/ml for MDA; and 49 ng/ml for MDMA. Due to the low number of positive cases, sensitivity was not calculated.

Additionally, two FP screening results were found among the 32 negative cases. One of these cases was negative for all substances, while the other one was positive for Δ^9-THC (81 ng/ml). In two TN

cases MDMA was detected at a concentration below the DRUID cut-off, none of theother TN cases dhad any traces of amphetamines. Specificity was 94%.

Methamphetamines

None of the OF samples were positive for methamphetamines; therefore sensitivity could not be assessed. Two screening results were FPs; consequently, specificity was 94%.

Opiates

None of the OF samples were positive for opiates, so sensitivity could not be assessed. All 35 screening results were TNs, so specificity was 100%.

Benzodiazepines

Only one out of the 35 OF samples was positive for benzodiazepines. The screening result for this positive case was FN. The concentration found in the positive sample was 3.7 ng/ml nitrazepam. This value is very low and only just above the DRUID cut-off. When applying the device cut-off, the FN screening result became a TN. Sensitivity was not assessed since the number of positive cases was too low.

The remaining 34 screening results were TNs, so specificity was 100%.

Rapid STAT 'next version' in the coffeeshop

Table 31. Analytical evaluation of the Rapid STAT in the coffeeshop.

	DRUID cut-offs						Device cut-offs					
	AMP	**MAMP**	**CAN**	**OPI**	**COC**	**BZO**	**AMP**	**MAMP**	**CAN**	**OPI**	**COC**	**BZO**
TP	3	0	35	3	2	0	3	0	35	2	2	0
FP	2	0	2	0	0	0	2	0	2	1	0	0
TN	35	44	2	40	38	44	37	44	5	40	38	44
FN	4	0	5	1	4	0	3	0	2	1	4	0
No of tests	44	44	44	44	44	44	44	44	44	44	44	44
Failed tests	0	0	0	0	0	0	0	0	0	0	0	0
Sensitivity	43%	n.a.	88%	n.a.	33%	n.a.	43%	n.a.	95%	n.a.	33%	n.a.
Specificity	95%	100%	50%	100%	100%	100%	95%	100%	71%	98%	100%	100%
Accuracy	86%	100%	84%	98%	91%	100%	86%	100%	91%	95%	91%	100%
Prevalence	16%	0%	91%	9.1%	14%	0%	16%	0%	84%	6.8%	14%	0%
PPV	83%	100%	66%	n.a.	100%	n.a.	90%	100%	71%	n.a.	100%	n.a.
NPV	59%	n.a.	87%	n.a.	95%	n.a.	83%	n.a.	88%	n.a.	95%	n.a.

n.a. = calculation not applicable

The second test series with Rapid STAT was conducted in a coffeeshop. Subjects were asked to participate by a researcher. The test with Rapid STAT was carried out by a representative of the manufacturer, under supervision of the main researcher. From each subject, OF for confirmation analysis was collected by a member of the research team by using spit cups. Evaluation results are shown in Table 31.

Cannabis

Altogether 40 OF samples (91%) were positive for Δ^9-THC. The concentration range was 4.5-9092.6 ng/ml. 35 screening results were TPs, relating to a concentration range of 4.5-9093 mg/ml; five screening results were FNs, relating to a concentration range of 7.9-99 ng/ml. Sensitivity was high when applying DRUID cut-offs (88%), and even very high (95%) when applying the device cut-off.

Two screening results were FPs, resulting in 50% specificity. When applying the device cut-off, specificity increased to 71%.

Cocaine

Six OF samples (14%) were positive for cocaine. The concentration range was 17-1847 ng/ml. Two screening results were TPs, relating to concentrations of 1566 and 1847 ng/ml, respectively. The four FN results related to a concentration range of 17-1154 ng/ml. Sensitivity was very low: 33%.

As a consequence of the absence of FP screening results, specificity was 100%.

Amphetamines

Seven OF samples (16%) were positive for amphetamines, with a concentration range of 79-797 ng/ml. Three screening results were TPs, relating to a concentration range of 81-797 ng/ml. The four FN screening results related to a concentration range of 79- 509 ng/ml. Sensitivity for amphetamines was 43%.

28 OF samples were negative for amphetamines. Three of these contained traces of amphetamines but below the DRUID cut-offs. Two screening results were FPs, resulting in 95% specificity. Both FP results related to OF samples containing amphetamines below the DRUID cut-off levels.

Methamphetamines

None of the OF samples were positive for methamphetamines and all 44 screening results were TNs. Therefore, sensitivity could not be calculated and specificity was 100%. In one sample, methamphetamines were found in the laboratory, but at a concentration below the DRUID cut-off levels.

Opiates

Four OF samples were positive for 6-MAM, with concentrations ranging from 20-8259 ng/ml. One screening result was FN. The corresponding OF sample had a 6-MAM concentration of 4563 ng/ml and a codeine concentration of 346 ng/ml. Remarkably, screening results relating to this sample were also FN for cannabis (28 ng/ml) and cocaine (1154 ng/ml). The three TP screening results related to concentrations ranging from 20-8,259 ng/ml. Sensitivity was not calculated because of the too low number of opiate-positive samples.

The remaining 40 screening results were all FNs when applying the DRUID cut-offs, resulting in 100% specificity. When applying the device cut-off, one screening result was FP, resulting in 98% specificity.

Benzodiazepines

No OF samples were positive for benzodiazepines. Consequently, sensitivity could not be calculated. All 44 screening results were TNs, resulting in 100% specificity.

7.3.2. Evaluation of DrugTest 5000

DrugTest 5000 at the roadside

Table 32. Analytical evaluation of the DrugTest 5000 at the roadside.

	DRUID cut-offs						Device cut-offs					
	AMP	**MAMP**	**CAN**	**OPI**	**COC**	**BZO**	**AMP**	**MAMP**	**CAN***	**OPI**	**COC**	**BZO**
TP	0	0	10	1	3	2	0	0	10	1	3	2
FP	2	0	5	0	0	0	2	0	5	0	0	0
TN	62	63	41	63	58	62	62	63	44	63	60	62
FN	0	0	8	0	3	0	0	0	5	0	1	0
No of tests	64	63	64	64	64	64	64	63	64	64	64	64
Failed tests	0	1	0	0	0	0	0	1	0	0	0	0
Sensitivity	n.a.	n.a.	56 %	n.a.	50%	n.a.	n.a.	n.a.	67%	n.a.	n.a.	n.a.
Specificity	97%	100%	89%	100%	100%	100%	97%	100%	90%	100%	100%	100%
Accuracy	95%	100%	80%	100%	95%	100%	95%	100%	84%	100%	98%	100%
Prevalence	0%	0%	28%	1.6%	9.4%	3.1%	0%	0%	23%	1.6%	6.3%	3.1%
PPV	n.a.	n.a.	93%	n.a.	100	n.a.	n.a.	n.a.	95%	n.a.	n.a.	n.a.
NPV	n.a.	n.a.	42%	n.a.	90	n.a.	n.a.	n.a.	49%	n.a.	n.a.	n.a.

n.a. = calculation not applicable, *(cut-off 25ng/ml for Δ^9-THC)

Table 32 gives an overview of the results of the analytical evaluation of the DrugTest 5000 at the roadside, which was carried out from early 2008 to mid-2009.

Cannabis

Eighteen OF samples (28%) were positive for Δ^9-THC, concentrations ranging from 2.4-8035 ng/ml. Ten screening results were TPs, relating to concentrations ranging from 44-8035 ng/ml. Eight results were FNs relating to concentrations ranging from 2.4-407 ng/ml. The resulting sensitivity was 56%. When applying the device cut-off, three FN results became TNs, resulting in 67% sensitivity.

Five FP results were observed among the 46 negative cases. As a consequence specificity was 89%.

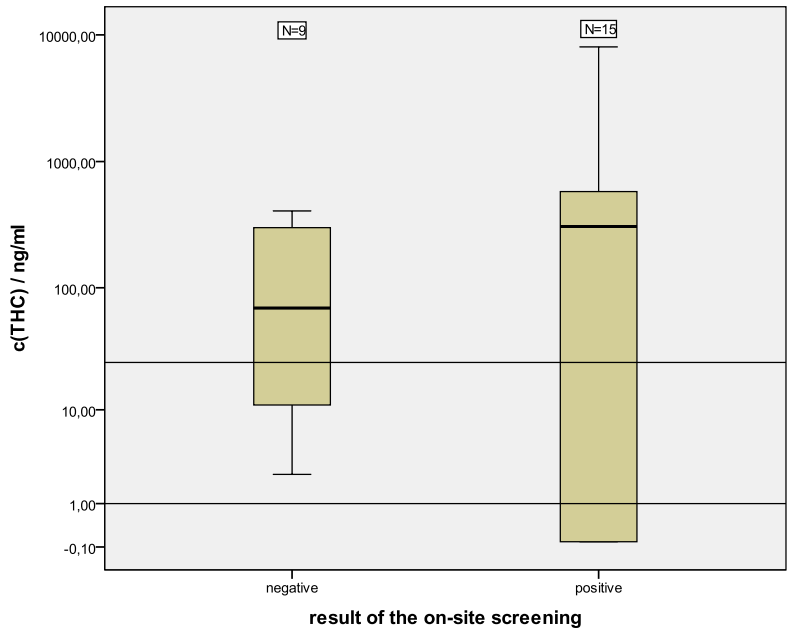

Figure 34. Box and whisker plot cannabis DrugTest 5000 at the roadside. 40 cases with 0 ng/ml of Δ^9-THC in their OF were tested negative for Δ^9-THC. These cases are not included in the plot. Horizontal line at 25 ng/ml indicates test cut-off.

The box and whisker plot for the DrugTest 5000 shows that there is a large overlap in the Δ^9-THC concentrations of positive and negative screening results. Therefore, the seperating power of the screening device can be regarded as low.

Cocaine

Six OF samples (9%) were positive for cocaine. Five of them had cocaine concentrations ranging from 253-6381 ng/ml. The sixth positive sample contained only benzoylecgonine, with a concentration of 95 ng/ml. Three screening results were TPs. Two of these related to cocaine concentrations of 253 and 2825 ng/ml, respectively; the third one related to the sample that only contained benzoylecgonine. The three FN results related to cocaine concentrations ranging from 389-6381 ng/ml. As a result, sensitivity was 50%.

The remaining 58 screening results were TN, resulting in 100% specificity.

Amphetamines

All 64 OF samples were negative for amphetamines. Therefore, sensitivity could not be assessed. Two screening results were FPs. As a consequence, specificity was 97%.

Methamphetamines

All OF samples were negative for methamphetamines. Sensitivity could therefore not be assessed. One of the screening tests failed to give a result for methamphetamines, probably due to a low concentration of saliva. The remaining 63 screening results were TNs, resulting in 100% specificity.

Opiates

Only one OF sample was positive for opiates, containing 1909 ng/ml of 6-MAM and 319 ng/ml of codeine. The corresponding screening result was TP. Sensitivity could not be assessed. The remaining 63 screening results were TNs, resulting in 100% specificity.

Benzodiazepines

Two OF samples were positive for benzodiazepines, containing 68 and 202 ng/ml diazepam, respectively. Both corresponding screening results were TPs. Sensitivity was not assessed.

Among the 62 negative samples, two contained traces of benzodiazepines, but below the DRUID cut-offs. All 62 corresponding screening results were TNs and as a consequence specificity was 100%.

DrugTest 5000 in the coffeeshop

Table 33. Analytical evaluation of DRUGTEST 5000 in the coffeeshop.

	DRUID cut-offs						Device cut-offs					
	AMP	**MAMP**	**CAN**	**OPI**	**COC**	**BZO**	**AMP**	**MAMP**	**CAN***	**OPI**	**COC**	**BZO**
TP	0	0	13	0	0	0	0	0	13	0	0	0
FP	0	0	0	0	0	0	0	0	0	0	0	0
TN	19	20	3	20	20	20	20	20	5	20	20	20
FN	1	0	4	0	0	0	0	0	2	0	0	0
No of tests	20	20	20	20	20	20	20	20	20	20	20	20
Failed tests	0	0	0	0	0	0	0	0	0	0	0	0
Sensitivity	n.a.	n.a.	76%	n.a.	n.a.	n.a.	n.a.	n.a.	87%	n.a.	n.a.	n.a.
Specificity	100%	100%	n.a.	100%	100%	100%	100%	100%	n.a.	100%	100%	100%
Accuracy	95%	100%	80%	100%	100%	100%	100%	100%	90%	100%	100%	100%
Prevalence	5.0%	0%	85%	0%	0%	0%	0%	0%	75%	0%	0%	0%
PPV	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
NPV	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.

n.a. = calculation not applicable, *(cut-off 25ng/ml for Δ^9-THC)

DrugTest 5000 was also tested in the coffeeshop, although not many tests could be carried out due to a low number of available test cassettes. Prevalence was low for all substances except for cannabis (see Table 33), making a proper sensitivity assessment only possible for this substance. All tests with

the DrugTest 5000 were carried out by a representative of the manufacturer, under supervision of the main SWOV-researcher.

Cannabis

Seventeen OF samples (85%) were positive for Δ^9-THC. The concentration ranged from 1.3-12063 ng/ml. 13 screening results were TPs, relating to a concentration range of 25-12063 ng/ml. Four other results were FNs, relating to a concentration range of 1.3-119 ng/ml. Sensitivity was 76%, but increased to 87% by applying the device cut-off.

There were only three negative samples. All three corresponding screening results were TNs. Specificity was not assessed due to the low number of negative samples.

Cocaine

None of the twenty OF samples were positive for cocaine. Therefore, sensitivity could not be assessed. All screening results were TNs, resulting in 100% specificity.

Amphetamines

Only one OF sample (5%) was positive for amphetamines, containing a concentration of 66 ng/ml of MDA. The corresponding screening result was FN. Sensitivity was not assessed. All other screening results were TNs, resulting in 100% specificity.

Methamphetamines

None of the twenty samples were positive for methamphetamines. Therefore, sensitivity could not be assessed. All screening results were TNs, resulting in 100% specificity.

Opiates

None of the twenty samples were positive for opiates. Therefore, sensitivity could not be assessed. All screening results were TNs, resulting in 100% specificity.

Benzodiazepines

None of the twenty samples were positive for benzodiazepines. Therefore, sensitivity could not be assessed. All screening results were TNs, resulting in 100% specificity.

7.3.3. Evaluation of Oratect III

Table 34. Analytical evaluation of the Oratect III in the coffeeshop.

