



Review article

Roadside screening tests for cannabis use: A systematic review



Erica Wennberg^{a,b}, Sarah B. Windle^{a,c}, Kristian B. Filion^{a,c,d},
 Brett D. Thombs^{a,c,d,e,f}, Genevieve Gore^g, Andrea Benedetti^{c,d,h}, Roland Grad^{a,i},
 Carolyn Ells^{a,d,f,i,j}, Mark J. Eisenberg^{a,c,d,k,*}

^a Lady Davis Institute for Medical Research, Jewish General Hospital/McGill University, Montreal, QC, Canada

^b Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

^c Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, QC, Canada

^d Department of Medicine, McGill University, Montreal, QC, Canada

^e Departments of Psychiatry, Psychology, and Counselling and Educational Psychology, McGill University, Montreal, QC, Canada

^f Biomedical Ethics Unit, McGill University, Montreal, QC, Canada

^g Schulich Library of Science and Engineering, McGill University, Montreal, Quebec, Canada

^h Respiratory Epidemiology and Clinical Research Unit, Centre for Outcomes Research and Evaluation, Research Institute of the McGill University Health Centre, Montreal, QC, Canada

ⁱ Department of Family Medicine, McGill University, Montreal, QC, Canada

^j Division of Experimental Medicine and Department of Social Studies of Medicine, McGill University, Montreal, QC, Canada

^k Division of Cardiology, Jewish General Hospital/McGill University, Montreal, QC, Canada

ARTICLE INFO

Keywords:

Cannabis
 Impaired driving
 DUI
 Systematic review

ABSTRACT

As more countries legalize recreational cannabis, roadside screening programs are imperative to detect and deter driving under the influence of cannabis. This systematic review evaluated roadside screening tests for cannabis use. We searched six databases (inception-March 2020) and grey literature sources for primary studies evaluating test characteristics of roadside screening tests for cannabis use compared to laboratory tests for cannabinoids in blood or oral fluid. The synthesis was focused on sensitivity and specificity of delta-9-tetrahydrocannabinol (THC) detection. 101 studies were included. Oral fluid tests were higher in specificity and lower in sensitivity compared to urine tests when evaluated against blood laboratory tests. Oral fluid tests were higher in sensitivity and similar in specificity compared to observational tests when evaluated against blood and oral fluid laboratory tests. Sensitivity was variable among oral fluid tests; two instrumented immunoassays (Draeger DrugTest 5000 [5 ng/mL THC cut-off] and Alere DDS 2 Mobile Test System) appeared to perform best, but definitive conclusions could not be drawn due to imprecise estimates. Specificities were similar. Overall, oral fluid tests showed the most promise for use in roadside screening for blood THC levels over legal limits; their continued development and testing are warranted. Urine tests are generally inadvisable, and observational tests require sensitivity improvements.

* Corresponding author. Divisions of Cardiology and Clinical Epidemiology, Jewish General Hospital/McGill University, 3755 Côte Ste-Catherine Road, Suite H421, Montreal, Quebec, H3T 1E2, Canada.

E-mail address: mark.eisenberg@mcgill.ca (M.J. Eisenberg).

<https://doi.org/10.1016/j.heliyon.2023.e14630>

Received 4 August 2022; Received in revised form 20 December 2022; Accepted 13 March 2023

Available online 22 March 2023

2405-8440/© 2023 Published by Elsevier Ltd.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Driving under the influence of cannabis (DUIC) is an established public health risk. Studies have demonstrated a link between DUIC and reduced driving ability, as well as an increased risk of motor vehicle crashes [1–4]. In countries that have legalized (or are on the path to legalizing) recreational cannabis, talk of legalization has been met with concern regarding its potential influence on cannabis-impaired driving. Evidence on the effect of legalization on DUIC is mixed at present [5–9]. However, the potential for post-legalization reductions in road safety remains a cause for concern and has led to government interest in pairing legalization with DUIC prevention efforts. Roadside detection by law enforcement (e.g., at sobriety checkpoints) has been effective in deterring alcohol-impaired driving [10,11], and the same strategy could be effective in preventing DUIC [12]. However, many questions remain concerning the performance of roadside screening tests for DUIC (e.g., field sobriety tests, oral fluid tests). To our knowledge, only two systematic reviews are available, both of which were limited to oral fluid screening tests. One included only 15 studies [13], and the other limited the matrix used for the laboratory reference test to oral fluid [14]. To address this gap in the literature, we performed a systematic review to assess test characteristics of roadside screening tests for cannabis use compared to laboratory-based confirmatory tests in blood and oral fluid. Laboratory tests for delta-9-tetrahydrocannabinol (THC) in blood (and to a lesser extent in oral fluid) are the most frequently used confirmatory tests in legal practice internationally. Presence of THC in blood and in oral fluid are well correlated, though THC levels in blood and oral fluid are not [15]. Laboratory testing for blood THC is currently viewed as the standard for determination of cannabis impairment, though its use is imperfect; positive results are indicative of prior cannabis use, but THC levels in blood (and in other matrices) are not reliable indicators of impairment [16,17]. We additionally synthesized data on the feasibility, acceptability, and cost-effectiveness of identified roadside screening tests as available.

2. Methods

Our review was conducted based on a protocol developed prior to initiating the review (<https://osf.io/q57y4/?show=view>) and is reported following the Preferred Reporting Items for Systematic reviews and Meta-Analyses of Diagnostic Test Accuracy (PRISMA-DTA) guideline [18].

2.1. Inclusion and exclusion criteria

Articles and abstracts were included if they were published in English or French and reported primary data on the test characteristics of one or more roadside screening tests for DUIC. We did not restrict by population. Eligible tests included any biochemical and observational measures and had to be designed for on-site use (e.g., at roadside or police stations); laboratory-based tests were excluded. Eligible laboratory reference tests included laboratory-based tests in blood or oral fluid to reflect the biological matrices currently used for confirmation of DUIC charges in countries with DUIC laws in place. Eligible publications reported at least one diagnostic test characteristic (e.g., sensitivity, specificity, positive predictive value) specific to cannabis detection. All diagnostic accuracy study designs were eligible.

2.2. Data sources

We searched the Embase, PsycINFO, Web of Science, TRID and Criminal Justice Database (ProQuest) databases from inception to March 31, 2020, as well as the MEDLINE database from inception to March 20, 2020. We also searched the websites of governmental (e.g., the National Highway Traffic Safety Administration) and non-governmental (e.g., the International Transport Forum) traffic safety and drug policy organizations selected as relevant by EW. Additional articles were identified by scanning the references of included publications.

2.3. Search strategy and Study Selection

The search was designed by a health sciences librarian (G.G.) and peer-reviewed using Peer Review of Electronic Search Strategies [19]. It included a combination of cannabis terms, drug testing terms, driving terms, and diagnostic test characteristic terms. The complete search strategies are presented in Appendix A. Duplicates were removed in EndNote X9. Screening was performed using Covidence (Veritas Health Innovation, Melbourne, Australia). Two reviewers independently reviewed the titles and abstracts of all identified publications in random order. If either reviewer deemed a publication potentially eligible, the publication was reviewed in full text by both reviewers independently. Disagreements after full-text review were resolved by consensus.