	DRUID cut-offs						Device cut-offs					
	AMP	**MAMP**	**CAN**	**OPI**	**COC**	**BZO**	**AMP**	**MAMP**	**CAN**	**OPI**	**COC**	**BZO**
TP	0	0	16	0	0	0	0	0	16	0	0	0
FP	0	1	0	0	0	0	0	1	0	0	0	0
TN	56	57	8	57	55	58	56	57	22	57	55	58
FN	2	0	34	1	3	0	2	0	20	1	3	0
No of tests	58	58	58	58	58	58	58	58	58	58	58	58
Failed tests	0	0	0	0	0	0	0	0	0	0	0	0
Sensitivity	n.a.	n.a.	32%	n.a.	n.a.	n.a.	n.a.	n.a.	44 %	n.a.	n.a.	n.a.
Specificity	100%	98%	100%	100%	100%	100%	100%	98%	100%	100%	100%	100%
Accuracy	97%	98%	41%	98%	95%	100%	97%	98%	65%	98%	95%	100%
Prevalence	3.4%	0%	86%	1.7%	5.2%	0%	3.4%	0%	62%	1.7%	5.2%	0%
PPV	n.a.	n.a.	100%	n.a.	n.a.	n.a.	n.a.	n.a.	100%	n.a.	n.a.	n.a.
NPV	n.a.	n.a.	35%	n.a.	n.a.	n.a.	n.a.	n.a.	39%	n.a.	n.a.	n.a.

n.a. = calculation not applicable

Cannabis

50 OF samples (86%) were positive for Δ^9-THC, concentrations ranging from 1.2-13,934 ng/ml. Sixteen screening results were TPs, relating to a concentration range of 46-13934 ng/ml. Another 34

results were FNs, relating to a concentration range of 1.2-2941 ng/ml. Sensitivity was 32%. By applying the device cut-off, sensitivity increased to 44%

None of the screening results were FPs, so specificity was 100%.

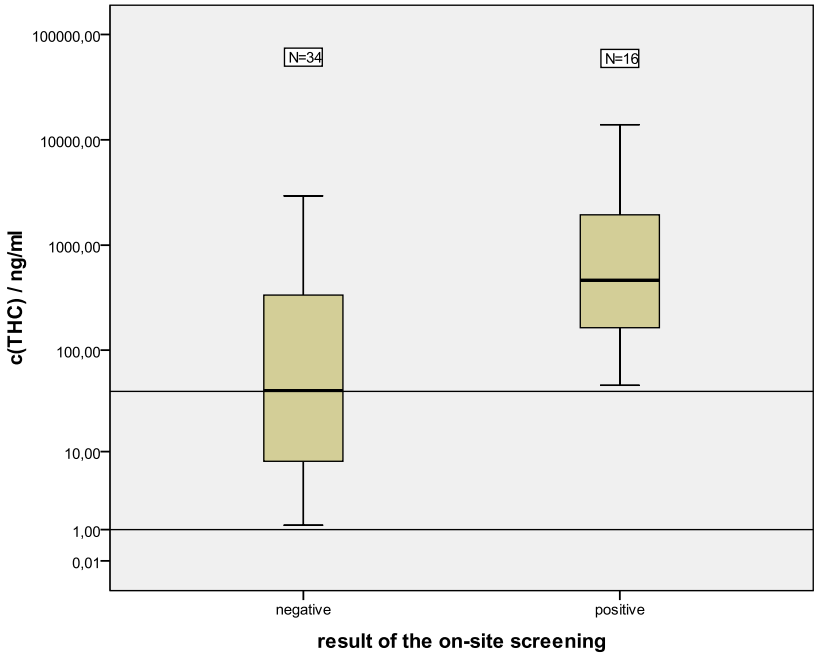

Figure 35. Box and whisker plot cannabis Oratect III in the 'coffeeshop'. 8 cases with 0 ng/ml of Δ^9-THC in their OF were tested negative for THC. These cases are not included in the plot. Horizontal line at 40 ng/ml indicates test cut-off.

The box and whisker plot shows no clear distinction between the Δ^9-THC concentrations of the positive and negative screening results.

Cocaine

Three OF samples (5%) were positive for cocaine, concentrations ranging from 22-49 ng/ml. All three corresponding screening results were FNs. Sensitivity was not assessed due to the low number of positive cases. The remaining 55 screening results were TNs, resulting in 100% specificity.

Amphetamines

Two OF samples were positive for amphetamines containing concentrations of 43 and 73 ng/ml respectively. Both corresponding screening results were FNs. Sensitivity was not assessed since only two cases were positive. The remaining 56 screening results were TNs, resulting in 100% specificity.

Methamphetamines

None of the OF samples were positive for methamphetamines. As a consequence, sensitivity could not be assessed. One screening result was FP, the remaining 57 being TNs. The FP result related to a sample with a Δ^9-THC concentration of 454 ng/ml, with no other substances being present. Consequently, specificity was 98%.

Opiates

One OF sample was positive for opiates, containing a concentration of 27 ng/ml morphine, just above the DRUID cut-off level of 20 ng/ml. The corresponding screening result was FP. Sensitivity was not assessed. All 57 remaining screening results were TNs, resulting in 100% specificity.

Benzodiazepines

None of the oral samples were positive for benzodiazepines. Therefore, sensitivity could not be assessed. All 58 screening results were TNs, resulting in 100% specificity.

7.3.4. Evaluation of BIOSENS

The BIOSENS Dynamic was evaluated only in the coffeeshop since it was included only in the final stage of the trial. The manufacturer of BIOSENS could not provide cut-off levels. Therefore the device was only evaluated against the DRUID cut-offs. During the first session in the coffeshop, sensitivity of the device for cannabis turned out to be very low. The manufacturer claimed that this was caused by a too high concentration of antibodies due to a production error (see Annex 5). It was decided to conduct a second test session and see whether sensitivity would increase significantly. Evaluation results are shown in Table 35.

Table 35. Analytical evaluation of BIOSENS in the coffeeshop.

	DRUID cut-offs					
	AMP/ MAMP	**CAN1***	**CAN2****	**OPI**	**COC**	**BZO**
TP	0	8	18	2	0	0
FP	9	0	1	3	0	0
TN	99	8	2	111	101	118
FN	10	63	18	2	3	0
No of tests	118	79	39	118	104	118
Failed tests	0	0	0	0	14	0
Sensitivity	0%	11%	50%	n.a.	n.a.	n.a.
Specificity	92%	100%	n.a.	97%	100%	100%
Accuracy	84%	20%	45%	96%	100%	100%
Prevalence	8.5%	90%	92%	3.4%	2.9%	0%
PPV	n.a.	100%	n.a.	n.a.	n.a.	n.a.
NPV	n.a.	29%	n.a.	n.a.	n.a.	n.a.

* first session; **second session
n.a. = calculation not applicable

Cannabis

During the first coffeeshop session, 71 OF samples (90%) were positive for Δ^9-THC, concentrations ranging from 1.3-12063 ng/ml. Eight corresponding screening results were TPs, relating to a concentration range of 1374-12063 ng/ml. The other 63 corresponding screening results were FNs, relating to a concentration range of 1.3-11119 ng/ml. Sensitivity was only 11%, which was remarkably low.

The remaining eight screening results were TNs, resulting in 100% specificity.

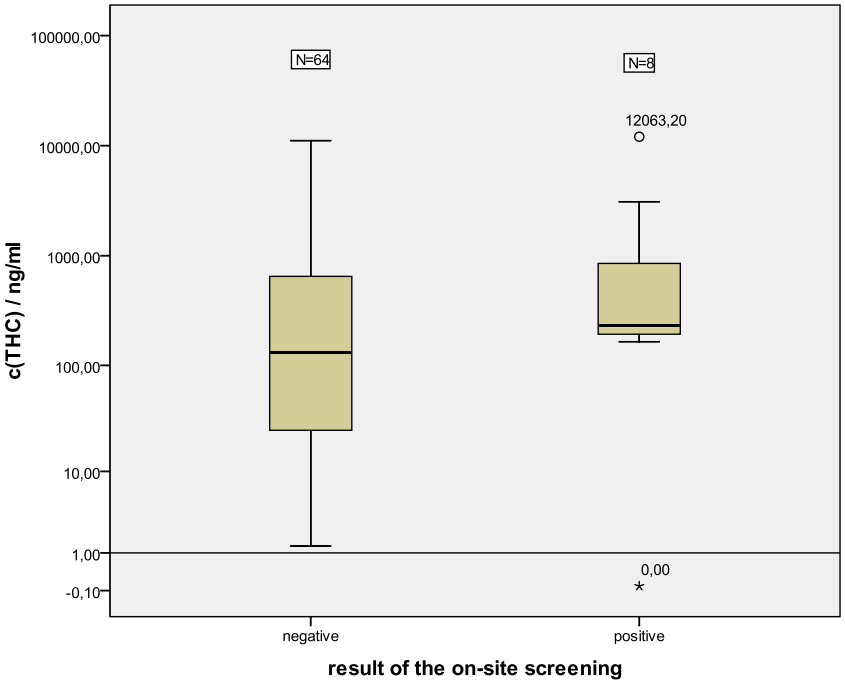

Figure 36. Box and whisker plot cannabis first evaluation BIOSENS in the coffeeshop. 7 cases with 0 ng/ml of Δ^9-THC in their OF were tested negative for THC. These cases are not included in the plot. No test cut-off is available, horizontal line at 1ng/ml indicates DRUID cut-off.

The box and whisker plot shows no clear distinction between the Δ^9-THC concentrations of the positive and negative screening results.

During the second coffeeshop session, 36 OF samples (92%) were positive for Δ^9-THC, concentrations ranging from 4.5-9093 ng/ml. Half of the corresponding screening results were TPs, relating to a concentration range of 226-9093 ng/ml. The other half were FNs, relating to a concentration range of 4.5-1721 ng/ml. The resulting 50% sensitivity was much higher than during the first session (11%), but still relatively low.

Only three OF samples were negative. One corresponding screening result was FP, the other two being TNs. Specificity was not assessed due to the low number of negative cases.

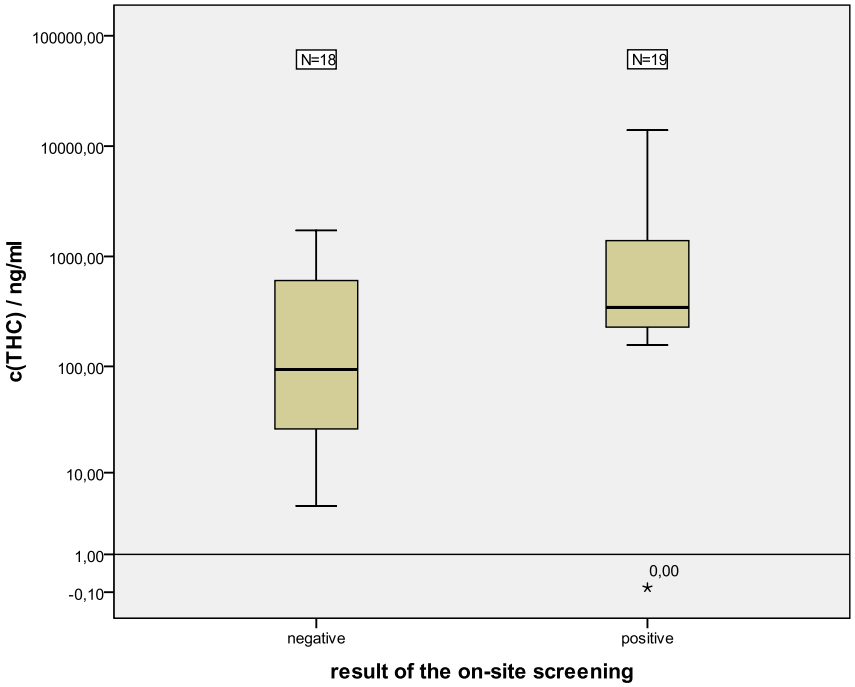

Figure 37. Box and whisker plot cannabis second evaluation BIOSENS in the coffeeshop. 2 cases with 0 ng/ml of Δ^9-THC in their OF were tested negative for THC. These cases are not included in the plot. No test cut-off is available, horizontal line at 1 ng/ml indicates DRUID cut-off.

The box and whisker plot improved somewhat when compared with the first session but still shows no clear distinction between the THC concentrations of the positive and negative screening results.

Cocaine

During the first coffeeshop session, none of the 79 OF samples were positive for cocaine. All corresponding screening results were TNs.

During the second coffeeshop test, a component for the screening of cannabis broke down. This component was replaced by the component that originally belonged to the screening of cocaine. Therefore, after 25 tests, testing for cocaine had to be aborted.

In total, three samples (2.9%) were positive for cocaine, concentrations ranging from 17 to 1847 ng/ml. All three corresponding screening results were FNs. Sensitivity was not assessed due to the low number of positive cases.

For the 101 negative OF samples the corresponding screening results were TNs. Consequently, specificity was 100%.

Three negative OF samples contained small concentrations of cocaine below the cut-off.

Methamphetamines/Amphetamines

Ten OF samples (8.5%) were positive for methamphetamines or amphetamines. Nine samples were positive for amphetamines with a concentration range of 79-1081 ng/ml and one was positive for MDA with a concentration of 66 ng/ml. All corresponding screening results were FNs. Consequently sensitivity was zero.

Regarding the 108 negative OF samples, 99 corresponding screening results were TNs, while nine others were FPs. Six of these FP screening results related to OF samples containing small concentrations of (meth)amphetamines below the DRUID cut-offs. The two remaining FP screening results related to OF samples containing THC, cocaine and 6-MAM, but not (meth)amphetamines . The resulting specificity was 92%.

Opiates

Four OF samples (3.4%) were positive for 6-MAM with a concentration range of 20-8259 ng/ml. Two screening results were FNs, relating to concentrations of 36 and 4563 ng/ml, respectively. The other two corresponding screening results were TPs, relating to concentrations of 20 and 8259 ng/ml, respectively. Sensitivity was not assessed due to the low number of positive cases.

Regarding the 114 negative OF samples, three screening results were FPs. One result related to a sample that was negative for all substances and the other two to samples that were positive for Δ^9-THC only (1374 and 2444 ng/ml, respectively). The resulting specificity was 97%.

Benzodiazepines

None of the OF samples were positive for benzodiazepines. All 118 screening results were TNs, resulting in 100% specificity. Sensitivity could not be assessed.

7.4. Results based on confirmation analysis in blood

For the evaluation based on comparison of on-site screening results with the results of confirmation analysis of blood, only DRUID cut-off levels could be used. These levels are presented in Table 36. All blood samples originate from the roadside. In the coffeeshop only OF was collected.

Table 36.DRUID cut-off levels for blood.

Substance	DRUID cut-off levels in blood
Amphetamine	20 ng/ml
Metamphetamine	20 ng/ml
Opiates	10 ng/ml
Cannabis (THC)	1 ng/ml
Cocaine	10 ng/ml
Benzodiazepines	2-50 ng/ml

7.4.1. Evaluation of Rapid STAT 'first version'

Table 37. Analytical evaluation of Rapid STAT at the roadside .

	DRUID cut-offs in blood					
	AMP	MAMP	CAN	OPI	COC	BZO
TP	0	0	3	0	2	0
FP	2	1	0	0	1	0
TN	26	27	20	28	25	28
FN	0	0	5	0	0	0
No of tests	28	28	28	28	28	28
Failed tests	0	0	0	0	0	0
Sensitivity	n.a.	n.a.	38%	n.a.	n.a.	n.a.
Specificity	93%	96%	100%	100%	96%	100%
Accuracy	93%	96%	82%	100%	96%	100%
Prevalence	0%	0%	98%	0%	7.1%	0%
PPV	n.a.	n.a.	35%	n.a.	n.a.	n.a.
NPV	n.a.	n.a.	29%	n.a.	n.a.	n.a.

n.a. = calculation not applicable

Table 37 gives an overview of the results of the analytical evaluation of Rapid STAT 'first version' at the roadside.

Cannabis

Eight blood samples (29%) were positive for Δ^9-THC. Five of these contained Δ^9-THC with a concentration range of 1.2-61 ng/ml. Three other samples only contained THCC with a concentration

range of 11-67 ng/ml. Three corresponding screening results were TPs. These TP results related to blood samples containing 10 ng/ml Δ^9-THC and 17 ng/ml THCC; 10 ng/ml Δ^9-THC plus 22 ng/ml THCC; and 16 ng/ml THCC only. The remaining five corresponding screening results were FNs. Three of these related to blood samples containing 1.2 - 61 ng/ml Δ^9-THC and the other two to blood samples containing only 11 and 67 ng/ml THCC, respectively. The resulting sensitivity was 38%.

For the 20 negative blood samples the corresponding screening results were TNs. Consequently, specificity was 100%.

Cocaine

Two blood samples were positive for cocaine. One sample contained 60.5 ng/ml cocaine plus 1186 ng/ml benzoylecgonine; the other one contained 37 ng/ml cocaine plus 762 ng/ml benzoylecgonine. Sensitivity was not assessed due to the low number of positive cases.

Regarding the 26 negative blood samples, one corresponding screening result was FP, resulting in 96% specificity.

Amphetamines

None of the blood samples were positive for amphetamines. Sensitivity could therefore not be assessed. For two negative blood samples the corresponding screening result was FP. As a consequence, specificity was 93%.

Methamphetamines

None of the blood samples were positive for methamphetamines. Sensitivity could therefore not be assessed. One screening result was FP, resulting in 96% specificity.

Opiates

None of the blood samples were positive for opiates. Sensitivity could therefore not be assessed. All 28 screening results were TNs, resulting in 100% specificity.

Benzodiazepines

None of the blood samples were positive for benzodiazepines. Sensitivity could therefore not be assessed. All 28 screening results were TNs, resulting in 100% specificity.

7.4.2. Evaluation of DrugTest 5000

Table 38. Analytical evaluation of the DrugTest 5000 at the roadside.

	DRUID cut-offs blood					
	AMP	**MAMP**	**CAN***	**OPI**	**COC**	**BZO**
TP	1	0	8	1	1	1
FP	0	0	1	0	1	1
TN	62	63	40	63	62	61
FN	1	0	15	0	0	1
No of tests	64	63	64	64	64	64
Failed tests	0	1	0	0	0	0
Sensitivity	n.a.	n.a.	35%	n.a.	n.a.	n.a.
Specificity	100%	100%	98%	100%	98%	98%
Accuracy	98%	100%	75%	100%	98%	98%
Prevalence	3.1%	0%	36%	1.6%	1.6%	3.1%
PPV	n.a.	n.a.	98%	n.a.	n.a.	n.a.
NPV	n.a.	n.a.	35%	n.a.	n.a.	n.a.

n.a. = calculation not applicable, *(cut-off 25ng/ml for Δ^9-THC)

Table 38 gives an overview of the results of the analytical evaluation of the DrugTest 5000 at the roadside.

Cannabis

23 blood samples (36%) were positive for Δ^9-THC, with a concentration range of 1.4-45 ng/ml. Eight corresponding screening results were TPs, relating to a concentration range of 3.6-45 ng/ml. 15 corresponding screening results were FNs, twelve of these relating to Δ^9-THC concentrations ranging from 1.4-27 ng/ml, and the remaining three relating to THCC concentrations ranging from 4.6-6.6 ng/ml. The resulting sensitivity was 35%.

Regarding the 41 negative blood samples, only one screening result was FP, resulting in 98% specificity.

Cocaine

One blood sample was positive for cocaine, containing 405 ng/ml cocaine and 3671 ng/ml benzoylecgonine. The corresponding screening result was TP. Sensitivity was not assessed.

Regarding the 63 negative blood samples, one corresponding screening result was FP. The 62 other ones were TNs, resulting in 98% specificity.

Amphetamines

Two blood samples were positive for amphetamines. One corresponding screening result was TP, relating to a concentration of 87 ng/ml amphetamine. The FN screening result related to a concentration of 83 ng/ml MDMA. Sensitivity was not assessed due to the low number of positive blood samples.

Regarding the 62 negative blood samples, all screening results were TNs. As a consequence, specificity was 100%.

Methamphetamines

No blood samples were positive for methamphetamines. Sensitivity could therefore not be assessed. One of the screening tests failed to give a result for methamphetamines, probably due to a too low concentration of OF.

Regarding the 63 negative blood samples, all screening results were TNs. As a consequence, specificity was 100%.

Opiates

One blood sample was positive for opiates, with a concentration of 31 ng/ml of morphine. The positive corresponding screening result was TP. Sensitivity was not assessed.

Regarding the 63 negative blood samples, all corresponding screening results were TNs. As a consequence, specificity was 100%.

Benzodiazepines

Two blood samples were positive for benzodiazepines, one containing 12 ng/ml alprazolam and the other 89 ng/ml of oxazepam. The screening result relating to the alprazolam was TP, but the one relating to the oxazepam was FN. Sensitivity was not assessed, due to the low number of positives.

Regarding the 62 negative blood samples, one corresponding screening result was FP, which resulted in 98% specificity.

7.5. Checklist of clinical signs of impairment

Police officers who were using the CSI had to conclude each checklist with one of the following three options: "the subject did recently use drugs", "the subject did not recently use drugs", or "it is not clear whether the subject did use drugs or not". Police officers were instructed to check the first option in case of self-reported drug use. However, in practice the instructions were often interpreted in a way that if signs of impairment were lacking, the second or third option was chosen. Therefore two tables will be presented: one based on the options chosen by the police officers, and one where all self-reported use is considered as resulting in the first option, regardless what the police officers have filled in.

Table 39. Analytical evaluation of the CSI checklist at the roadside.

	CSI by police officer	CSI by police officer plus self reported use
TP	33	86
FP	26	87
TN	4500	4444
FN	233	185
No of tests	4792	4802
Sensitivity	13%	32%
Specificity	99%	98%
Accuracy	95%	95%
Prevalence	5.6%	5.6%

Table 39 presents the results of the analytical evaluation of the CSI checklist that was assessed at the roadside. In total, 266 blood or OF samples collected at the roadside were positive. Only 33 of the corresponding police decisions were TPs, while the remaining 233 decisions were FNs. The resulting sensitivity of the CSI was only 13%.

Regarding the 4526 negative samples, the outcome of the checklist was 4500 TNs and 26 FPs, which resulted in 99% specificity.

If self reported drug use was considered as a positive outcome, sensitivity increased to 32% and specificity slightly dropped to 98%.

7.6. Discussion

Four different devices have been evaluated in the Dutch trial of on-site OF screening devices: Rapid STAT, DrugTest 5000, Oratect III and BIOSENS Dynamic. Comparison of the results was only possible for cannabis, since in most cases prevalence of the other substances was too low to calculate sensitivity. Table 40 gives an overview of the results of the analytical evaluation for cannabis.

Table 40. Comparison of oral fluid screening devices for cannabis (DR=DRUID cut-off, DE=device cut-off).

	Rapid STAT				DrugTest 5000*				Oratect III		BIOSENS	
	Roadside		Coffeeshop		Roadside		Coffeeshop		Roadside		Coffeeshop	
	DR	DE	DR	DE	DR	DE	DR	DE	DR	DE	DR**	DR***
Sensitivity	27%	30%	88%	95%	56%	67%	76%	87%	32%	44%	11%	50%
Specificity	100%	100%	50%	71%	89%	90%	n.a.	n.a.	100%	100%	100%	n.a.
Accuracy	77%	80%	84%	91%	80%	84%	80%	90%	41%	65%	20%	45%
Prevalence	31%	29%	91%	84%	28%	23%	85%	75%	86%	62%	90%	92%

* (cut-off 25ng/ml for Δ^9-THC), **first session; ***second session
n.a. calculation not applicable

Sensitivity for cannabis varied between 27% and 56% at the roadside when using OF for confirmation and applying the DRUID cut-off. When applying the device cut-off, sensitivity for cannabis at the roadside varied between 30% and 67%.

Sensitivity for cannabis was higher in the coffeeshop than at the roadside. Rapid STAT sensitivity for cannabis increased from 27% at the roadside to 88% in the coffeeshop and DrugTest 5000 sensitivity increased from 56% to 76%. The better results in the coffeeshop can be explained by the higher median Δ^9-THC concentrations in the coffeeshop (1047 ng/ml, versus 715.3 ng/ml at the roadside).

Furthermore, results from the Dutch evaluation indicate that higher sensitivity results in lower specificity, and vice versa. This is an expected outcome since low sensitivity decreases the risk of a false positive result.

Table 41 shows the evaluation results of Rapid STAT and DrugTest 5000 for cannabis based on confirmation analysis of OF and blood, respectively. For the other devices no blood samples were

available. The DrugTest 5000 results indicate that sensitivity is lower for confirmation analysis in blood than for confirmation analysis in OF. This could be expected since Δ^9-THC concentrations are higher in OF than in blood.

Table 41. Comparison of evaluation results using oral fluid (OF) versus blood (B) for confirmation (DRUID cut-offs applied).

	Rapid STAT		**DrugTest 5000**	
	Street		**Street**	
	OF	**B**	**OF**	**B**
Sensitivity	27%	38%	56%	35%
Specificity	100%	100%	89%	98%
Accuracy	77%	82%	80%	75%
Prevalence	31%	29%	28%	36%

Remarkably, Rapid STAT results do no support these findings. Rapid STAT sensitivity is a little higher for confirmation analysis of blood, but the difference is not statistically significant due to the low absolute numbers. The evaluation results of the Dutch CSI checklist were not very encouraging, though. Sensitivity was only 13%, and even when the results of a question on self-reported use were added, sensitivity increased to no more than 32%. This means that two-thirds of all drug-positive drivers were not detected by using this CSI checklist.

An earlier evaluation of the same checklist as part of the EU research project IMMORTAL (21) resulted in sensitivity scores that were twice as high as the scores in the DRUID study (25.3% for clinical signs of impairment alone and 61.1% for the combination of clinical signs and self-reported use). The difference between the results of the evaluation in IMMORTAL and in DRUID could be explained by the fact that in IMMORTAL most CSI checklist tests have been conducted by one and the same police officer, who got more and more experienced throughout the project, whereas the CSI checklist in the DRUID project was used by many different police officers. Results of the CSI checklist may improve by better training of police officers and by selecting times and places with high drug driving incidence.

7.7. Acknowledgements

We would like to express our gratitude to the hundreds of Dutch police officers who made it possible to collect data and body fluid samples at the roadside, and who tested the checklist of signs of impairment. Furthermore, we would like to thank the staff of the coffeeshop who allowed us to test their clients. And, last but not least, we also would like to thank the manufacturers who provided their on-site OF screening devices for free and who cooperated in many other ways

8. Integrated results and discussion

8.1. Integration of results for sensitivity, specificity and accuracy

The study for Task 3.2 of Work package 3 was carried out in three separate countries: Belgium, Finland and the Netherlands. In addition, two of the devices, the Mavand Rapid STAT and the Dräger DrugTest 5000, were tested in two, or more, countries participating in the study. In order to allow better comparability of the Task 3.2 evaluation findings the results for each separate device are presented in this section. 80% was set as a desirable target value for sensitivity, specificity and accuracy. This was done to enable analysis of the devices against a value that would be desirable for DUI testing, the value 80% is not an established criteria to determine 'success or failure' of a device.

All results of the sensitivity, specificity and accuracy according to the DRUID cut-offs were combined into tables. A separate table was made for each substance (Annex 8). For devices tested in two or three countries (Rapid STAT and DrugTest 5000), the results of different countries were combined to obtain a more comprehensive evaluation. Positive predictive value and negative predictive value were omitted from the tables because they are dependent of prevalence. Average values for sensitivity, specificity and accuracy were calculated based on results of all the devices with applicable calculations. The two sets of results for the BIOSENS Dynamic cannabis test were not combined because the manufacturer had claimed there was an error in production that decreased the sensitivity of the system in the earlier results.

The respective sensitivity, specificity and accuracy scores for the devices need to be considered together when interpreting the results, along with the study population used for the testing. The accuracy of a device is the percentage of true negative or true positive results out from all the test results. However, a high accuracy alone does not imply that the devices are necessarily good at correctly identifying positive cases since testing a population with a lot of non-users of a specific substance can be expected to result in a very high proportion of true negative cases, thus resulting in a high accuracy score. Therefore, the sensitivity, or proportion of positive cases that are correctly identified by the test, as well as the number of positive cases included to the study, also have to be examined to allow a better understanding of how the device is performing. Similarly, the specificity, or proportion of negative cases that are correctly identified by the test and the number of negative cases included to the study are important. It is notable that in this study all of the specificities for individual drug tests of the devices are in the range 90–100%, except for the OrAlert opiates test (81%). The study populations in general contained a high number of negative cases; notable exceptions to this are populations with a high proportion of cannabis and opiates positive cases in the Netherlands and Belgium respectively. Despite the need to consider all the results together sensitivity remains perhaps the primary parameter of interest. Combined sensitivity plots, with 95% confidence intervals, for all the relevant devices are shown in the related section of this discussion for each type of test (Figures 38-42). The confidence intervals were calculated with the modified Wald method (22). It should be noted that when there are no true positive or false negative results for a device test sensitivity can not be calculated and no error bar is calculated.

8.1.1. Amphetamine-Type Substance tests

The sensitivities of the amphetamines tests for each device are shown in Figure 38. For the amphetamine tests, the average sensitivity, specificity and accuracy was 60%, 97% and 93% respectively. The DrugWipe 5^{+} showed best performance reaching over 80% values in sensitivity (87%), specificity (95%) and accuracy (93%). DrugTest 5000 and Cozart DDS were above the average device performance for specificity and accuracy (98-99% for both devices), but already had problems with sensitivity (67% for both devices). OraLab6 and Rapid STAT were below average in sensitivity (58% and 54% respectively), but above average in specificity (100% and 97% respectively) and accuracy (94% and 92% respectively). For the evaluation of BIOSENS, the sensitivity was 0%, but the specificity was very good (92%) and accuracy was good (84%). Not enough amphetamine positive cases were gathered to properly evaluate sensitivity for the OrAlert and Oratect III devices (shown as 33% and 0% respectively). The standard error bars of the plots show a large variability which can be attributed to the small numbers of positive cases generally. For the OrAlert the possible

variability in sensitivity is 6.2-80% and for the Oratect III 0-71%. The margin of error is relatively small for the DrugWipe 5+ device that was solely evaluated in Finland, a country with a relatively high prevalence of amphetamine use.