2.4. Data extraction

Data were extracted independently by two individuals using piloted forms in Covidence (piloting was conducted on five studies by two reviewers). Any discrepancies were resolved via consensus. Variables included: sample characteristics; sample size; study design; cannabis consumption method and dose/timing (experimental studies); co-consumption of alcohol or other substances; screening tests evaluated; laboratory reference tests and positivity thresholds; and funding sources. Studies were categorized by the type of screening test they examined and their design. Screening tests were divided into three major categories: oral fluid screening tests, urine screening tests, and observational tests. Study design was considered at two levels. The first level considered whether or not cannabis was

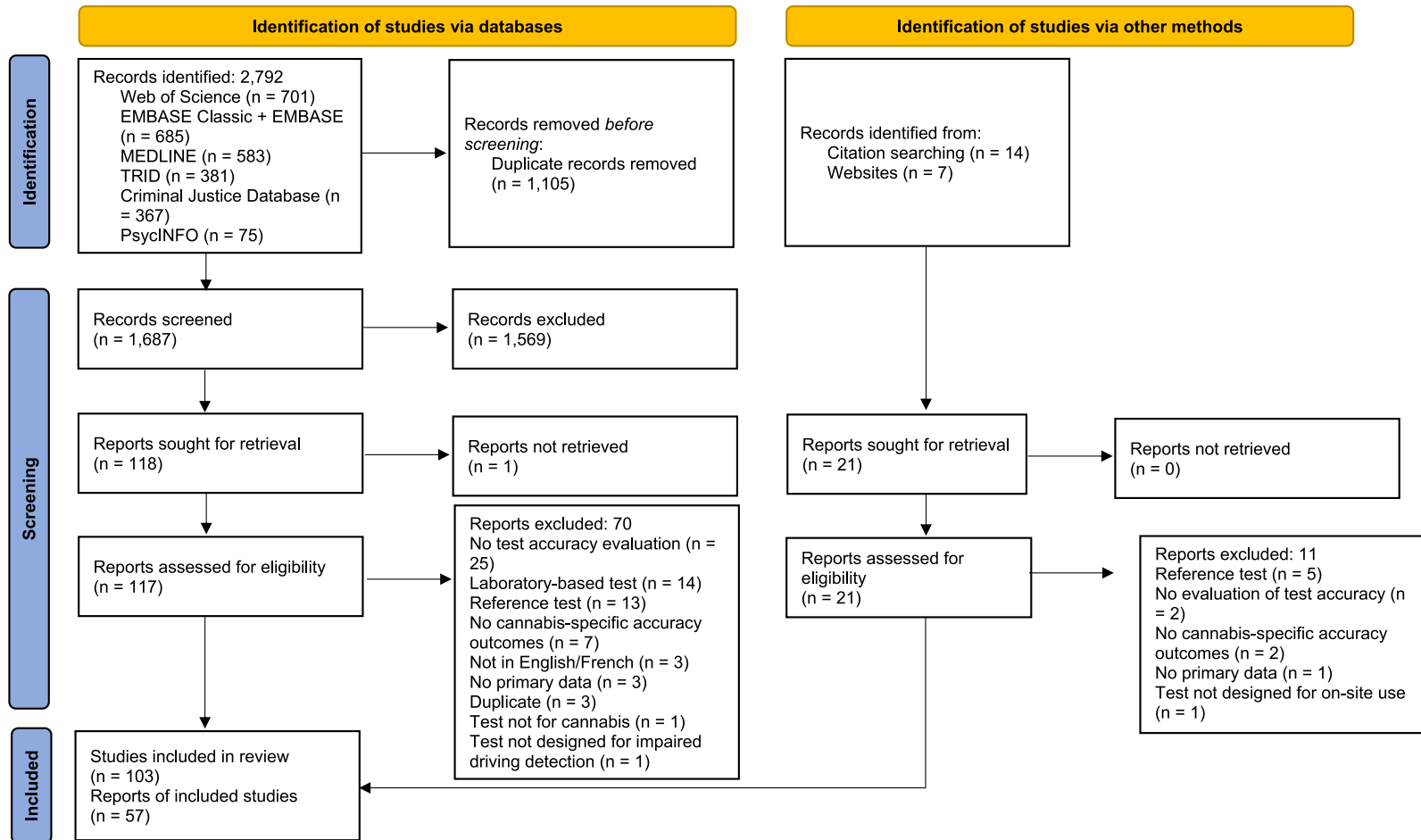


Fig. 1. PRISMA flow of study selection.

administered to participants (experimental vs. observational studies). Experimental studies fell into three major categories: between subjects (distinct participant groups were administered and not administered cannabis), within subjects pre-post (all participants were administered active cannabis doses, with screening test[s] performed pre- and post-administration), within subjects post only (all participants were administered active cannabis doses, with screening test[s] performed post-administration), and within subjects placebo-controlled crossover (participants were administered placebo and one or more active cannabis dose[s] over several sessions). The second level considered the diagnostic test accuracy study design (single test vs. comparative [non-randomized or randomized] and diagnostic cross-sectional vs. diagnostic case-control). Reported test accuracy characteristics (sensitivity, specificity, etc.) were extracted for all tests assessed. Confidence intervals (CIs) for test accuracy characteristics, as well as number of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN), were extracted as available. TP was defined as cannabis presence on the screening test and the laboratory reference test. TN was defined as cannabis absence on the screening test and the laboratory reference test. FP was defined as cannabis presence on the screening test but cannabis absence on the laboratory reference test. FN was defined as cannabis absence on the screening test but cannabis presence on the laboratory reference test. If test sensitivity or specificity were not provided, they were calculated using available information, as were 95% CIs [20]. 95% CIs were used as a marker of the precision of individual sensitivity and specificity estimates. Data on test feasibility, acceptability, and cost-effectiveness were extracted as secondary outcomes as available, as were test characteristics in frequent vs. occasional cannabis users and with cannabis use alone vs. polydrug use (cannabis and other drugs).

2.5. Risk of bias assessment

The risk of bias of included studies was assessed independently by two reviewers using the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) [21] tool, with disagreements resolved by consensus. The QUADAS-2 tool assesses risk of bias in primary diagnostic accuracy studies in four domains: patient selection (here, participant selection), index test (here, screening test), reference standard (here, laboratory reference test), and flow and timing. It also includes an assessment of concerns regarding applicability to the research question in the first three of the four aforementioned domains.

2.6. Data synthesis

Meta-analysis of diagnostic test characteristics was not conducted due to insufficient numbers of studies that evaluated the same screening tests using the same laboratory reference tests. Data were instead synthesized narratively. Data from experimental and observational single test studies were synthesized separately. Due to a large amount of available data, the synthesis was focused on test sensitivity and specificity compared to laboratory tests for THC in blood or oral fluid, as THC is the major cannabinoid of interest for determination of recent cannabis use. Other test characteristics and laboratory test cannabinoids were described but were not focused on for the synthesis. Data on test feasibility and acceptability were likewise synthesized narratively.

3. Results

3.1. Search results

Our database searches generated 2792 records (Fig. 1). After duplicate removal and title and abstract screening, 117 reports were screened in full text. Forty-seven met inclusion criteria. Ten additional eligible reports were identified through the grey literature search and hand-searching the reference lists of included publications. Fourteen reports described more than one study. Thirteen reports described the multi-study Rosita [22,23], Rosita-2 [24–28], and DRUID projects [29–34] (hereafter grouped under the project name). Therefore, 103 studies were included in our systematic review, described by 57 reports.

3.2. Study characteristics

Fourteen experimental studies [35–47] (Table B1) and 89 observational studies [22–34,43,48,49] (Table B2) were included. Experimental studies were performed primarily in subjects with a history of cannabis use (occasional, frequent, or both) ($n = 10$). Four used cannabis-fortified oral fluid samples. Experimental studies were primarily within-subjects in design (pre-post: $n = 4$; placebo-controlled crossover: $n = 4$; post only: $n = 2$); a minority were between-subjects ($n = 4$). Observational studies were primarily performed in drivers suspected of or arrested for impaired driving ($n = 47$). Others were performed in drivers ($n = 10$), people with drug addiction ($n = 6$), cannabis coffeeshop customers ($n = 5$), club/bar customers ($n = 1$), hospital patients ($n = 1$), music festival attendees ($n = 1$), prison inmates ($n = 1$), recent cannabis users ($n = 1$), and mixed samples ($n = 16$). Oral fluid screening tests were evaluated in the majority of included observational studies ($n = 78$) and all experimental studies. Most studies evaluating oral fluid screening tests compared them solely to laboratory tests of oral fluid (experimental: $n = 12$; observational: $n = 44$). The remainder used laboratory tests of blood/serum/plasma alone (observational: $n = 21$) or blood/serum and oral fluid (experimental: $n = 2$; observational, $n = 13$). Urine screening tests were evaluated in six observational studies, all of which compared them to laboratory tests of blood/plasma. Observational tests were evaluated in seven observational studies. They were compared to laboratory tests of blood ($n = 3$), blood or urine ($n = 3$), and oral fluid ($n = 1$).