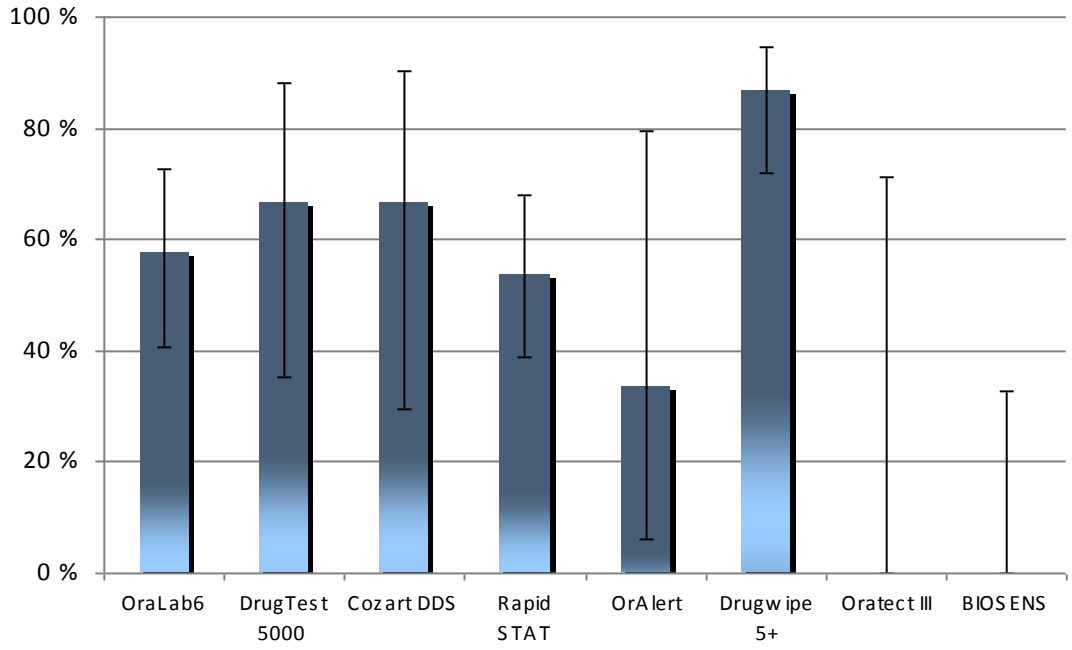

Figure 38. Combined sensitivity plots for amphetamines tests.

8.1.2. Cannabis tests

The sensitivities of the cannabis tests for each device are shown in Figure 39. For the DrugTest 5000 results for the devices with a 5 ng/ml cut-off (tested in Belgium) and 25 ng/ml cut-off (tested in the Netherlands) for Δ^9-THC have been consolidated. For cannabis tests the average sensitivity, specificity and accuracy was 38%, 95% and 73% respectively. None of the devices were near 80% sensitivity. The best sensitivity (59 %) was observed for the DrugTest 5000. In addition to this, three devices, Rapid STAT (56%), BIOSENS (50% in the successful evaluation) and DrugWipe 5+ (43%) had above average sensitivity. All the other tests had a sensitivity of less than 50%. Attention should be paid to the fact that the DRUID cut-off for THC (1ng/ml) is very low, and indeed much lower than the manufacturers cut-offs (typically 25-50ng/ml), which can explain the low sensitivity for all the devices. However, it should also be borne in mind that for the study in the Netherlands at least part of the evaluation of devices (DrugTest 5000, Rapid STAT, BIOSENS and Oratect III), and in the case of BIOSENS all of it, was performed in coffeeshops. In this setting very recent use of cannabis is to be expected, resulting in higher concentrations of THC in oral fluid which are easier to detect. For specificity, all devices had very good to excellent values (between 90-100%). The problems in sensitivity are, to some extent, reflected in the accuracy results. Six of the eight devices got moderate, low or very low accuracy scores. DrugWipe 5+ had the highest accuracy score (88%) and the second highest score was 82 % for the DrugTest 5000. Cozart DDS (71%), OrAlert (78%) and Rapid STAT (78%) were all above average for accuracy. However, the high proportion of negative cases in the study populations for DrugWipe 5+ (113 negative cases from 134 tests) and OrAlert (83 from 110) should be noted. The error bars of the plots show less variability than for the amphetamines, which reflects the higher prevalence of cannabis in the sampled populations.

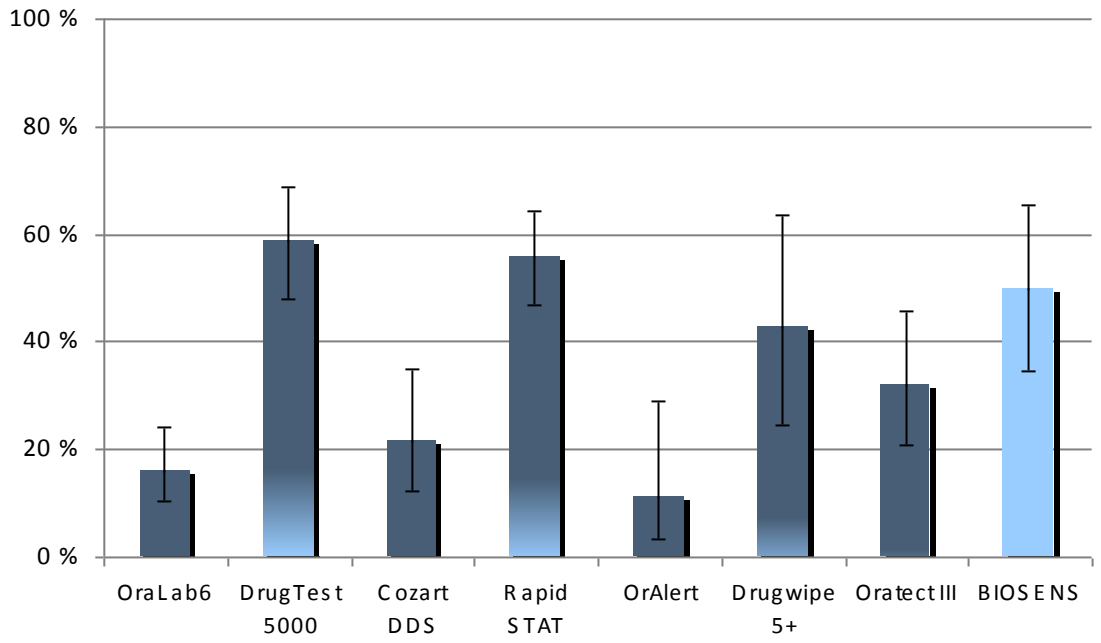

Figure 39. Combined sensitivity plots for cannabis tests

8.1.3. Cocaine tests

The sensitivities of the cocaine tests for each device are shown in Figure 40. Average sensitivity of the cocaine tests was very low at 36%. The highest scores were achieved by the DrugTest 5000 and OrAlert devices (both 50%). For the Oratect III and BIOSENS devices not enough positive cases were found to properly evaluate sensitivity (both shown as zero) and the error bars suggest this value could lie between 0-62%. No positive cases at all were found in the evaluation of the DrugWipe 5^{+}. The proportion of cocaine positive cases sampled for the devices was low, rising to a maximum of 22% for the OraLab6 (54 positive cases from 249 tests). Therefore it is not surprising that the specificity scores are 99-100% for all devices and accuracy is either very good, or excellent, for all devices except the OraLab6 (86%). Again the small number of positive cases in the study populations gives large variability for the standard error bars. In addition, the concentrations of cocaine (and benzoylecgonine) found in the positive cases were generally low which can at least partially explain the low sensitivity of the devices for cocaine.

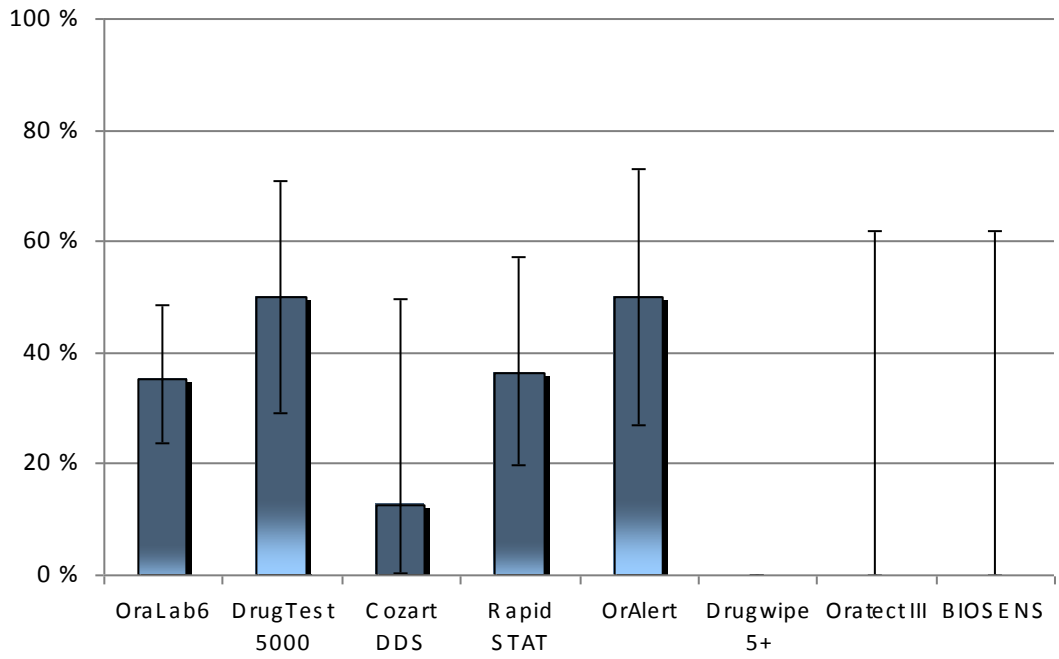

Figure 40. Combined sensitivity plots for cocaine tests

8.1.4. Opiates tests

The sensitivities of the opiates tests for each device are shown in Figure 41. Concerning the opiates tests, three devices reached 80% or more in sensitivity, specificity and accuracy. These devices were Rapid STAT (sensitivity 90%, specificity 97% and accuracy 96%), DrugTest 5000 (sensitivity 89%, specificity 94% and accuracy 92%) and Cozart DDS (sensitivity 83%, specificity 95% and accuracy 91%). For the OraLab6 and OrAlert devices sensitivity was low (69%) or moderate (73%) respectively, but values for specificity (98% and 81% respectively) and accuracy (84% and 75% respectively) were higher. Attention should be paid to the fact that the devices evaluated in the Belgian study (DrugTest 5000, Rapid STAT, OraLab6, Cozart DDS and OrAlert) were all evaluated in treatment centres for drug addiction, where a high prevalence of opiates can be expected. For DrugWipe 5^{+}, Oratect III and BIOSENS, there were not enough positive cases to reliably make the sensitivity calculation (shown as 0% for the DrugWipe 5^{+} and Oratect III devices and 50% for BIOSENS), however specificity and accuracy was very good or excellent for each of these devices (96-100%). The error bars for these devices suggest sensitivity is between 0-82.9% for the former two and 15.4-84.6% for the BIOSENS. Variability of the standard error bars for the devices tested in Belgium is small.

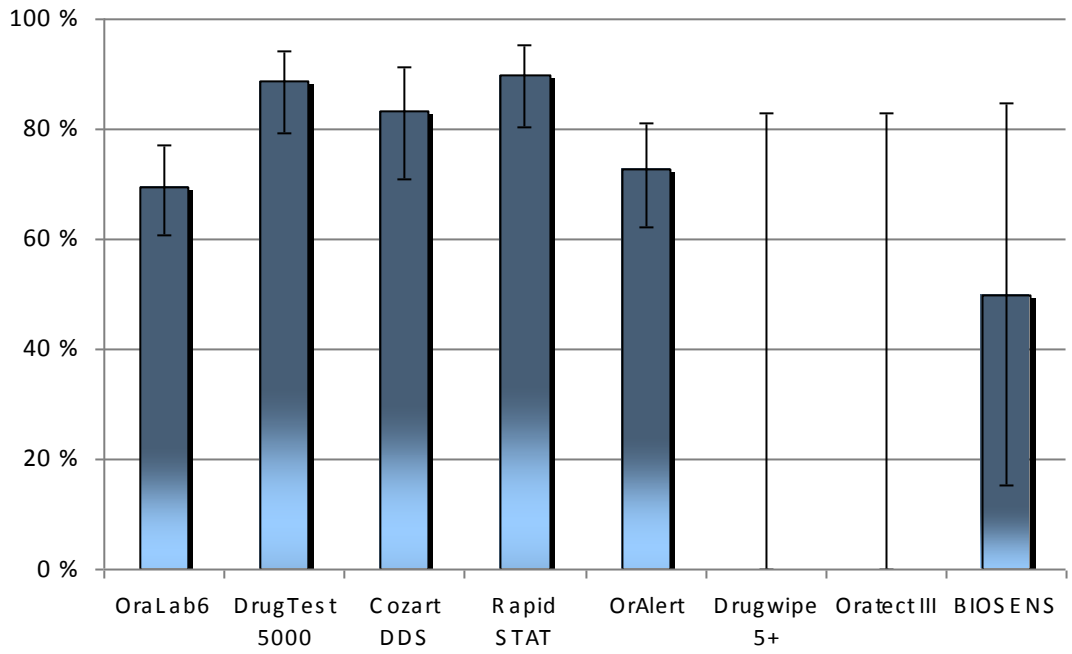

Figure 41. Combined sensitivity plots for opiates tests

8.1.5. Benzodiazepines tests

The sensitivity of the benzodiazepines tests for each device is shown in Figure 42. None of the tests for benzodiazepines had over 80 % values for all calculations. The average sensitivity was 62%. Rapid STAT and DrugTest 5000 both had above average, but still low, scores (67% and 65% respectively) whilst the Cozart DDS had very low sensitivity (48%). Specificity was again very high, or excellent, for each of these devices, and the accuracy of the Rapid STAT and DrugTest 5000 devices was very high (90% and 92% respectively) and for the Cozart DDS moderate (77%). For the BIOSENS and Oratect III benzodiazepine tests sensitivity calculations were not possible because there were no positive benzodiazepine cases sampled for either device. Both of these tests had excellent scores for specificity and accuracy, probably partly due to the absence of any positive cases. The variability of the error bar for the DrugTest 5000 is larger than for the Rapid STAT, this may be partly expected due to the fact that the latter device was evaluated in Finland, a country which has a relatively high DUI prevalence for benzodiazepine use, as well as Belgium and the Netherlands. The DrugTest 5000 was evaluated only in the latter two countries.

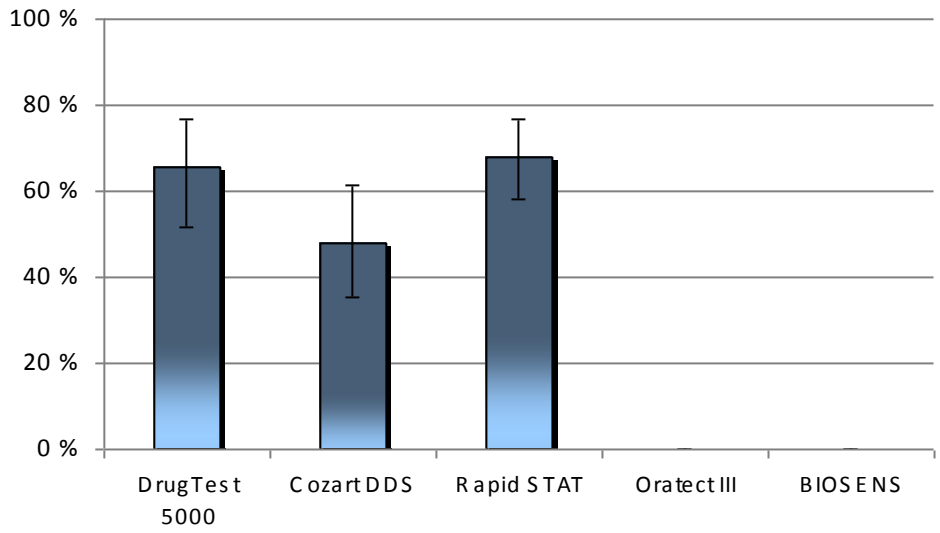

Figure 42. Combined sensitivity plots for benzodiazepines tests

8.1.6. Methamphetamine, MDMA and PCP tests

Not enough positive cases were gathered to successfully evaluate any of the devices with methamphetamines tests, and none for the Cozart MDMA test. Similarly, due to the extreme rarity of PCP in Europe, no positive cases were gathered to allow evaluation of the OraLab6 and OrAlert PCP tests. It is notable the OrAlert PCP test gave a small number of false positive results, at least some of which are probably attributable to the presence of the anti-depressant venlafaxine and its main metabolite o-desmethylvenlafaxine. Specificity and accuracy of all the methamphetamine, MDMA and PCP tests were 98% or above.

8.1.7. Device failures

A number of device failures were observed in the study. The reasons for device failure may vary, for example, the device is used incorrectly or only part of the integrated device is successful (i.e. there is no control line, indicating a successful negative or positive screening, for one of the test strips). Therefore, for at least some of the tested devices, only some of the individual drug tests failed. 15 OrAlert devices were observed to fail in the Belgian study, a smaller number of Rapid STAT, DrugTest 5000 and OraLab6 devices also completely failed (5, 2 and 1 respectively). In the Finnish study one DrugWipe 5^{+} device failed, except for the amphetamines test and six Rapid STAT devices either completely, or partially, failed. In the Netherlands the only unsuccessful tests were for the BIOSENS cocaine test (15 failed tests on the second analysis) and one DrugTest 5000 methamphetamine test. The roadside analysis of the Oratect III was also aborted due to the failure of a number of tests, however in the coffeeshop all tests with the Oratect III were success.

8.1.8. Comparison to the Rosita-2 project

In the Rosita-2 project the sensitivities of the devices for amphetamines varied from 40% to 83% and specificity from 80% to 100%. The amphetamine results obtained in DRUID are slightly better for sensitivity and specificity (sensitivity 54-87%, specificity 88-100%). Also for opiate and benzodiazepine tests some minor development can be seen. Opiate sensitivity was 51-100% in the Rosita-2 project and 69-90% in DRUID (average sensitivity 79%). For benzodiazepines, sensitivity was 33-60% in the Rosita-2 study whereas in DRUID values of 48-68% were seen. Specificity values for the opiate and benzodiazepine tests obtained in DRUID are somewhat higher that the values obtained for these in Rosita-2, but it should be remembered that for many of the devices tested in DRUID there were no positive cases for these substances. Unfortunately, no significant improvements can be seen for the

cannabis and cocaine tests since the Rosita-2 project. On the contrary, the sensitivity values for cocaine did not reach the level obtained in the earlier study.

It should, however, also be remembered that at least some of the devices tested in Rosita-2 are from different manufacturers than those tested in this DRUID study. Nonetheless, in general, devices can be expected to improve in performance as the field develops. Comparison of the analytical cut-offs used in Rosita-2 and DRUID shows that they are generally about the same, or only slightly lower for the DRUID project. Thus comparison of Rosita-2 results to the DRUID results gives a good insight to the direction of the development of on-site testing. The main differences observed in cut offs are for the benzodiazepines (e.g. 1ng/ml instead of 5ng/ml for nordiazepam, lorazepam and clonazepam). In addition, some of the benzodiazepines analysed in the Rosita-2 project were not included as compulsory substances in DRUID so direct comparison is impossible. Nevertheless, better results against lower cut-off values would appear to indicate progress for these benzodiazepines tests.

8.1.9. Summary of individual substance test results

Taken as a whole, the results of Task 3.2 indicate that the opiate tests seem to be at moderate or good levels even when DRUID cut-offs, which are lower than most of the manufacturers cut-offs for opiates, are used as a target level. In addition, there was one promising amphetamine test. For benzodiazepines, the tests did not perform very well when considering sensitivity. In particular it is evident that cannabis and cocaine tests of the devices still lack sensitivity when evaluated against the DRUID cut-offs. It is worth noting that when there is a large number of compounds included in a substance category, e.g. benzodiazepines, the on-site tests do not necessarily cover all the compounds that should be screened for. It is also possible that the tests might have cross reactivity for some substances even though the manufacturers do not report this. This can occur simply because the manufacturer has not tested the device for all the compounds within a category, for example, benzodiazepines (e.g. bromazepam), or due to unexpected cross reactivity (e.g. venlafaxine for PCP tests).

Somewhat disturbingly, none of the devices in the DRUID study performed well (at above 80% for sensitivity, specificity and accuracy) for all of the separate tests that they comprised. Nonetheless, the DrugTest 5000 had the best overall results, with a sensitivity, specificity and accuracy of 50%, or above, for all substances. The next best device was Rapid STAT that gave sensitivity, specificity and accuracy scores of above 50% for all drug groups with the exception of the sensitivity for cocaine. Neither of these devices was successfully evaluated for MDMA. When comparing the results for different devices from the separate countries at the device cut-offs, the performance of the tests look a bit more promising. For many of the substance groups the individual test cut-offs of the devices are still too high. In contrast the DRUID cut-offs were set at the beginning of the DRUID project, for an epidemiological study, at a sufficiently low level to allow optimal detection of positive DUI cases.

8.1.10. Use in police enforcement activities

In practice, for example, in police enforcement, emphasis should be put on the composition, and likely substances of abuse to be found, of the tested population when choosing which on-site test to use. If most of the tested persons are likely to be cannabis consumers, a test with good sensitivity, specificity and accuracy for cannabis should be chosen. Also, when planning and evaluating a study of on-site devices particular attention should be paid to the selection of the study population. If, for example, a methamphetamine test is to be evaluated as part of the device, effort should be made to obtain enough positive cases to enable sensitivity calculations. Furthermore, if the study population (e.g. a randomly selected group) differs greatly from the population for which the device is intended to be used (e.g. DUI suspects) then positive predictive values and negative predictive values, used to ascertain the usefulness of a device, should be calculated using the relevant drug prevalence of the latter population. For the positive predictive value and negative predictive value calculations contained within the individual country reports of this evaluation DUI prevalences were used. The values calculated in the individual country reports indicate the usefulness of the screening device for DUI populations in each separate country involved in the study and so shall not be discussed further here.

8.2. Checklist for clinical signs of impairment evaluation

Oral fluid screening devices for the detection of drugged drivers are relatively expensive and their use is time-consuming. Furthermore, according to the results of the EU research project IMMORTAL, the drug-related road toll is a lot smaller than the alcohol-related one (21). In order to make drug-driving enforcement cost beneficial some kind of preselection of suspected drivers seems to be imperative. Such a preselection procedure is already common practice in Victoria, Australia, where times and places with high drug-driving incidence are selected for police enforcement activities. Subsequently, drivers are preselected for drug testing, based on clinical signs and symptoms of drug-related impairment (23). A further increase in cost-effectiveness may be achieved by drug testing alcohol positive drivers in particular, not only because of their higher exposure to drugs than alcohol negative drivers (15% vs. 5%), but also because of the extremely high risk of combined alcohol and drug use (24).

Evaluation results of the Belgian checklist showed that the correlation between signs and symptoms of drug use and the actual presence of drugs was very low. The pupil reaction tests seemed to be the best predicting parameters, especially for amphetamines and THC. Furthermore, it was observed that concentrations leading to a significant correlation were often a lot higher than the DRUID cut-off. The clinical signs of impairment checklist outcomes were mainly true positive in subjects who either took drugs very recently and/or in high quantities.

For the Finnish clinical signs of impairment evaluation it was notable that as many as half of the drug positive DUI cases examined displayed three or fewer symptoms from the checklist. It is debateable whether or not so few possible signs of impairment, which may have been due to other factors, could be said to be suspicious. It was notable that the cases in which most signs of impairment were detected were almost exclusively those from the DUI sampling by police, relatively few symptoms, or often none at all, were observed in the non suspect cases performed by researchers. Overall, the fact that the majority of cases included to the Finnish clinical signs of impairment study were non-suspect and there were relatively few assessed suspected drug driver cases, makes it hard to draw any meaningful conclusions.

The evaluation results of the Dutch clinical signs of impairment checklist for sensitivity for detecting drug use, were low (13%). Including the self-reported drug users that the checklist performers had judged as non-users increased the sensitivity somewhat, but still to only quite a low level (32%).This compares quite poorly to previous assessments of the same checklist in for the EU research project IMMORTAL, in which most of the assessments were made by police officers working continuously in the project, who thus gained experience.

Results of the checklist for clinical signs of impairment evaluation may be improved by training of police officers and by selecting times and places with a higher drug driving incidence. Police officers evaluating clinical signs of impairment can also be expected to more efficiently detect signs of impairment through gaining experience.

8.3. Overall evaluation for police enforcement

8.3.1. Method of evaluation

In addition to evaluating the individual substance tests of the on-site devices an ‘overall evaluation’ was performed to gage the usefulness of the devices as a means of strengthening a police officers possible initial suspicion of drug use in a DUI suspect. For this evaluation a true positive was considered to be a positive screening result from the device, for any substance category, followed by a positive finding for a substance in the oral fluid confirmation sample, even if the substance found in the confirmation sample was not detected by the device (e.g. a case with a positive screening result for amphetamines only is found to be positive for cannabis only in the confirmation sample). In drug enforcement practice false positive results could be viewed as only problematic if no substances at all are detected in the confirmation analysis. In a case where another substance is detected the result of a false positive screening for a specific substance can be regarded as a true positive for drugs overall. This comparison is somewhat complicated by the fact that the devices do not test for all the same

substance groups and do not necessarily comprise the same number of individual substance tests (e.g. DrugWipe 5^+).

Only the second 'successful' evaluation of the BIOSENS device was included for this evaluation since the failure of the cannabis test to function properly, as in the first evaluation, could be expected to give a lower bias to the sensitivity results as a result of testing in a coffeeshop, where the principal narcotics findings are reasonably expected to be Δ^9-THC.

8.3.2. Overall evaluation results

The resulting sensitivities, specificities and accuracies of the devices can be seen in Table 61 - Table 63 (Annex 9). The sensitivities and their error bars are shown in Figure 43.

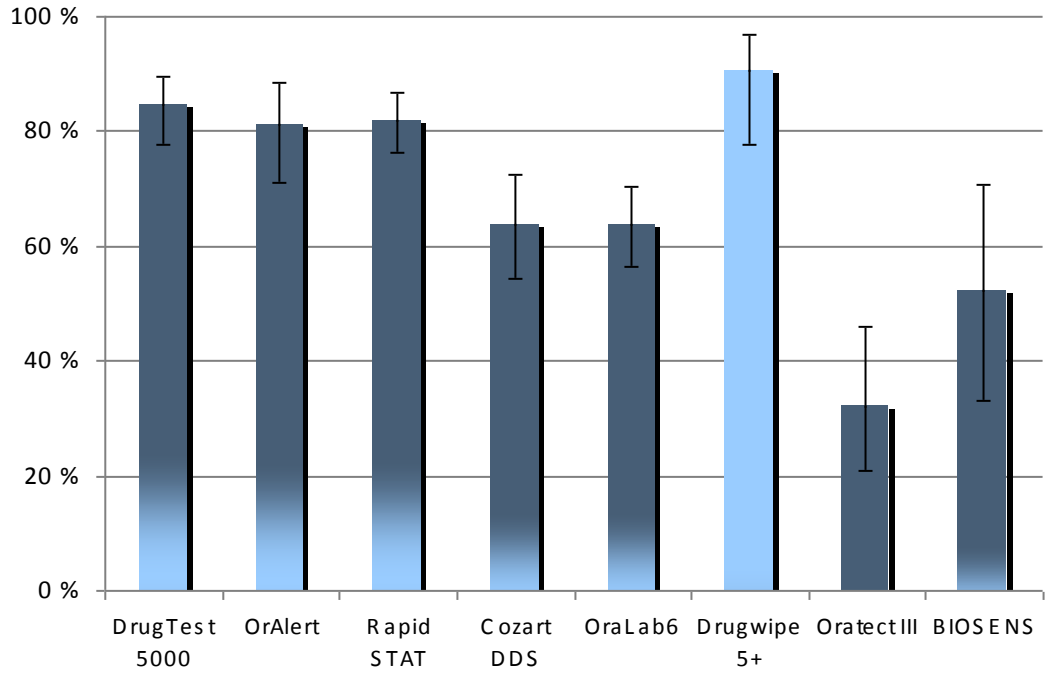

Figure 43. 'Overall' sensitivity for each device in the study, with error bars.

The sensitivity of each device as a function of specificity is shown in Figure 44. Using the above model of evaluation it can be seen that the DrugWipe 5^+ delivers the best results for sensitivity (91%) whilst also performing very highly in terms of specificity (95%). However the margins of error (95% confidence interval) displayed in Figure 43 show that this value could vary between 78-97%, this margin of error would seem to be due to the size of the study population (135 tests performed) since the device was only tested in Finland. The strong results for this device probably reflect largely on the device's high performing individual amphetamines test in a country with a relatively high prevalence for amphetamines. However, this overall sensitivity is still higher than the individual sensitivity of the amphetamines test for DrugWipe 5^+ (87%) indicating that the device was successful in screening for other drugs. Both DrugTest 5000 and Rapid STAT also performed strongly in this evaluation both for sensitivity (85% and 82% respectively) and specificity (86% and 88% respectively), which is a reflection of their generally relatively good performance for each individual substance test. The sensitivity error margins are also somewhat narrower for these two devices that were tested on a greater number of subjects (220 and 342 tests performed respectively). The OrAlert device also performs at a high level of sensitivity (81%) in this evaluation, however the specificity is somewhat lower at 70% - which is the lowest score for any of the devices. The sensitivities of the other four

devices included in the study range between 64% and 32%, which are quite low values. The specificities are, however, very high, or excellent, at between 93% and 100%. The relatively large error bars for the Oratect III device and BIOSENS can be attributed to the number of successful evaluations (58 and 25 respectively).

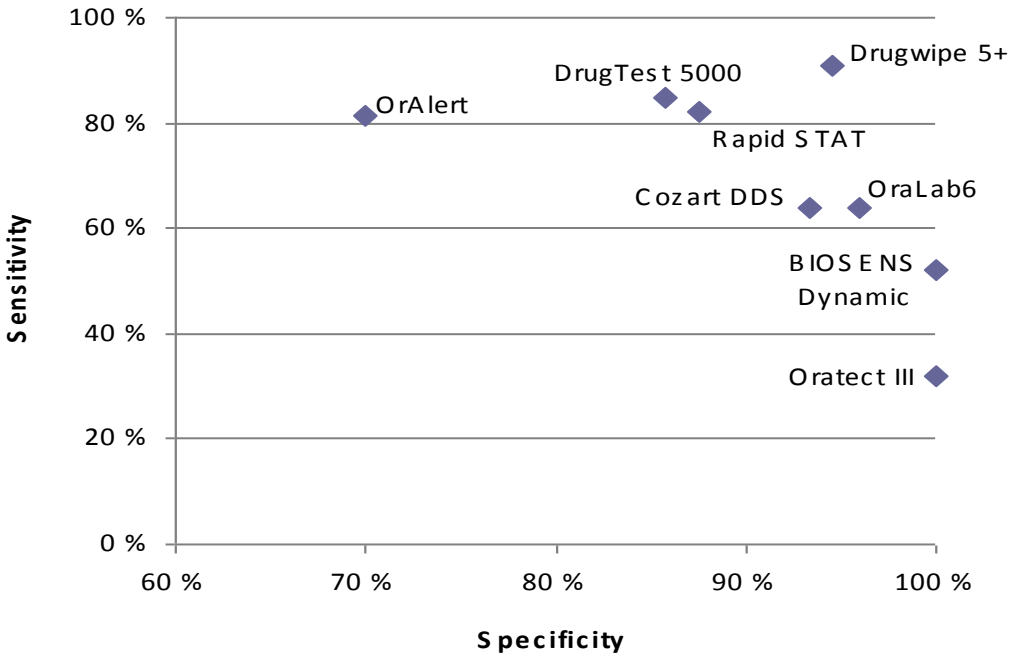

Figure 44. Sensitivity vs. specificity for each device.