3.3. Roadside screening test characteristics

Thirty-six oral fluid screening tests, six urine screening tests, and two observational tests were investigated (Table B3). The majority of investigated oral fluid screening tests (n = 24), and all investigated urine screening tests, consisted of non-instrumented visually-interpreted test sticks that test for the presence of THC and/or other cannabinoids via immunoassay. Oral fluid screening tests also included instrumented (involving a portable computer) computer-interpreted immunoassay tests (n = 8) and two devices still in development. Investigated observational tests included Drug Recognition Expert (DRE) evaluations and an observational test battery. These tests consist of a series of assessments performed by a trained examiner that cumulate in a decision of non-impairment or impairment in one or more drug classes.

3.4. QUADAS-2 assessment

The QUADAS-2 assessment is summarized in Fig. 2 (see Table B4 for complete assessment). No studies had low risk across all risk of bias domains. Participant selection was rated at high risk of bias for all experimental studies (n = 14) and 67% of observational studies (n = 60/89), largely due to restriction to people with specific cannabis use histories and non-consecutive or non-random sampling methods, respectively. The screening test domain was rated at high risk of bias in 28% of experimental studies and 1% of observational studies, and as unclear in 0% and 15%. The laboratory reference test domain was rated at high risk of bias in 21% of experimental and observational studies and as unclear in 50% and 76%. The large proportions of unclear ratings were due to many studies lacking

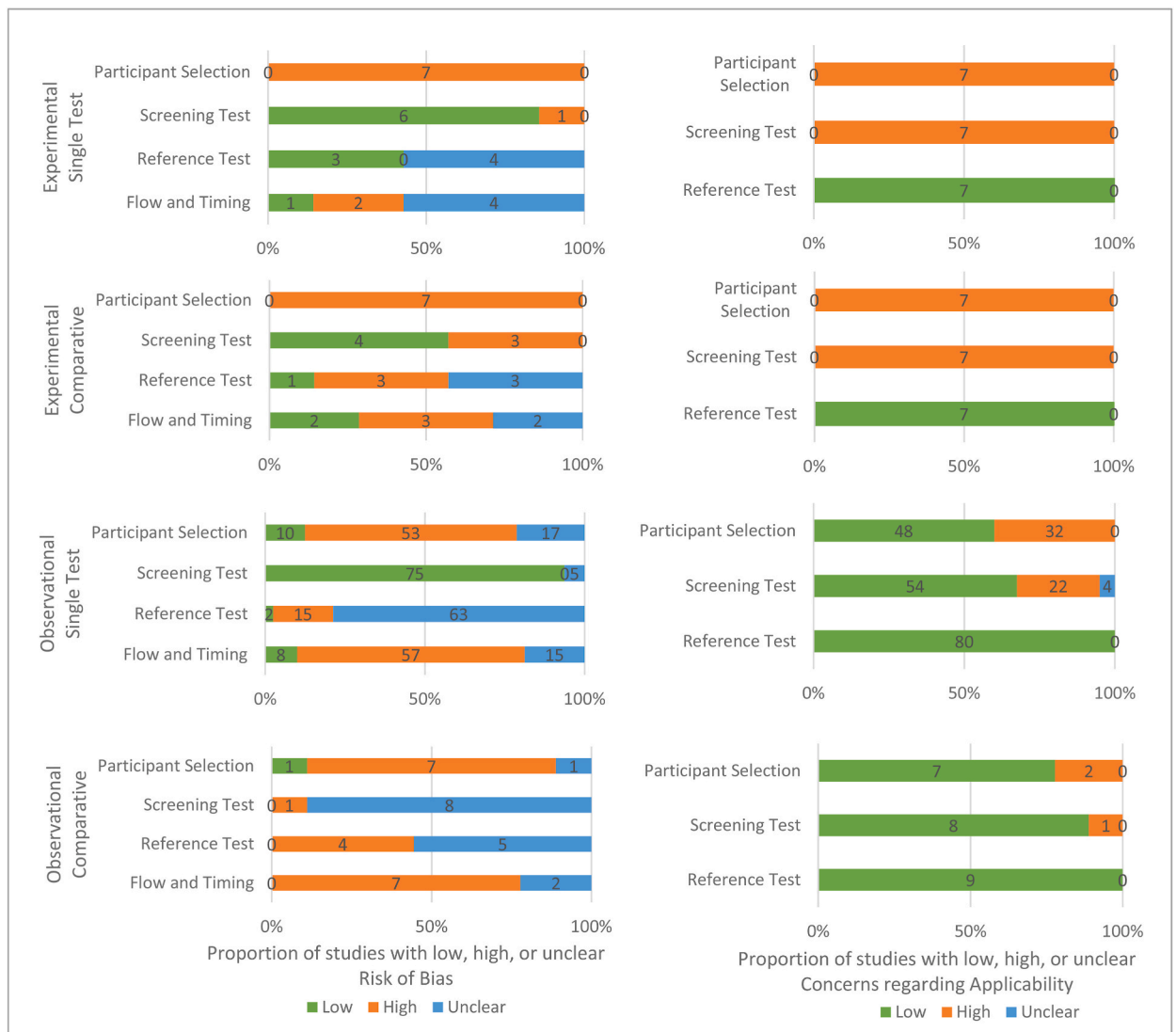


Fig. 2. Summary of QUADAS-2 quality assessment results across all included studies by study design.

reporting of whether laboratory test interpretation was blinded to screening test results. Flow and timing was rated at high risk of bias in 36% of experimental and 72% of observational studies and as unclear in 43% and 19%. The large proportion of unclear ratings was due to many studies lacking reporting of time delays between the screening and laboratory tests.

No studies had low risk across all concerns regarding applicability domains. All experimental studies and 38% of observational studies were rated at high risk in the participant selection domain (0% unclear). All experimental studies and 26% of observational studies were rated at high risk in the screening test domain (4% unclear). All studies were rated at low risk in the laboratory reference test domain.

3.5. Sensitivity and specificity of roadside screening tests for cannabis use

3.5.1. Test sensitivity and specificity overall

Two experimental studies evaluated oral fluid screening tests compared to laboratory tests for THC in blood or serum (summarized for the most common legal limits [1.0–5.0 ng/mL THC] in Table 1; see Table B5 for complete results). Definitive conclusions on the performance of these tests from the two experimental studies were not possible due to imprecise sensitivity and specificity estimates within the studies. One within-subjects experimental comparative study compared the Draeger DrugTest 5000 with 5 ng/mL positivity threshold (DDT5000-5) and the Alere DDS 2 (DDS2) against a laboratory test in blood (Table 2) [40]. Focusing on positivity thresholds between 1.0 and 5.0 ng/mL THC, the DDS2 achieved higher sensitivity and slightly lower specificity than the DDT5000-5, but sensitivity and specificity estimates were imprecise (SENS [95% CI]: 43.6–85.7% [33.1–95.0] vs. 34.7–73.7% [24.8–88.2]; SPEC: 82.1–77.0% [90.0–68.4] vs. 87.1–82.1% [94.9–72.6]). In observational single-test studies (Table B6), three studies of the DDT5000-5 in drivers suspected of driving under the influence of drugs (DUID) reported sensitivities between 82.9 and 87.1% [74.3–91.6] at a positivity threshold of 1 ng/mL THC in blood or serum [43,50,51]; specificity estimates were imprecise for all but one (88.7% [83.6–92.3]). For the DDS2, sensitivity estimates for drivers suspected of or arrested for DUID were similar to the DDT5000-5, but laboratory tests used were heterogeneous, and the estimates were imprecise. The second experimental study evaluated the Securetec DrugWipe 5S against a laboratory test in serum at 1 ng/mL and 5.0 ng/mL THC; sensitivity and specificity estimates were again imprecise (SENS: 50.0% [35.5–64.5] and 80.8% [62.1–91.5]; SPEC: 90.0% [74.4–96.5] and 93.5% [82.5–97.8]). The DrugWipe 5S was not evaluated in any observational studies.