The findings of this overall evaluation largely reflect the results discussed in the previous sections, nonetheless as a means of assessing the devices it should be remembered that the results from this analysis may rely, to some extent, on chance. A device which falsely detects one substance whilst missing another cannot be said to be analytically reliable.

As previously noted, the overall evaluation performance of the DrugWipe 5^{+} can be largely attributed to the strong individual performance of the device's amphetamines test and the prevalence of these substances in the study population. Similarly, it is worth reflecting upon the fact that the overall sensitivity results for devices in the Belgian study are significantly reduced when the opiates screening results are not considered. The sensitivity of each of these devices is therefore enhanced, to some extent, by the fact that the Belgian study was largely carried out with samples collected from drug addiction centres with a high prevalence of opiates. A similar outcome can also be expected to be true for devices tested in coffeeshops in the Netherlands, due to the high prevalence and sample concentrations for cannabis as mentioned above. While a high prevalence of an individual substance, or group of substances, in the sampling group should be considered it would be extremely difficult, or impossible, to test all the devices on a study population with a high prevalence of all the substances concerned, even more so since this study was carried out in three countries.

9. Conclusions

To date, oral fluid screening devices for the detection of drugs have been used in only a few countries, but an increasing number are planning to introduce them as a legal screening device. The benefits of using oral fluid for drug screening purposes is that recent drug use can better be detected in oral fluid than in urine, sweat or hair. On top of that, oral fluid collection is much less invasive than urine collection.

Theoretically, the largest general deterrence effects on drug driving may be expected from large-scale random drug testing, as is the case with random breath testing for alcohol (25). However, the time-consuming process of on-site oral fluid screening, in combination with the quite high cost of the devices and the relatively low sensitivity for cannabis, which in many countries is the most frequently used illegal drug, will probably prevent large-scale random drug testing in practice.

The results of the evaluation of each device need to be viewed in the context of the study population on which they were tested. For some of the devices, a full performance evaluation was not possible for all of the test strips on the panel due to low prevalence of the substance(s) in question. Sensitivity is usually enhanced to some extent if the study population has a high prevalence for the screened drug and if the concentrations of the drugs contained in the samples from the study population are high because of recent consumption. Conversely, when interpreting specificity values it should be noted that when a population with a low prevalence of the desired substance is tested, specificity can be expected to be high. Such a population can also be expected to result in higher accuracy results in a similar manner. Also, it should be borne in mind that sensitivity, specificity and accuracy are specific for this study and the study populations investigated in this study. Positive predictive values and negative predictive values, calculated with drug prevalences for the population for which the screening is intended, should be considered as factors too when selecting which on-site device to use.

It is disturbing that the sensitivities of the cannabis and cocaine tests were all quite low, although further testing of the cocaine tests is desirable due to the low prevalences and the low concentrations encountered in this study. There are several countries in Central and Southern Europe for which these two substance classes are of special interest. On the other hand, it seems the sensitivities of the devices are generally better for amphetamines, a frequently encountered drug class among the DUI drivers in the Nordic countries. The suitability of the device for the intended national DUI population should also be considered, for example, PCP is rarely, if ever, found in Europe, therefore at the current time utilising a PCP test is unnecessary. Since the on-site tests are relatively expensive the suitability of all the individual substance tests incorporated in the device should be considered.

The evaluation showed that none of the evaluated tests is on a desirable level (>80% for sensitivity, specificity and accuracy) for all of the separate tests that they comprised. However, there were tests that performed already on a promising level for one or more substance classes. The DrugTest 5000 had the best overall results. The next best device was Rapid STAT, which performed at a similar level, except for the cocaine test which was somewhat less sensitive. Clearly the best device in terms of sensitivity for amphetamines was the DrugWipe 5^{+}.

It should be noted that during the study sampling for screening and confirmation analysis took place almost simultaniously. In the enforcement practice however blood sampling for comfirmation purposes may be seriously delayed. For substances with a low half-life in particular, e.g. cocaine, this may result in an increased number of false positive screening results.

For cost-benefit purposes, a working method to preselect suspected drivers for on-site drug screening, would be desirable. Unfortunately, the evaluation of the CSI checklist in this project did not give very encouraging results. Apparently, at least for persons with little observational training for clinical signs of impairment, or only relatively short-term experience of this, symptoms of drug use remain easily undetected. Also, correlation between signs of impairment and findings in oral fluid was not very good. However, proper training and long-term experience of observing clinical signs of impairment could reasonably be expected to yield better results.

The effectiveness of drug driving enforcement can be further enhanced by preselecting times and places with a likelihood of elevated numbers of drug positive drivers and by targeting alcohol positive drivers. This is not only because alcohol positive drivers are likely to have a higher exposure to drugs

than alcohol negative drivers, but also due to the fact that the risk of combined alcohol and drug use is extremely high.

References

General section:

1. Aps JKM, Martens LC. Review: The physiology of saliva and transfer of drugs into saliva. Forensic Sci Int. 2005;150:119-31.
2. Menkes DB, Howard RC, Spears GFS, Cairns ER. Salivary THC following cannabis correlates with subjective intoxication and heart rate. Psychopharmacology (Berl). 1991;103:277-9.
3. Cone EJ, Oyler J, Darwin WD. Cocaine disposition in saliva following intravenous, intranasal, and smoked administration. J Anal Toxicol. 1997;21(6):465-75.
4. Samyn N, A.G. V, van Haeren C, Kintz P. Analysis of Drugs of Abuse in Saliva. Forensic Science Reviews. 1999;11(1):1-19.
5. Schramm W, Smith RH, Craig PA. Drugs of abuse in saliva: a review. J Anal Toxicol. 1992;16:1-9.
6. Verstraete AG, Raes E, editors. Rosita-2 project Final Report. Gent: Academia Press; 2006.
7. Blencowe T, Lillsunde P. Protocol of "Workshop on drug driving detection by means of oral fluid screening". DRUID report; 2009.
8. Kuijten C. Evaluation of oral fluid Screening devices by TISPOL to Harmonise European police Requirements (ESTHER). DRUID Report; 2009.
9. Langel K, Engblom C, Pehrsson A, Gunnar T, Ariniemi K, Lillsunde P. Drug testing in oral fluid - Evaluation of sample collection devices. J Anal Toxicol. 2008;32(6):393-401.
21. Mathijssen MPM, Houwing S. The prevalence and relative risk of drink and drug driving in the Netherlands: a case-control study in the Tilburg police district. Leidschendam: SWOV; 2005.
22. Agresti A, Coull BA. Approximate is better than "Exact" for interval estimation of bionomial proportions. The American Statistician. 1998;52:119-26.
23. Boorman M, Owens K. The Victorian legislative framework for the random testing drivers at the roadside for the presence of illicit drugs: an evaluation of the characteristics of drivers detected from 2004 to 2006. Traffic Injury Prevention. 2009;10:16-22.
24. Mathijssen MPM, Houwing S. Risks posed by drugs in traffic. European Union Research Project IMMORTAL. The Risk of Drink and Drug Driving - Results of a Case-Control Study Conducted in the Netherlands. In: Transportation Research Circular E-C096. Transportation Research Board, Washington, DC; 2006.
25. Homel R. Policing and Punishing the Drinking Driver: A Study of General and Specific Deterrence. Research in Criminology. Springer Verlag; 1998.

Country report Belgium:

10. Matuszewski B. Standard line slopes as a measure of a relative matrix effect in quantitative HPLC-MS bioanalysis. J Chromatogr B Analyt Technol Biomed Life Sci. 2006;830(2):293-300.
11. Raes E, Verstraete AG. Usefulness of roadside urine drug screening in drivers suspected of driving under the influence of drugs (DUID). J Anal Toxicol. 2005;29(7):632-6.
12. Santos PM, Lopez-Garcia P, Navarro JS, Fernandez AS, Sadaba B, Vidal JP. False positive phencyclidine results caused by venlafaxine. Am J Psychiatry. 2007;164(2):349.
13. Sena SF, Kazimi S, Wu AHB. False-positive phencyclidine immunoassay results caused by venlafaxine and O-desmethylvenlafaxine. Clin Chem Lab Med. 2002;48(4):676-7.

Country report Finland:

14. Rajala P, Ariniemi K, Mykkanen S, Lillsunde P. Screening for abused drugs in blood samples using modified EMIT immunoassay. 6th Conference of the International Society of Addiction Medicine; Helsinki. 2004.

15. Gunnar T, Ariniemi K, Lillsunde P. Fast gas chromatography-negative-ion chemical ionization mass spectrometry with microscale volume sample preparation for the determination of benzodiazepines and alpha-hydroxy metabolites, zaleplon and zopiclone in whole blood. J Mass Spectrom. 2006;41(6):741-54.

16. Gunnar T, Engblom C, Ariniemi K. Pressure-adjusted continual flow heart-cutting for the high throughput determination of amphetamine-type stimulant drugs in whole blood by fast multidimensional gas chromatography-mass spectrometry. J Chromatogr A. 2007;1166(1-2):171-80.

17. Gunnar T, Mykkanen S, Ariniemi K, Lillsunde P. Validated semiquantitative/quantitative screening of 51 drugs in whole blood as silylated derivatives by gas chromatography-selected ion monitoring mass spectrometry and gas chromatography electron capture detection. J Chromatogr B Analyt Technol Biomed Life Sci. 2004;806(2):205-19.

18. Kankaanpaa A, Gunnar T, Ariniemi K, Lillsunde P, Mykkanen S, Seppala T. Single-step procedure for gas chromatography-mass spectrometry screening and quantitative determination of amphetamine-type stimulants and related drugs in blood, serum, oral fluid and urine samples. J Chromatogr B Analyt Technol Biomed Life Sci. 2004;810(1):57-68.

19. Pehrsson A, Gunnar T, Engblom C, Seppa H, Jama A, Lillsunde P. Roadside oral fluid testing: Comparison of the results of DrugWipe 5 and DrugWipe Benzodiazepines on-site tests with laboratory confirmation results of oral fluid and whole blood. Forensic Sci Int. 2008;175(2-3):140-8.

Country report the Netherlands:

20. Möller M. Drogenerkennung im Strassenverkehr. Berichte der Bundesanstalt für Strassenwesen, Mensch und Sicherheit M 96. Bergisch Gladbach: Bundesanstalt für Strassenwesen BASt; 1998.

21. Mathijssen MPM, Houwing S. The prevalence and relative risk of drink and drug driving in the Netherlands: a case-control study in the Tilburg police district. Leidschendam: SWOV; 2005.

Annex 1

Table 42. DRE checklist used in evaluation of Varian OraLab6.

Signs of impairment	**O** unsteady on one's feet, swaggering **O** uncontrolled movements **O** dizzy, sleepy **O** euphoric **O** not understanding instructions **O** incoherent speech **O** excessive talking **O** mumbling **O** low, scraping voice **O** scratching face **O** trembling **O** shaking leg **O** excited, agressive behaviour **O** bloodshot eyes **O** red nostrils **O** trembling eyelids **O** sniffing **O** excessive sweating **O** swallowing **O** smell of marihuana **O** small pupils: < 3.0 mm **O** large pupils: > 6.5mm
Nystagmus test	**O** nomal reaction **O** shaking movement of pupil
Reaction of pupil to light	**O** normal reaction **O** slow reaction

Annex 2

Table 43. General parameters used for mass spectrometry.

ES+ Source	Capillary Voltage	0.8 kV
	Extractor Voltage	4 V
	RF Lens	0 V
	Source Temperature	140 °C
	Desolvation Temperature	450 °C
	Desolvation Gas Flow	1000 l/h
	Cone Gas Flow	50 l/h
Analyser	Collision Gas Flow	0.15 ml/min

Table 44. MRM transitions, dwell times, cone voltage, collision energy and retention times for all standards and internal standards.

Substance	Q 1	Q 3	Dwell time (msec)	Cone (V)	Collision Energy (eV)	Retention time (min)
6-acetylmorphine	328.12	152.08	35	47	61	3.61
6-acetylmorphine-D 3	331.10	164.90	35	45	37	3.58
7-amino-clonazepam	286.08	222.05	35	41	25	2.79
7-amino-clonazepam-D4	290.02	226.00	35	39	27	2.78
7-amino-flunitrazepam	284.14	226.86	35	39	25	2.96
7-amino-flunitrazepam-D7	291.11	230.30	35	39	27	2.94
Alprazolam	309.01	204.94	25	49	39	4.34
Amphetamine	136.07	119.05	35	15	9	3.43
Amphetamine-D5	141.01	92.90	35	17	27	3.37
Benzoylecgonine	290.14	168.00	35	33	19	2.64
Benzoylecgonine-D3	293.10	171.00	35	33	19	2.63
Bromazepam	316.03	181.96	30	41	29	3.91
Citalopram	325.11	262.10	25	37	25	4.79
Clonazepam	316.08	269.96	30	45	25	4.02
Cocaine	304.11	182.10	15	31	19	4.66
Cocaine-D 3	307.10	185.00	15	31	19	4.64
Codeine	300.14	165.01	30	41	43	3.84
Codeine-D 3	303.10	215.00	30	45	25	3.83
Diazepam	285.08	222.02	25	43	27	4.71
Diazepam-D 5	290.08	227.00	25	43	27	4.69
Flunitrazepam	314.08	268.09	30	41	25	4.09
Lorazepam	321.02	229.03	40	29	29	4.30
MDA	180.02	105.03	35	15	21	3.28
MDA-D5	185.01	110.00	35	17	21	3.23
MDEA	208.10	162.97	30	23	13	3.79
MDEA-D 5	213.07	162.90	30	21	13	3.69
MDMA	194.10	162.95	35	21	13	3.43
MDMA-D5	199.10	135.20	35	21	21	3.34
Methamphetamine	149.96	90.95	30	21	17	3.61
Methamphetamine-D5	155.00	120.90	30	19	11	3.55

Substance	Q 1	Q 3	Dwell time (msec)	Cone (V)	Collision Energy (eV)	Retention time (min)
Morphine	286.11	152.10	35	45	53	3.10
Morphine-D3	289.08	164.90	35	43	37	3.10
Nordiazepam	271.02	208.03	25	41	25	4.58
Oxazepam	287.05	241.01	40	33	25	4.31
THC	315.18	193.05	60	31	21	5.44
THC-D3	318.13	196.00	20	31	25	5.44

Table 45. Validation parameters for UPLC-MS/MS confirmation method.

	R^2	Extraction yield* (%)	Inaccuracy* (%)	Imprecision* (%)	Absolute matrix effect (%)[£]	Relative matrix effect[$] (CV)
6-acetylmorphine	0.999	75.1	-3.6	3.4	-11.5	1.8
7-amino-clonazepam	0.997	70.7	+1.1	4.3	11.1	1.6
7-amino-flunitrazepam	0.996	77.4	-1.2	4.3	10.2	1.9
Alprazolam	0.998	77.6	-1.6	5.3	14.9	2.6
Amphetamine	0.993	54.8	+6.2	7.9	16.2	2.0
Benzoylecgonine	0.983	2.8	+2.4	3.4	8.7	1.8
Bromazepam	0.997	77.2	-2.1	5.7	16.7	3.0
Clonazepam	0.998	84.2	-2.9	5.1	30.6	3.1
Cocaine	0.999	78.3	-2.2	3.8	3.1	1.0
Codeine	0.999	70.4	-0.8	4.8	19.7	1.6
Diazepam	0.998	84.8	-0.4	4.1	10.4	0.6
Flunitrazepam	0.998	83.8	-1.5	4.9	27.9	2.0
Lorazepam	0.996	86.5	+6.5	6.0	11	2.6
MDA	0.998	58.1	+1.5	6.6	1.3	1.8
MDEA	0.997	70.3	-3.5	4.5	11.5	1.7
MDMA	0.997	65.9	+1.8	6.8	4.2	1.5
Methamphetamine	0.997	56.5	+0.4	6.8	6.6	1.4
Morphine	0.997	39.5	+0.5	4.4	6	1.1
Nordiazepam	0.995	77.4	-4.1	5.1	17.3	2.2
Oxazepam	0.995	79.9	-1.2	5.8	21.6	2.5
THC	0.998	52.6	-3.9	6.7	93.5	1.1

* extraction yield, inaccuracy and imprecision at medium concentration (20 ng/mL)

£ absolute matrix effect determined at 100 ng/mL

$ CV of slopes of standard lines from five different sources

Annex 3

STANDARDIZED FIELD SOBRIETY OBSERVATION SHEET

Date

Concerns R-report nr.	Laboratory sample nr (filled out by KTL)
Surname and initial names	Social security nr.

OBSERVATIONS REGARDING WAY OF DRIVING, WEATHER AND ROADWAY

Way of driving

☐ No own observations ☐ Secure ☐ Unsteady ☐ Inappropriate speed ☐ Violation of way of priority

☐ Winding, deviation from straight line up to _____ meters.

Number of deviations: on a _____ meters of observation

☐ Other attentions

Control of devices of vehicle

☐ Driving with low revolutions ☐ Insecure use of gears ☐ Roaring of motor

☐ Other

Fault and defects of vehicle

☐ No ☐ Yes, what?

Weather and lighting

☐ Rain ☐ Hard wind / storm ☐ Snow / sleet ☐ Fog ☐ Daylight ☐ Dusk ☐ Dark

Roadway

☐ Good ☐ Poor ☐ Construction on way ☐ Good lighting ☐ Poor lighting ☐ Dry ☐ Wet ☐ Icy / snowy

OBSERVATIONS DURING STOPPING AND CONFRONTING

Reactivity

☐ Normal ☐ Slow ☐ Very slow

Physical deviations

☐ None ☐ Sweating ☐ Tremor ☐ Vomiting ☐ Restlessness

Appearance

☐ Neat ☐ Shabby ☐ Filthy

Speaks Finnish or Swedish

☐ Yes ☐ No ☐ Faltering

Speech

☐ Clear ☐ Sputtering ☐ Thick ☐ Lisping

Communication, sense of time and place

☐ Clear sense of time ang place ☐ Drowsy ☐ Wakes up ☐ Deep sleep / unconscious ☐ Altered

Behaviour

☐ At ease, behaved ☐ Agitated ☐ Aggressive ☐ Matey ☐ Frivolous ☐ Uninterested ☐ Defiant ☐ Weepy

Rising out of vehicle

☐ Normal ☐ Balance disturbed ☐ Has to lean on vehicle

Walking

☐ Secure ☐ Dragging ☐ Wobbly ☐ Balance disturbed

Smell of alcohol

☐ Yes ☐ No

Alcometer test

☐ Yes: time ☐ No ☐ cannot be done ☐ Refused

Positive on site tests

Cozart	Time	☐ Amphetamine	☐ Opiates	☐ Benzodiazepines	☐ Cocaine	☐ THC
Drugwipe	Time	☐ Amphetamine	☐ Opiates	☐ Benzodiazepines	☐ Cocaine	☐ THC
Other, what?		☐ Amphetamine	☐ Opiates	☐ Benzodiazepines	☐ Cocaine	☐ THC

Eyes

☐ Nothing abnormal ☐ Conjunctivas reddish ☐ Watery / gleaming ☐ Restless

Pupils

☐ Normal ☐ Dilated ☐ Contracted

Reaction to light

☐ Slow ☐ Fast

Nystagmus

☐ Jerky movement ☐ No jerking observed

Right about mm Left about mm

Lighting conditions on test site

☐ Daylight ☐ Dusk ☐ Night, streetlights ☐ Night, indoor

☐ Other, what?

Conspicuous behaviour

☐ Did not change during evaluation ☐ Increased during evaluation ☐ Decreased during evaluation

Test started: time	Test ended	**The ability of the driver**
		☐ Is not impaired ☐ Is impaired ☐ Is considerably impaired

Further information: like other observations, confiscated substances, pills, paraphernalia etc.

Time and place	Signature and name of observer

Annex 4

Validation parameters for the GC-MS methods

Table 46. Validation results for Fraction 1. For all analytes, the concentration levels used in precision and bias experiments were ULOQ, medium level and LLOQ. Recovery was determined at medium concentration level.

Analyte	Concentration range / ng/ml	Linearity / R^2	Precision / %	Bias / %	Recovery / %
Amphetamine	25-1250	0.9999	7.15/6.19/5.21	7.43/7.15/5.75	83.8
Methamphetamine	25-1250	0.9990	7.64/8.45/9.34	4.51/10.8/5.65	89.7
MDA	25-1250	0.9995	9.54/9.82/9.50	4.07/3.05/9.23	94.6
MDMA	25-1250	0.9990	8.53/7.20/7.23	6.06/8.94/4.48	90.6
MDEA	25-1250	0.9933	9.25/11.4/10.6	-0.16/7.41/-1.27	97.1
Cocaine	10-500	0.9999	6.04/5.04/8.54	7.87/4.68/4.62	85.2
THC	1-50	0.9996	7.94/6.98/13.2	8.88/5.75/-0.01	78.1
Codeine	5-250	0.9998	8.60/6.96/5.96	9.69/4.09/5.90	92.0
Morphine	5-250	0.9970	7.18/13.8/13.9	3.75/5.57/-2.38	58.1
6-MAM	1-50	0.9989	8.63/5.46/16.8	12.8/4.92/14.8	92.2

Table 47. Validation results for Fraction 2. The concentration levels used for precision and bias experiments were ULOQ, 100 ng/ml amd LLOQ. Recovery was determined at medium concentration level.

Analyte	Concentration range / ng/ml	Linearity / R^2	Precision / %	Bias / %	Recovery / %
Benzoylecgonine	10-500	0.9927	12.0/8.81/18.8	-0.77/-2.08/-9.44	25.6

Table 48. Validation results for Fraction 3. The concentration levels used for precision and bias experiments were ULOQ, medium level and LLOQ. Recovery was determined at medium concentration level.

Analyte	Concentration range / ng/ml	Linearity / R^2	Precision / %	Bias / %	Recovery / %
Diazepam	0.5-25	0.9974	6.54/6.87/13.0	9.31/4.17/0.24	89.3
Nordiazepam	0.5-25	0.9988	4.04/7.54/11.3	1.04/-3.81/-3.13	84.6
Midazolam	0.5-25	0.9860	6.57/4.59/12.7	2.80/4.78/-8.84	88.4
Flunitrazepam	0.5-10	0.9980	9.28/7.22/5.82	4.42/-2.60/-10.4	101
Phenazepam	1.25-25	0.9990	10.85/5.34/7.83	7.53/6.70/-0.75	91.2
Oxazepam	0.5-25	0.9995	6.07/1.98/2.15	5.77/1.24/0.63	94.7
Nitrazepam	0.5-25	0.9987	6.42/3.11/17.9	12.2/5.85/-10.55	78.6
Temazepam	0.5-25	0.9993	6.44/2.44/4.23	3.75/2.70/6.61	89.6
Chlordiazepoxide	2-100	0.9944	21.4/13.9/35.4	19.3/22.5/34.2	63.3
Lorazepam	0.2-10	0.9930	9.64/9.11/13.7	-7.99/-3.21/2.66	118
Clonazepam	0.5-25	0.9993	6.50/5.51/15.0	3.11/4.10/-3.66	87.3
Alprazolam	0.5-25	0.9904	11.8/5.95/10.3	-9.15/8.14/0.90	94.4
a-OH-alprazolam	0.1-5	0.9964	21.3/25.9/19.9	28.7/13.9/8.34	66.0
Bromazepam	1-50	0.9925	12.14/11.50/13.13	16.47/-6.45/6.69	88.27

Annex 5

Deviation Report: 1st round test results Druid 2009-06-26

Date of report: 2009-06-29
Reporting person: Max Zigliara
Responsible person: Per Månsson, Jonas Åkesson
Status: closed

Biosens system data
Biosens Dynamic serial no: C120346
Activator batch no: A000326
Biocell batch no: 750553
Eluent batch no: 145010

Background
The Biosens Dynamic was tested in Holland (Doetinchem) 2009-06-26 by Sjoerd Houwing (SWOV) and Max Zigliara (Biosensor Applications) to evaluate its capability to detect drugs of abuse by oral sampling. The testing is part of the DRUID project to evaluate advances in drugs of abuse testing equipment for European traffic police.
The Biosens Dynamic has in previous field tests with in the ESTHER project, and in several independent clinical studies showed a very good overall performance regarding both THC and other illicit substances.
During the tests in Holland at the coffee-shop where THC users were tested with the Biosens instrument, the system performance was not up to normal standard.

Analysis of deviation
Analysis of the detailed run files from the test stored on the system hard-drive show an abnormal binding of THC antibodies onto the Biocell sensor surface. This is detected through the frequency shift in resonance frequency of the Biocell sensors before and after the binding of antibodies.
The very high binding leads to a decreased sensitivity.
The root cause of this problem was discovered to originate from the Activator bottle, containing an abnormal concentration of THC antibodies.

Conclusion
The system sensitivity is depending on the correct concentration of antibodies being used in order to utilize the dynamic range of the Biocell sensors.
An error in production resulted in the wrong activator being used in the tests, which decreased the sensitivity of the system.
The system otherwise performed correctly, other consumables performed correctly and handling of the system was done correctly.

Corrective actions

Production will go through routines and set-up to make sure the error will not occur again. Corrective actions will be presented at the next production quality steering meeting.

Annex 6

CHECKLIST DETECTION ILLEGAL DRUGS © SWOV, 2006

Survey session No. …, Day ……., Date ……., Time span ……..
Police observer: ……..…. (police team ……) **Subject No. …..**

1. Subject gender: **O** male **O** female

2. Signs of impairment

Observation at some distance when subject enters and leaves mobile research unit:

O unsteady on one's feet, swaggering (opiates)
O uncoordinated movements (opiates)
O drowsy, sleepy (opiates)
O euphoria (opiates, XTC, amphetamine, cocaine)
O not understanding instructions (opiates)
O incoherent speech(opiates)
O chattering (XTC, amphetamine, cocaine)
O slurred speech (THC)
O low, rasping voice (opiates)
O scratching one's face (opiates)
O trembling (XTC, amphetamine, cocaine)
O shaking leg (XTC, amphetamine, cocaine)
O excited, aggressive behaviour (amphetamine, cocaine)

Observation at close range before and during breath test for alcohol:

O bloodshot eyes (THC, XTC, amphetamine, cocaine)
O red nostrils (cocaine)
O trembling eyelids (XTC, amphetamine, cocaine)
O sniffing (cocaine)
O undue perspiring (opiates)
O swallowing (THC, opiates, XTC, amphetamine, cocaine)
O smell of hash (THC)
O pinpoint pupils: < 3.0 mm (opiates)
O **dilated pupils: > 6.5 mm** (XTC, amphetamine, cocaine, sometimes THC)

3. Breath test for alcohol; result (BAC) ..... g/l (in **two** decimals)

4. Question (while waiting for breath test result):

Did you during the past 24 hours use hash, weed or other illegal drugs?

O no

Oyes,namely.:.....……
Nameof medicine, if used:...…….….....(checklist benzos!)

5. Test pupil reaction to light (only if pupils are dilated and BAC is below 0.2 g/l):

O normal reaction

O slowed reaction (XTC, amphetamine, cocaine, THC)

6. Police officer's final judgement about the use of impairing drugs other than alcohol

O did not use

O did use; substance: .....…….

O in doubt between use or not, because:…….....

SEE OVERLEAF FOR SOME IMPORTANT REMARKS

SOME IMPORTANT REMARKS

1. Most of the listed signs of impairment may also be caused by other chronic or acute factors than impairing drug use, e.g. alcohol use, physical or mental disability, illness, sleeplessness, fatigue, climatologically circumstances, etc.
2. The following signs are nearly always associated with impairing drug use:
   a. smell of hash
   b. pinpoint pupils, especially under adverse daylight conditions
   c. slowed reaction of dilated pupils to light.

Therefore, these signs deserve special attention of the observer!

Annex 7

Materials and methods

DRUID Netherlands Forensic Institute

Analytical conditions

LCMS analysis was performed on a Water Acquity UPLC®-system with a Waters Quatro premier XE triple quadrupole mass spectrometer. Chromatography employed a reversed-phase UPLC ® column (BEH C-18, 100 x 2.1-mm i.d., 1.7 µm particle diameter) and a 17-min gradient elution (methanol / 10 mM ammonium bicarbonate pH 10.0, 5/95 to 95/5). The UPLC® injector was modified for on-line dilution of the injected sample to allow large injection volumes of acetone. The eluent was introduced to the electrospray source of the triple quadrupole MS instrument at a flow-rate of 500 µL/min. Molecular ions were fragmented using optimized collision-induced dissociation voltages for each compound (9 to 50 eV, positive ion mode). For each target-compound two MRM were monitored and for each deuterated internal standard one MRM was monitored.

Table 49. Internal standards, protonated molecules [M+H]+, monitored fragments, retention times (RT), and collision energies (CE) of the drugs of abuse in the order of retention time.