Fourteen experimental studies evaluated oral fluid screening tests compared to laboratory tests for THC in oral fluid (summarized for the most common legal limits [0.0–5.0 ng/mL THC] in Table 3; see Table B5 for complete results). Test sensitivity was highly variable, both for the same screening test (i.e., across studies and laboratory test positivity thresholds) and between screening tests, and was highest at or above the positivity threshold provided by the manufacturer. Test specificity was more consistent. Overall, the DDT5000-5 and DDS2 showed higher sensitivity compared to other oral fluid tests when evaluated against laboratory tests of oral fluid, but imprecise sensitivity and specificity estimates within studies made definitive conclusions impossible. Focusing on positivity thresholds between 0.5 and 5.0 ng/mL THC, one within-subjects experimental comparative study found the DDT5000-10 had higher sensitivity and slightly lower specificity compared to the Securetec DrugWipe 5S (SENS: 39.7–44.0% [31.2–54.2] vs. 25.4–28.0% [18.4–37.8]; SPEC: 97.9–90.3% [99.6–81.3] vs. 100.0–94.4% [100.0–86.6]) (Table 4) [35]. Another within-subjects experimental comparative study found the DDT5000-10 had lower sensitivity and higher specificity compared to the DDT5000-5, but the sensitivity and specificity estimates were imprecise [37]. Two within-subjects experimental comparative studies compared the DDT5000-5 and the DDS2 [40,42]. One showed similar performance between the DDS2 and the DDT5000-5 (SENS: 53.2–84.4% [47.2–89.3] vs 57.5–80.0% [51.4–85.7]; SPEC: 98.2–94.5 [99.2–91.8] vs 98.7–91.9 [99.5–88.8]). The other study favoured the DDS2 in sensitivity, but the sensitivity estimate was imprecise. For comparison of oral fluid screening tests to non-THC cannabinoid laboratory tests, see the Appendix (Tables B7–B9).

Four observational studies evaluated urine screening tests against a laboratory test for THC in blood, all in drivers suspected of DUID (Table B10) [22,51]. One found the sensitivity and specificity of the DrugScreen Multi-5 urine test were 92.9% [89.5–95.2] and 40.0% [24.6–57.7]. Using the same blood positivity threshold (1 ng/mL THC), comparisons with other observational studies in drivers suspected of DUID showed higher sensitivity and substantially lower specificity in the urine test compared to the DDT5000-5 oral fluid test (SENS: absolute differences of 4.1%–10.1%; SPEC: 48.7% to –7.2%). The three remaining studies used a positivity threshold of 0.3 ng/mL THC; test sensitivities and specificities for the three urine tests evaluated ranged from 86.0 to 97.6% [73.8–99.6] and 65.3–79.6% [51.3–88.5] [22]. Comparisons with the sole oral fluid test with sensitivity and specificity results at this blood positivity threshold (Cozart Rapiscan 200 ng/mL, also evaluated in drivers suspected of DUID) [22] again showed lower specificity in the urine screening tests (absolute differences: 22.7% to –8.4%), but much higher sensitivity (57.4%–69.0%).

Observational tests were evaluated in seven observational studies [52,53,54,55,56,57,58] (Table B11). Six examined DRE evaluation. The most precise assessment of DRE evaluation against laboratory tests in blood found a sensitivity of 59.7% [48.6–70.0] and specificity of 86.5% [78.2–91.9] [52]. One study found using presence of two of the following four measures as indicative of test positivity for cannabis maximized DRE evaluation sensitivity and specificity (97.0% [94.4–98.4] and 96.7% [94.0–98.2] at 1 ng/mL THC in blood): (1) ≥ 3 misses, finger to nose test; (2) eyelid tremors, Modified Romberg Balance test; (3) ≥ 2 clues, one leg stand; (4) ≥ 2 clues, walk and turn [55]. An observational comparative study demonstrated significantly lower sensitivity and similar specificity in a battery of observed signs of impairment compared to the DDT5000-25 oral fluid screening test when evaluated against a laboratory test in oral fluid (SENS: 19.3% [13.5–26.7] vs. 76.3% [68.5–82.7]; SPEC: 94.9% [89.4–97.6] vs. 93.2% [87.2–96.5] at 27 ng/mL THC) (Table 4) [54]. Another observational comparative study of DRE evaluation and the DDS2 oral fluid screening test that compared to a laboratory test in blood had similar findings, but sensitivity and specificity estimates were imprecise (Table 2) [53].

Table 1

Sensitivity and specificity of investigated oral fluid screening tests compared to laboratory reference tests of THC in blood or serum at positivity thresholds ranging from 0 to 5 ng/mL THC, as reported by 2 experimental studies of roadside tests for cannabis use.

Screening Test (positivity threshold, ng/mL THC)	Study (sample)	Laboratory Reference Test (n tests; n reference test positives)	Sensitivity and specificity (% [95% CI]) at varying positivity thresholds (ng/mL THC)							
			0–0.5		1.0		2.0		5.0	
			SENS	SPEC	SENS	SPEC	SENS	SPEC	SENS	SPEC
Alere DDS										
Alere DDS 2 Mobile Test System (25)	Newmeyer 2017 (adult occasional and frequent cannabis users)	LC-MS/MS in blood (134; 1.0: 78, 2.0: 41, 5.0: 21)	-	-	43.6 (33.1–54.6)	82.1 (70.2–90.0)	65.9 (50.6–78.4)	81.7 (72.7–88.3)	85.7 (65.4–95.0)	77.0 (68.4–83.8)
Draeger DrugTest										
Draeger DrugTest 5000 (5)	Newmeyer 2017 (adult occasional and frequent cannabis users)	LC-MS/MS in blood (103; 1.0: 72, 2.0: 53, 5.0: 19)	-	-	34.7 (24.8–46.2)	87.1 (71.2–94.9)	43.4 (31.0–56.7)	88.0 (76.2–94.4)	73.7 (51.2–88.2)	82.1 (72.6–88.9)
Securetec DrugWipe										
Securetec DrugWipe 5S (15)	Wille 2015 (adult regular cannabis users)	GC-MS in serum (72; 1.0: 42, 5.0: 26)	-	-	50.0 (35.5–64.5)	90.0 (74.4–96.5)	-	-	80.8 (62.1–91.5)	93.5 (82.5–97.8)

CI = confidence interval; DDS = Drug Detection System; GC-MS = gas chromatography mass spectrometry; LC-MS/MS = liquid chromatography tandem mass spectrometry; THC = delta-9-tetrahydrocannabinol.

Comparative studies are demarcated in bold.

Table 2

Direct comparisons of test sensitivity and specificity compared to laboratory reference tests of THC in blood at varying positivity thresholds (ng/mL THC), as reported by 1 randomized experimental and 1 non-randomized observational comparative study of roadside tests for cannabis use.

Screening Test (positivity threshold, ng/ mL THC)	Study (sample)	Laboratory Reference Test (n tests; n reference test positives)	Sensitivity and specificity (% [95% CI]) at varying positivity thresholds (ng/mL THC)													
			0.2		0.5		1.0		2.0		5.0		10.0		27.0	
			SENS	SPEC	SENS	SPEC	SENS	SPEC	SENS	SPEC	SENS	SPEC	SENS	SPEC	SENS	SPEC
Oral fluid screening tests vs. observational tests																
Alere DDS 2 (25)	Edwards 2017 (subjects arrested for an alleged OWI)	EIA and GCMS- NPD in blood (15; 9),	-	-	-	-	-	-	-	-	-	-	88.9 (56.5–98.0)	66.7 (30.0–90.3)	-	-
DRE evaluation (N/A)			-	-	-	-	-	-	-	-	-	-	55.6 (26.7–81.1)	66.7 (30.0–90.3)	-	-
Oral fluid screening tests vs. oral fluid screening tests																
Alere DDS 2 (25)	Newmeyer 2017 (adult occasional and frequent cannabis users)	LC-MS/MS in blood (DDS2: 134; 1.0: 78, 2.0: 41, 5.0: 21, 10.0: 11)	-	-	-	-	43.6 (33.1–54.6)	82.1 (70.2–90.0)	65.9 (50.6–78.4)	81.7 (72.7–88.3)	85.7 (65.4–95.0)	77.0 (68.4–83.8)	81.8 (52.3–94.9)	71.5 (63.0–78.8)	-	-
Draeger DrugTest 5000 (5)		LC-MS/MS in blood (103; 1.0: 72, 2.0: 53, 5.0: 19, 10.0: 10)	-	-	-	-	34.7 (24.8–46.2)	87.1 (71.2–94.9)	43.4 (31.0–56.7)	88.0 (76.2–94.4)	73.7 (51.2–88.2)	82.1 (72.6–88.9)	70.0 (39.7–89.2)	76.3 (66.8–83.8)	-	-

CI = confidence interval; DDS = Drug Detection System; DRE = Drug Recognition Expert; EIA = enzyme immunoassay; GCMS-NPD = gas chromatography mass spectrometry with nitrogen phosphorus detector; LC-MS/MS = liquid chromatography tandem mass spectrometry; OWI = operating while intoxicated; THC = delta-9-tetrahydrocannabinol.

Experimental studies are demarcated in bold. Crouch 2005, Walsh 2003, and Walsh 2007 were comparative in design but did not evaluate devices against the same reference test positivity thresholds and were thus excluded from this table.

Table 3

Sensitivity and specificity of investigated oral fluid screening tests compared to laboratory reference tests of THC in oral fluid at positivity thresholds ranging from 0 to 5 ng/mL THC, as reported by 9 experimental studies of roadside tests for cannabis use.

Screening Test (positivity threshold, ng/mL THC)	Study (sample)	Laboratory Reference Test (n tests; n reference test positives)	Sensitivity and specificity (% [95% CI]) at varying positivity thresholds (ng/mL THC)							
			0–0.5		1.0		2.0		5.0	
			SENS	SPEC	SENS	SPEC	SENS	SPEC	SENS	SPEC
Alere DDS										
Alere DDS 2 Mobile Test System (25)	Newmeyer 2017 (adult occasional and frequent cannabis users)	LC-MS/MS in OF (134; 0.2: 93, 1.0: 60, 2.0: 60, 5.0: 31)	44.1 (34.4–54.2)	92.7 (80.6–97.5)	61.7 (49.0–72.9)	90.5 (81.7–95.3)	61.7 (49.0–72.9)	90.5 (81.7–95.3)	96.8 (83.8–99.4)	86.4 (78.5–91.7)
	Swortwood 2017 (adult occasional and frequent cannabis users)	LC-MS/MS in OF (545; 0.2: 397, 1.0: 265, 2.0: 212, 5.0: 147)	36.5 (31.9–41.4)	99.3 (96.3–99.9)	53.2 (47.2–59.1)	98.2 (95.9–99.2)	65.1 (58.5–71.2)	97.6 (95.3–98.8)	84.4 (77.6–89.3)	94.5 (91.8–96.3)
Draeger DrugTest										
DrugTest 5000 (10)	Desrosiers 2012 (adult current cannabis smokers)	2D-GC-MS in OF (66; 0.5: 58, 1.0: 57, 2.0: 54)	75.9 (63.5–85.0)	100.0 (67.6–100.0)	77.2 (64.8–86.2)	100.0 (70.1–100.0)	81.5 (69.2–89.6)	100.0 (75.8–100.0)	–	–
	Arnell 2019 (adult infrequent cannabis users)	LC-MS/MS in OF (163; 1.0: 116, 2.0: 91)	–	–	39.7 (31.2–48.8)	97.9 (88.9–99.6)	44.0 (34.2–54.2)	90.3 (81.3–95.2)	–	–
DrugTest 5000 (5)	Toennes 2013 (heavy cannabis users)	GC-MS in OF (282; 269)	–	–	–	–	–	–	94.4 (91.0–96.6)	15.4 (4.3–42.2)
	Newmeyer 2017 (adult occasional and frequent cannabis users)	LC-MS/MS in OF (103; 0.2: 83, 1.0: 56, 2.0: 56, 5.0: 28)	34.9 (25.6–45.7)	100.0 (83.9–100.0)	50.0 (37.3–62.7)	97.9 (88.9–99.6)	50.0 (37.3–62.7)	97.9 (88.9–99.6)	89.3 (72.8–96.3)	94.7 (87.1–97.9)
	Swortwood 2017 (adult occasional and frequent cannabis users)	LC-MS/MS in OF (551; 0.2: 401, 1.0: 252, 2.0: 208, 5.0: 145)	36.9 (32.3–41.7)	99.3 (96.3–99.9)	57.5 (51.4–63.5)	98.7 (96.6–99.5)	66.3 (59.7–72.4)	96.8 (94.4–98.2)	80.0 (72.8–85.7)	91.9 (88.8–94.2)
	Desrosiers 2012 (adult current cannabis smokers)	2D-GC-MS in OF (66; 0.5: 58, 1.0: 57, 2.0: 54)	86.2 (75.1–92.8)	75.0 (40.9–92.8)	87.7 (76.8–93.9)	77.8 (45.3–93.7)	90.7 (80.1–95.6)	75.0 (46.8–91.1)	–	–
	Desrosiers 2014 (adult occasional and frequent cannabis users) (Oral-Eze/ StatSure) ^a	2D-GC-MS in OF (394; 1.0: 303, 2.0: 258/395; 1.0: 294, 2.0: 258)	–	–	66.7 (61.2–71.7)/ 66.7 (61.1–71.8)	98.9 (94.0–99.8)/ 93.1 (86.4–96.6)	75.6 (70.0–80.4)/ 74.0 (68.4–79.0)	94.1 (88.8–97.0)/ 91.2 (85.3–94.9)	–	–
	Hartman 2015 (adult occasional cannabis users) (all sessions/ high-dose THC sessions)	2D-GC-MS in OF (1710; 1.0: 1125, 2.0: 961, 5.0: 719/ 546; 1.0: 456, 2.0: 411, 5.0: 333)	–	–	40.4 (37.5–43.2)/ 48.7 (44.1–53.3)	99.8 (99.0–100.0)/ 100.0 (95.9–100.0)	47.0 (43.9–50.2)/ 53.8 (48.9–58.5)	99.6 (98.8–100.0)/ 99.3 (95.9–99.9)	60.8 (57.2–64.3)/ 64.9 (59.6–69.8)	98.2 (97.2–98.8)/ 97.2 (94.0–98.7)
Securetec DrugWipe										
DrugWipe 5S (15)	Wille 2015 (adult regular cannabis users)	UPLC-MS/MS in OF (79; 60)	–	–	43.3 (31.6–55.9)	100.0 (83.2–100.0)	–	–	–	–

(continued on next page)

Table 3 (continued)

Screening Test (positivity threshold, ng/mL THC)	Study (sample)	Laboratory Reference Test (n tests; n reference test positives)	Sensitivity and specificity (% [95% CI]) at varying positivity thresholds (ng/mL THC)							
			0–0.5		1.0		2.0		5.0	
			SENS	SPEC	SENS	SPEC	SENS	SPEC	SENS	SPEC
DrugWipe 5S (10)	Arkell 2019 (adult infrequent cannabis users)	LC-MS/MS in OF (165; 1.0: 118, 2.0: 93)	–	–	25.4 (18.4–34.0)	100.0 (92.4–100.0)	28.0 (19.8–37.8)	94.4 (86.6–97.8)	–	–
Tests in Development										
Disposable screen- printed carbon electrode (0) ^b	Wanklyn 2016 (cannabis smokers and non-cannabis smokers)	LC-MS/MS in OF (508; 337)	28.2 (23.6–33.2)	98.8 (95.8–99.7)	–	–	–	–	–	–

2D-GC-MS = 2-dimensional gas chromatography mass spectrometry; CI = confidence interval; DDS = Drug Detection System; GC-MS = gas chromatography mass spectrometry; LC-MS/MS = liquid chromatography tandem mass spectrometry; OF = oral fluid; THC = delta-9-tetrahydrocannabinol; UPLC-MS/MS = ultra-performance liquid chromatography tandem mass spectrometry. Comparative studies are demarcated in bold.