Analyte	Internal standaard	RT *min*	$[M+H]^+$ *m/z*	CV *(V)*	Fragm. 1 *m/z*	CE1 (eV)	Fragm. 2 *m/z*	CE2 (eV)
Benzoylecgonine	Methylecgonine-d3	3.3	290.2	25	168.1	20	105	30
Aminoflunitrazepam	7-aminoflunitrazepam-d7	3.8	284.2	35	135.2	25	240.2	30
Morphine	Morfine-d3	3.9	286.2	36	165.1	40	153.1	30
MDA	MDA-d5	4.4	180.2	10	163	10	105	20
Amphetamine	Amfetamine-d8	4.6	135.9	10	119	10	90.9	15
6-MAM	6-monoacetylmorfine-d6	4.8	328.2	35	165.1	35	211.2	25
MDMA	MDMA-d5	5.1	194.2	15	163.1	15	105	25
Codeine	Codeine-d3	5.4	300.3	35	215.3	25	165.1	30
Bromazepam	Desmethylflunitrazepam-d4	5.4	316.1	30	182.2	35	209.2	25
Metamphetamine	Methylamfetamine-d8	5.5	150	15	91	15	119.1	10
Nitrazepam	Nitrazepam-d5	5.6	282.1	25	236.2	25	180.2	35
Clonazepam	Nitrazepam-d5	5.8	316.1	35	270.2	25	214.2	35
MDEA	MDEA-d5	6.1	208.3	15	163.1	15	105	25
Flunitrazepam	Flunitrazepam-d7	6.2	314.2	30	268.3	25	239.2	35
Clobazam	Oxazepam-d5	7.3	301.2	25	259.2	20	224.2	35
Oxazepam	Oxazepam-d5	7.5	287.2	25	241.2	25	104	35
Lorazepam	Lorazepam-d4	7.8	321.1	25	275.2	25	229.2	35
Alprazolam	Alprazolam-d5	8.4	309.3	35	281.3	25	274.2	25
Triazolam	Triazolam-d4	8.7	343	35	239.2	40	111.1	50
Desalkylflurazepam	Desalkylflurazepam-d4	8.8	289.2	35	140	30	226.2	30
Temazepam	Temazepam-d5	8.9	301.2	20	255.2	20	177.1	35
Chlordiazepoxide	Desmethyldiazepam-d5	9.9	300.2	20	227.2	25	165.1	45
Lormetazepam	Desmethyldiazepam-d5	10.1	335	20	289.2	20	177.1	40
Desmethyldiazepam	Desmethyldiazepam-d5	10.8	271.2	35	140	25	165	25
Cocaine	Cocaine-d3	11.4	304.2	25	182.2	20	105	35
Diazepam	Desmethyldiazepam-d5	11.6	285.2	30	154	25	193.2	30
Midazolam	Desmethyldiazepam-d5	11.8	326.2	35	244.2	25	291.3	25
THC-COOH	11-nor-9-carboxy-delta9-THC-d9	12.1	345.2	25	299.4	20	327.3	15
Flurazepam	Desalkylflurazepam-d4	12.6	388.1	25	315.2	25	288.3	25
THC	Delta 9-THC-d3	15.3	315.3	25	193.2	20	259.5	30

Validation

Selectivity:
No interfering compounds were present in blank blood and oral fluid samples.

Precision:
The within-day precision was determined on two concentration levels B and F by repeated analysis (n=8). All compounds showed a within-day precision of < 20 % in both blood and oral fluid (data not shown).

Table 50. Validation results of drugs of abuse in blood. Analytes are shown in the order of retention time.

Analyte	Calibration range (mg/L)	Correlation coefficient (R^2) *	Reproducibility (%)	Accuracy (%)	LOD (µg/L)	LOQ (µg/L)
Benzoylecgonine	0.015-3.0	1.000 (2^{nd})	6	117	0.1	15
Aminoflunitrazepam	0.001-0.2	0.999	10	94	0.1	1
Morphine	0.005-1.0	0.999	6	96	0.1	5
MDA	0.01-2.0	0.999	5	101	0.1	10
Amphetamine	0.01-2.0	0.999	16	95	0.1	10
6-MAM	0.002-0.4	0.999	7	92	0.02	2
MDMA	0.01-2.0	1.000 (2^{nd})	4	102	0.1	10
Codeine	0.02-4.0	0.999 (2^{nd})	6	103	0.2	20
Bromazepam	0.002-0.4	0.999	12	89	0.1	2
Metamphetamine	0.005-1.0	0.999	10	91	0.1	5
Nitrazepam	0.002-0.4	0.999	17	87	0.01	2
Clonazepam	0.001-0.2	0.999	11	96	0.04	1
MDEA	0.005-1.0	0.999	5	91	0.03	5
Flunitrazepam	0.001-0.2	0.999	15	99	0.02	1
Clobazam	0.005-1.0	0.999	5	106	0.02	5
Oxazepam	0.01-2.0	0.999	5	96	0.1	10
Lorazepam	0.002-0.4	1.000 (2^{nd})	7	112	0.1	4
Alprazolam	0.001-0.2	0.999	6	93	0.01	1
Triazolam	0.001-0.2	0.999	7	107	0.03	1
Desalkylflurazepam	0.002-0.4	0.999	7	95	0.01	2
Temazepam	0.01-2.0	0.999	5	94	0.04	10
Chlorediazepoxide	0.01-2.0	0.999	5	93	0.1	10
Lormetazepam	0.001-0.2	0.995	6	94	0.02	1
Desmethyldiazepam	0.005-1.0	0.999	7	104	0.1	5
Cocaine	0.005-1.0	0.999	5	106	0.04	5
Diazepam	0.005-1.0	0.999	7	92	0.4	5
Midazolam	0.002-0.4	0.999	8	86	0.03	2
9-COOH-THC	0.001-0.2	0.991	9	96	0.5	1
Flurazepam	0.001-0.2	0.999	15	78	0.01	1
THC	0.001-0.2	0.999	13	103	0.02	1

*) Linear fit except when labeled 2^{nd}: quadratic fit

Table 51. Validation results of drugs of abuse in oral fluid. Analytes are shown in the order of retention time.

Analyte	Calibration range (mg/L)	Correlation coefficient (R^2) *	Reproducibility (%)	LOD (µg/L)	LOQ (µg/L)
Benzoylecgonine	0.015-3.0	0.993(2nd)	3	0.1	15
Aminoflunitrazepam	0.001-0.2	0.990	ND	0.5	1
Morphine	0.005-1.0	0.986(2nd)	5	0.5	5
MDA	0.01-2.0	0.994	5	0.1	10
Amphetamine	0.01-2.0	0.991(2nd)	11	1.0	10
6-MAM	0.002-0.4	0.982	6	0.1	2
MDMA	0.01-2.0	0.995	5	0.1	10
Codeine	0.02-4.0	0.993	2	0.4	20
Bromazepam	0.002-0.4	0.992	7	0.1	2
Metamphetamine	0.005-1.0	0.993	10	0.1	5
Nitrazepam	0.002-0.4	0.992	4	0.04	2
Clonazepam	0.001-0.2	0.988	6	0.1	0.9
MDEA	0.005-1.0	0.992(2nd)	5	0.3	5
Flunitrazepam	0.001-0.2	0.992	3	0.1	1
Clobazam	0.005-1.0	0.991	5	0.04	5
Oxazepam	0.01-2.0	0.993	4	0.2	10
Lorazepam	0.002-0.4	0.994(2nd)	8	0.2	10
Alprazolam	0.001-0.2	0.992	2	0.2	1
Triazolam	0.001-0.2	0.994	2	0.2	1
Desalkylflurazepam	0.002-0.4	0.991	ND	0.2	2
Temazepam	0.01-2.0	0.993	3	0.2	10
Chlordiazepoxide	0.01-2.0	0.992	12	0.3	10
Lormetazepam	0.001-0.2	0.991	8	0.1	1
Desmethyldiazepam	0.005-1.0	0.992	6	0.3	5
Cocaine	0.005-1.0	0.993	5	0.1	5
Diazepam	0.005-1.0	0.993	13	0.5	5
Midazolam	0.002-0.4	0.986	14	0.05	2
9-COOH-THC	0.001-0.2	0.901(2nd)	ND	2.3	2
Flurazepam	0.001-0.2	0.991	10	0.03	1
THC	0.001-0.2	0.924	43	0.8	1

*) Linear fit except when labeled 2nd: quadratic fit.

Annex 8

Sensitivity, specificity and accuracy results of the devices per individual substance test

[1] Tested in Belgium

[2] Tested in Belgium and the Netherlands

[3] Tested in Belgium, Finland and the Netherlands

[4] Tested in Finland

[5] Tested in the Netherlands

n.a. Calculation not applicable.

Table 52. Amphetamine results.

	OraLab6[1]	DrugTest 5000[2]	Cozart DDS[1]	Rapid STAT[3]	OrAlert[1]	DrugWipe 5+[4]	Oratect III[5]	BIOSENS[5]	Average
TP	19	6	4	22	1	33	0	0	85
FP	0	2	1	8	10	5	0	9	35
TN	216	210	131	288	97	92	56	99	1189
FN	14	3	2	19	2	5	2	10	57
total	249	221	138	337	110	135	58	118	1366
failed	1	2	0	11	15	0	0	0	29
missing	0	0	0	1	0	1	0	0	2
sensitivity	58%	67%	67%	54%	n.a.	87%	n.a.	0%	60%
specificity	100%	99%	99%	97%	91%	95%	100%	92%	97%
accuracy	94%	98%	98%	92%	89%	93%	97%	84%	93%

Table 53. Methamphetamine results.

	OraLab6[1]	DrugTest 5000[2]	Rapid STAT[3]	OrAlert[1]	Oratect III[5]	Average
TP	0	0	4	0	0	4
FP	0	0	6	0	1	7
TN	249	220	325	110	57	961
FN	0	0	0	0	0	0
total	249	220	335	110	58	972
failed	1	3	8	15	0	27
missing	0	0	1	0	0	1
sensitivity	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
specificity	100%	100%	98%	100%	98%	99%
accuracy	100%	100%	98%	100%	98%	99%

Table 54. MDMA results.

	Cozart DDS[1]
TP	0
FP	0
TN	138
FN	0
total	138
failed	0
missing	0
sensitivity	n.a.
specificity	100%
accuracy	100%

Table 55. Cannabis results.

	OraLab6[1]	DrugTest 5000[2] *	Cozart DDS[1]	Rapid STAT[3]	OrAlert[1]	DrugWipe 5+[4]	Oratect III[5]	BIOSEN[5] **	Average
TP	18	47	11	67	3	9	16	18	189
FP	2	6	0	23	0	4	0	1	36
TN	135	134	87	198	83	109	8	2	756
FN	94	33	40	53	24	12	34	18	308
total	249	220	138	341	110	134	58	39	1289
failed	1	2	0	8	15	1	0	0	27
missing	0	0	0	0	0	1	0	0	1
sensitivity	16%	59%	22%	56%	11%	43%	32%	50%	38%
specificity	99%	96%	100%	90%	100%	96%	100%	n.a.	95%
accuracy	61%	82%	71%	78%	78%	88%	41%	51%	73%

* results for devices with both cut-offs consolidated, ** only the second successful analysis using BIOSENS is included.

Table 56. Cocaine results.

	OraLab6[1]	DrugTest 5000[2]	Cozart DDS[1]	Rapid STAT[3]	OrAlert[1]	DrugWipe 5+[4]	Oratect III[5]	BIOSENS[5]	Average
TP	19	9	1	8	7	0	0	0	44
FP	0	1	1	3	0	0	0	0	5
TN	195	202	129	305	96	123	55	101	1206
FN	35	9	7	14	7	0	3	3	78
total	249	221	138	330	110	123	58	104	1333
failed	1	2	0	10	15	1	0	14	43
missing	0	0	0	9	0	12	0	0	21
sensitivity	35%	50%	13%	36%	50%	n.a.	n.a.	n.a.	36 %
specificity	100%	100%	99%	99%	100%	100%	100%	100%	100%
accuracy	86%	95%	94%	95%	94%	100%	95%	97%	94%

Table 57. Opiate results.

	OraLab6[1]	DrugTest 5000[2]	Cozart DDS[1]	Rapid STAT[3]	OrAlert[1]	DrugWipe 5+[4]	Oratect III[5]	BIOSENS[5]	Average
TP	84	70	45	63	61	0	0	2	325
FP	3	9	4	7	5	0	0	3	31
TN	125	133	80	262	21	134	57	111	923
FN	37	9	9	7	23	1	1	2	89
total	249	221	138	339	110	135	58	118	1368
failed	1	2	0	5	15	1	0	0	24
missing	0	0	0	0	0	0	0	0	0
sensitivity	69%	89%	83%	90%	73%	n.a.	n.a.	n.a.	79%
specificity	98%	94%	95%	97%	81%	100%	100%	97%	97%
accuracy	84%	92%	91%	96%	75%	99%	98%	96%	91%

Table 58. Benzodiazepine results.

	DrugTest 5000[2]	Cozart DDS[1]	Rapid STAT[3]	Oratect III[5]	BIOSENS[5]	Average
TP	34	25	64	0	0	123
FP	0	5	3	0	0	8
TN	169	81	242	58	118	669
FN	18	27	31	0	0	75
total	221	138	340	58	118	875
failed	2	0	9	0	0	11
missing	0	0	0	0	0	0
sensitivity	65%	48%	67%	n.a.	n.a.	62%
specificity	100%	94%	99%	100%	100%	99%
accuracy	92%	77%	90%	100%	100%	91%

Table 59. PCP results.

	OraLab6[1]	OrAlert [1]	Average
TP	0	0	0
FP	5	0	5
TN	244	110	354
FN	0	0	0
total	249	110	359
failed	1	15	16
missing	0	0	0
sensitivity	n.a.	n.a.	n.a.
specificity	98%	100%	99%
accuracy	98%	100%	99%

Annex 9

Overall evaluation: sensitivity, specificity and accuracy of the device for detecting drug positive cases

[1] Tested in Belgium

[2] Tested in Belgium and the Netherlands

[3] Tested in Belgium, Finland and the Netherlands

[4] Tested in Finland

[5] Tested in the Netherlands

Table 60. Overall evaluation of on-site devices

	OraLab6[1]	DrugTest 5000[2]	Cozart DDS[1]	Rapid STAT[3]	OrAlert [1]	DrugWipe 5+ [4]	Oratect III[5]	BIOSENS [5]
TP	72	66	28	119	21	87	8	2
FP	3	11	2	17	9	5	0	0
TN	63	22	39	37	15	4	34	11
FN	111	121	69	169	65	39	16	12
Total	249	220	138	342	110	135	58	25
Prevalence	69%	65%	78%	60%	73%	32%	86%	92%

Table 61. Sensitivity of overall evaluation, with 95% standard error limits

	OraLab6[1]	DrugTest 5000[2]	Cozart DDS[1]	Rapid STAT[3]	OrAlert [1]	DrugWipe 5+ [4]	Oratect III[5]	BIOSENS [5]
Sensitivity	64%	85%	64%	82%	81%	91%	32%	52%
Upper limit	71%	90%	72%	87%	88%	97%	46%	71%
Lower limit	56%	78%	55%	76%	71%	78%	21%	33%

Table 62. Specificity of overall evaluation, with 95% standard error limits

	OraLab6[1]	DrugTest 5000[2]	Cozart DDS[1]	Rapid STAT[3]	OrAlert [1]	DrugWipe 5+ [4]	Oratect III[5]	BIOSENS [5]
Specificity	96%	86%	93%	88%	70%	96%	100%	100%
Upper limit	99%	92%	99%	92%	83%	99%	-	-
Lower limit	88%	76%	77%	81%	52%	88%	62%	29%

Table 63. Accuracy of overall evaluation, with 95% standard error limits

	OraLab6[1]	DrugTest 5000[2]	Cozart DDS[1]	Rapid STAT[3]	OrAlert [1]	DrugWipe 5+ [4]	Oratect III[5]	BIOSENS [5]
Accuracy	74%	85%	70%	84%	78%	93%	41%	56%
Upper limit	79%	89%	77%	88%	85%	97%	54%	73%
Lower limit	68%	80%	62%	80%	70%	88%	30%	37%