^a 2 OF samples for confirmatory analysis were collected per participant per test using two different collection devices (Oral-Eze and StatSure). OF samples collected with each device were analyzed separately and resulting sensitivity and specificity results were reported separately.

^b A cut-off value for the current was calculated using the CA3-CA1 current responses for samples containing 0 ng/mL Δ^9 -THC. The cut-off was defined as the average CA3-CA1 current plus 2 standard deviations. Results show sensor performance at different time points (s) in the CA3-CA1 response (method involves 3 CA steps, CA3-CA1). Results are reported for optimal performance (0.050–0.075).

Table 4

Direct comparisons of test sensitivity and specificity compared to laboratory reference tests of THC in oral fluid at varying positivity thresholds (ng/mL THC), as reported by 4 experimental and 8 observational randomized and non-randomized comparative studies of roadside tests for cannabis use.

Screening Test (positivity threshold, ng/mL THC)	Study (sample)	Laboratory Reference Test (n tests; n reference test positives)	Sensitivity and specificity (% [95% CI]) at varying positivity thresholds (ng/mL THC)															
			0.2		0.5		1.0		2.0		5.0		10.0		27.0			
			SENS	SPEC	SENS	SPEC	SENS	SPEC	SENS	SPEC	SENS	SPEC	SENS	SPEC	SENS	SPEC		
Oral fluid screening tests vs. observational tests																		
Draeger DrugTest 5000 (25)	Fierro 2014 ^a (drivers positive for THC at a cut-off of 1 ng/mL THC in OF)	UPLC-MS/MS in OF (253; 135)	-	-	-	-	-	-	-	-	-	-	-	-	-	76.3 (68.5–82.7)	93.2 (87.2–96.5)	
Observed signs of impairment																19.3 (13.5–26.7)	94.9 (89.4–97.6)	
Oral fluid screening tests vs. oral fluid screening tests																		
Draeger DrugTest 5000 (10)	Arkell 2019 (adult infrequent cannabis users)	LC-MS/MS in OF (163; 1.0: 116, 2.0: 91, 10.0: 45)	-	-	-	-	39.7 (31.2–48.8)	97.9 (88.9–99.6)	44.0 (34.2–54.2)	90.3 (81.3–95.2)	-	-	66.7 (52.1–78.6)	85.6 (78.1–90.8)	-	-		
Securetec DrugWipe 5S (10)		LC-MS/MS in OF (165; 1.0: 118, 2.0: 93, 10.0: 47)	-	-	-	25.4 (18.4–34.0)	100.0 (92.4–100.0)	28.0 (19.8–37.8)	94.4 (86.6–97.8)	-	-	44.7 (31.4–58.8)	92.4 (86.1–95.9)	-	-			
Draeger DrugTest 5000 (10)	Desrosiers 2012 (adult current cannabis smokers)	2D-GC-MS in OF (66; 0.5: 58, 1.0: 57, 2.0: 54)	-	-	75.9 (63.5–85.0)	100.0 (67.6–100.0)	77.2 (64.8–86.2)	100.0 (70.1–100.0)	81.5 (69.2–89.6)	100.0 (75.8–100.0)	-	-	-	-	-	-		
Draeger DrugTest 5000 (5)			-	-	86.2 (75.1–92.8)	75.0 (40.9–92.8)	87.7 (76.8–93.9)	77.8 (45.3–93.7)	90.7 (80.1–95.6)	75.0 (46.8–91.1)	-	-	-	-	-	-		
Cozart RapiScan] (50)	Drummer 2007 (drivers below alcohol limit and positive on screening for methamphetamine or THC)	GC-MS-SIM in OF (101; 73)	-	-	-	-	-	-	18.3 (11.4–28.0)	NC	-	-	-	-	-	-		
Securetec DrugWipe II (30)		GC-MS-SIM in OF/blood (85; 82)	-	-	-	-	-	-	57.5 (46.1–68.2)	NC	-	-	-	-	-	-		
Affiniton DrugWipe 5 (20)	Logan 2014 (subjects under arrest for DUI)	GC-MS in OF (90; 23)	-	-	-	-	-	-	43.5 (25.6–63.2)	92.5 (83.7–96.8)	-	-	-	-	-	-		
Draeger DrugTest 5000 (5)		GC-MS in OF (91; 24)	-	-	-	-	-	-	58.3 (38.8–75.5)	98.5 (92.0–99.7)	-	-	-	-	-	-		
Alere DDS 2 (25)	Newmeyer 2017 (adult occasional and frequent cannabis users)	LC-MS/MS in OF (134; 0.2: 93, 1.0: 60, 2.0: 60, 5.0: 31)	44.1 (34.4–54.2)	92.7 (80.6–97.5)	-	-	61.7 (49.0–72.9)	90.5 (81.7–95.3)	61.7 (49.0–72.9)	90.5 (81.7–95.3)	96.8 (83.8–99.4)	86.4 (78.5–91.7)	-	-	-	-		
Draeger DrugTest 5000 (5)		LC-MS/MS in OF (103; 0.2: 83, 1.0: 56, 2.0: 56, 5.0: 28)	34.9 (25.6–45.7)	100.0 (83.9–100.0)	-	-	50.0 (37.3–62.7)	97.9 (88.9–99.6)	50.0 (37.3–62.7)	97.9 (88.9–99.6)	89.3 (72.8–96.3)	94.7 (87.1–97.9)	-	-	-	-		
Concateno DDS (31)	Strano-Rossi 2012 (Group 1) (drivers stopped during roadside patrols)	UHPLC-MS/MS + GC-MS ^b in OF (NR)	-	-	-	-	89 (68–109)	NR	-	-	-	-	-	-	-	-		
Draeger DrugTest 5000 (5)			-	-	-	-	95 (86–104)	NR	-	-	-	-	-	-	-	-		

(continued on next page)

Table 4 (continued)

Screening Test (positivity threshold, ng/mL THC)	Study (sample)	Laboratory Reference Test (n tests; n reference test positives)	Sensitivity and specificity (% [95% CI]) at varying positivity thresholds (ng/mL THC)													
			0.2		0.5		1.0		2.0		5.0		10.0		27.0	
			SENS	SPEC	SENS	SPEC	SENS	SPEC	SENS	SPEC	SENS	SPEC	SENS	SPEC	SENS	SPEC
Draeger DrugTest 5000 (5)	Strano-Rossi 2012 (Group 2) (drivers stopped during roadside patrols)	UHPLC-MS/ MS + GC- MS ^b in OF (NR)	–	–	–	–	80 (67–92)	NR	–	–	–	–	–	–	–	–
Mavand RapidSTAT (15)	Strano-Rossi 2012 (Group 3) (drivers stopped during roadside patrols)	UHPLC-MS/ MS + GC- MS ^b in OF (NR)	–	–	–	–	85 (73–97)	NR	–	–	–	–	–	–	–	–
Mavand RapidSTAT (15)	Strano-Rossi 2012 (Group 4) (drivers stopped during roadside patrols)	UHPLC-MS/ MS + GC- MS ^b in OF (NR)	–	–	–	–	33 (7–60)	NR	–	–	–	–	–	–	–	–
Securetec DrugWipe 5+ (30)	Strano-Rossi 2012 (Group 4) (drivers stopped during roadside patrols)	UHPLC-MS/ MS + GC- MS ^b in OF (NR)	–	–	–	–	31 (6–56)	NR	–	–	–	–	–	–	–	–
Concateno DDS (31)	Strano-Rossi 2012 (Group 4) (drivers stopped during roadside patrols)	UHPLC-MS/ MS + GC- MS ^b in OF (NR)	–	–	–	–	31 (6–56)	NR	–	–	–	–	–	–	–	–
Securetec DrugWipe 5+ (30)	Strano-Rossi 2012 (Group 4) (drivers stopped during roadside patrols)	UHPLC-MS/ MS + GC- MS ^b in OF (NR)	–	–	–	–	31 (6–56)	NR	–	–	–	–	–	–	–	–
Alere DDS 2 Mobile Test System (25)	Swortwood 2017 (adult occasional and frequent cannabis users)	LC-MS/MS in OF (545; 0.2: 397, 1.0: 265, 2.0: 212, 5.0: 147)	36.5 (31.9–41.4)	99.3 (96.3–99.9)	–	–	53.2 (47.2–59.1)	98.2 (95.9–99.2)	65.1 (58.5–71.2)	97.6 (95.3–98.8)	84.4 (77.6–89.3)	94.5 (91.8–96.3)	–	–	–	–
Draeger DrugTest 5000 (5)	Strano-Rossi 2012 (Group 4) (drivers stopped during roadside patrols)	UHPLC-MS/ MS + GC- MS ^b in OF (NR)	–	–	–	–	31 (6–56)	NR	–	–	–	–	–	–	–	–
Alere DDS 2 Mobile Test System (25)	Swortwood 2017 (adult occasional and frequent cannabis users)	LC-MS/MS in OF (551; 0.2: 401, 1.0: 252, 2.0: 208, 5.0: 145)	36.9 (32.3–41.7)	99.3 (96.3–99.9)	–	–	57.5 (51.4–63.5)	98.7 (96.6–99.5)	66.3 (59.7–72.4)	96.8 (94.4–98.2)	80.0 (72.8–85.7)	91.9 (88.8–94.2)	–	–	–	–
Safecare Ora- Check (50)	Tang 2018 (adult substance abuse clinic and rehabilitation centre patients)	LC-MS in OF (285; 20)	–	–	–	–	0.0 (0.0–16.1)	100.0 (98.6–100.0)	–	–	–	–	–	–	–	–
Securetec DrugWipe 6S (20)	Tang 2018 (adult substance abuse clinic and rehabilitation centre patients)	LC-MS in OF (429; 32)	–	–	–	–	21.9 (11.0–38.8)	100.0 (99.0–100.0)	–	–	–	–	–	–	–	–
Ulti-Med SalivaScreen (50)	Strano-Rossi 2012 (Group 4) (drivers stopped during roadside patrols)	UHPLC-MS/ MS + GC- MS ^b in OF (NR)	–	–	–	–	0.0 (0.0–12.1)	100.0 (99.0–100.0)	–	–	–	–	–	–	–	–

2D-GC-MS = 2-dimensional gas chromatography mass spectrometry; CI = confidence interval; DDS = Drug Detection System; DUI = driving under the influence; DRE = Drug Recognition Expert; GC-MS = gas chromatography mass spectrometry; GC-MS-SIM = gas chromatography mass spectrometry with selected ion monitoring; LC-MS/MS = liquid chromatography tandem mass spectrometry; OF = oral fluid; THC = delta-9-tetrahydrocannabinol; UHPLC-MS/MS = ultra high-performance liquid chromatography mass spectrometry.

Experimental studies are demarcated in bold. Crouch 2005, Walsh 2003, and Walsh 2007 were comparative in design but did not evaluate devices against the same reference test positivity thresholds and were thus excluded from this table.

^a Results are for the THC alone group.

^b GC-MS (1 ng/mL positivity threshold) was used for THC-positive screening test results unconfirmed by UHPLC-MS/MS (2 ng/mL THC positivity threshold).

3.5.2. Test sensitivity and specificity over time

Four within-subjects experimental studies reported sensitivity and specificity of oral fluid screening tests over time following cannabis consumption compared to laboratory tests in oral fluid. Sensitivity decreased with time, while specificity was less impacted. Focusing on the common positivity threshold between these studies (2 ng/mL THC), one study found the sensitivities of the DDT5000-10 and Securetec DrugWipe 5S decreased from 70.6% [53.8–83.2] and 56% [39.6–70.5] at 10 m post-cannabis vaporization to 46.4% [29.5–64.2] and 21.4% [10.2–39.5] at 1 h and 16.7% [5.8–39.2] and 0.0% [0.0–17.6] at 2 h post-vaporization [35]. In another study, sensitivities of the DDS2 and DDT5000-5 decreased from 85.0% [70.9–92.9] and 77.1% [61.0–87.9] at timepoints ≤ 5 h post-cannabis consumption to 64.3% [51.2–75.5] and 50.9% [38.1–63.6] at timepoints ≤ 20 h post-consumption [40]. A third study found sensitivity and specificity of the DDT5000-5 decreased from 85.7% [80.0–90.0] at ≤ 6 h post-cannabis smoking to 83.9 [78.3–88.3] at ≤ 8 h post-smoking (results averaged between Oral-Eze and StatSure oral fluid collection devices) [38]. A fourth study with a positivity threshold of 1.0 ng/mL THC in oral fluid found sensitivity of the Securetec DrugWipe 5S decreased from 90.0% [59.6–98.2] 5 m after smoking a second cannabis dose to 50.0% [23.7–76.3] at 1 h 20 m post second dose [47].

3.5.3. Influence of frequency cannabis of use, drug co-consumption, and route of administration on test sensitivity and specificity

Three within-subjects experimental studies compared the performance of the DDS2 and/or DDT5000-5 oral fluid screening tests in frequent and occasional cannabis users (results for THC laboratory tests in Table B12; other cannabinoids Table B13) [38,40,42]. There was a trend towards higher sensitivity for THC in occasional compared to frequent cannabis users, notably when comparing to laboratory tests of blood versus oral fluid; however, estimates were often imprecise.

Three studies investigated the influence of alcohol co-consumption on diagnostic test characteristics (Table B14) [39,43,54]. There was some evidence for an impact on specificity, notably for observational tests. Two within-subjects experimental studies found co-consumption of alcohol did not affect oral fluid test accuracy [39,43], and an observational study found it produced similar sensitivity (0.2% absolute difference) and reduced specificity (THC: 93.2% [87.2–96.5] vs. THC + Ethanol: 80.0% [54.8–93.0]) [54]. This last study found alcohol co-consumption marginally increased the sensitivity (4.2% absolute difference) of an observational test and significantly impeded its specificity (THC: 94.9% [89.4–97.6] vs. THC + E: 66.7% [41.7–84.8]); however, sensitivity and specificity estimates for the THC + E group were imprecise due to small sample size. The significantly reduced specificity observed for the observational test could be explained by dual cannabis and alcohol consumption amplifying signs of impairment, leading to evaluators making decisions of cannabis impairment when THC levels were below the positivity threshold.

The influence of cannabis route of administration on device performance was evaluated in one within-subjects experimental comparative study of the DDS2 and DDT5000-5 (Table B15) [42]. Sensitivity varied by route. For the DDS2, it was highest with smoking and lowest with vaporization (45.0–88.9% [37.0–94.5] vs. 27.6–79.1% [20.5–88.6]). For the DDT5000, it was highest with oral consumption and lowest with vaporization (39.4–93.0% [31.6–97.6] vs. 31.4–72.1% [23.8–83.2]). Specificity values were similar across routes (maximum absolute differences: 8.7% for DDS2; 4.1% for DDT5000-5).

3.6. Feasibility and acceptability of investigated cannabis use detection methods

Seventy-one studies reported on test feasibility and acceptability (Table B16), all regarding oral fluid and urine screening tests. Sixty-one reported test failure proportions (oral fluid and urine screening tests commonly have a field indicating invalid results). All tests except seven (Avitar OralScreen, Branan Oratect, Innovacon OrAlert, Securetec DrugWipe 5+, Sun Biomedical Laboratories Oraline, Ultimed SalivaScreen, and Varian OraLab) had cumulative test failure proportions below the 10% benchmark recommended by the Rosita-2 project [24]. Test failure proportions for later versions of the Oratect, Oraline, and DrugWipe fell below 10%. Roadside screening tests were overall deemed feasible and acceptable by evaluators (police officers and research staff); tests with lower positivity thresholds, greater ease of use and interpretation, rapider results, less invalid results, and higher accuracy were preferred. No data were available on cost-effectiveness.

4. Discussion

This systematic review synthesized the sensitivity and specificity of available roadside screening tests for cannabis use compared to oral fluid and blood laboratory tests for THC. The results of our review showed that oral fluid screening tests were higher in specificity and lower in sensitivity compared to urine screening tests in evaluations against blood laboratory tests. Oral fluid tests had higher sensitivity and similar specificity for THC compared to observational tests in evaluations against both blood and oral fluid laboratory tests. Sensitivity was highly variable among oral fluid tests, while specificity was more consistent. Overall, the DDS2 and DDT5000-5 appeared to achieve the highest sensitivities among oral fluid tests when compared to oral fluid laboratory tests, but frequent imprecision of sensitivity and specificity estimates prevented definitive conclusions. Based on these results, oral fluid screening tests show the most promise for use in screening for blood THC levels above legal limits. Urine tests are generally inadvisable, and observational field tests require improvements.

The performance of oral fluid screening tests has improved over the past decade. In the 2000s, roadside screening tests for DUIC were examined as part of a series of international projects on impaired driving. The last of these (DRUID, 2007–2009), which included three European countries, found oral fluid screening tests inadvisable for DUIC detection [59]. Although all eight tests evaluated reached the DRUID specificity target of 80% when compared to laboratory tests in oral fluid, none reached the 80% sensitivity target. In contrast, several oral fluid screening tests evaluated in more recent studies included in our review had sensitivities exceeding 80% at the DRUID cut-off (1 ng/mL THC in oral fluid). In addition, while the DRUID project found that the sensitivity of oral fluid screening

tests to cannabis ranged from 11 to 59% at the DRUID cut-off, the corresponding range was 25–92% for more recent studies included in our review. At 1 ng/mL THC in blood, oral fluid test sensitivity ranged from 35 to 83% in more recent studies included in our review, and a multi-year observational study published since our search found sensitivities of 77% and 81% for two oral fluid screening tests at this cut-off in serum or plasma [60]. These results support continued investment in the development and evaluation of oral fluid screening tests, which have already been introduced into some North American drugged driving detection protocols. Alabama, Indiana, and Michigan (as part of a pilot program) currently use oral fluid screening tests for DUIC screening, and they are approved for use in DUIC screening in Canada. Importantly, positive oral fluid screening test results should always be confirmed with laboratory testing, as they should not be expected to achieve 100% specificity.

Compared to oral fluid screening tests, urine screening tests were found to have lower specificity for THC and higher sensitivity for THC, especially for blood THC levels <0.5 ng/mL. This is to be expected, as the major cannabinoid present in urine (and detected by screening tests) is not THC but a breakdown product (THC-COOH) that rises as blood THC levels fall and remains in urine long after blood THC levels are undetectable [61]. One consequence of this is that people with low levels of blood THC are likely to have THC-COOH in their urine; another is a high chance of false positives. Similar findings of low urine screening test specificity for blood THC detection were found in a recent study published after our search [60]. These results, taken together with the invasive nature of roadside urine collection and its hygiene and feasibility concerns, indicate that urine screening tests are generally inadvisable for roadside DUIC screening. However, they may be useful to jurisdictions with zero tolerance laws. Observational tests were challenged by their low THC sensitivity. This might have resulted from evaluators attributing signs of cannabis use to those of other drug classes, as DRE evaluations have been found to perform well in differentiating drug-impaired and non-drug impaired individuals [44,57,58]. There was also limited evidence for a reduction in observational test specificity with co-consumption of other substances, which should be investigated further. Tailoring observational tests to cannabis could improve their performance [55]. However, until improved performance is demonstrated, oral fluid screening tests are more promising.

Two potential limitations of oral fluid screening tests should be further investigated. First, test sensitivity could vary by frequency of cannabis use. There was a trend towards decreased sensitivity of oral fluid tests in frequent cannabis users, notably when screening for blood THC levels. While these differences must be explored further in larger studies, they can likely be explained by chronically elevated blood THC levels in frequent users. Whether this would present a limitation when screening for recent use is unclear, as blood THC levels in chronic users can remain elevated after several days of abstinence [62]. Second, the performance of oral fluid screening tests might vary depending on the time since cannabis consumption. In our review, sensitivity dropped rapidly post-consumption for some tests in comparisons to oral fluid laboratory tests. Future work is necessary on evaluating and improving this feature, and decision-makers should inquire about performance over time when selecting oral fluid tests.

This review had several limitations. First, the paucity of studies comparing screening tests to laboratory tests of blood limited our findings on urine and observational tests and may have limited our findings on oral fluid tests, as these were primarily evaluated against laboratory tests of oral fluid. Blood is the predominant bodily fluid used in legal cases, and while there is a good correlation between presence of THC in oral fluid and presence of THC in blood, there is not a good correlation between oral fluid and blood THC levels [15]. Second, our review is limited by the current lack of a gold standard for cannabis impairment detection. While laboratory tests for THC in blood are considered reliable indicators of recent cannabis intake and are currently used to lay DUIC charges, they are not reliable indicators of impairment [63]. Therefore, the sensitivity and specificity results summarized in our review should not be interpreted as indicative of tests' ability to differentiate between cannabis impairment and non-impairment. Third, imprecise sensitivity and specificity estimates hindered our conclusions for many tests; future studies should aim for enough reference test positives and negatives for precision. Fourth, inconsistency in laboratory test THC positivity prevalence across included studies might have led to variability in screening test performance, as test performance is impacted by the prevalence of positives in the examined sample [64]. We chose to focus our results on experimental studies and comparative studies as opposed to observational single-test studies in part to reduce this source of variability, as positivity prevalence within comparative studies and across experimental studies was more consistent. Lastly, as feasibility and acceptability assessments were extracted as available from included publications, our review may not represent the complete literature on these outcomes.

5. Conclusion

This systematic review synthesized and compared the diagnostic test characteristics of available roadside screening tests for cannabis use compared to laboratory tests for cannabinoids in oral fluid and blood. Our results support continued investment in the development and testing of oral fluid screening tests. These tests achieved higher specificity than urine screening tests in comparisons against blood laboratory tests for THC and higher sensitivity than observational tests (e.g., DRE evaluations) in comparisons against both blood and oral fluid laboratory tests for THC. Future studies should explore variation in oral fluid test performance by cannabis use frequency and over time following cannabis consumption, and attempt improvements to observational tests.

Declarations

Author contribution statement

All authors listed have significantly contributed to the development and the writing of this article.

Funding statement

Dr. Mark J Eisenberg was supported by the Canadian Institutes of Health Research [PJT-165867].

Data availability statement

Data will be made available on request.

Acknowledgements

The authors thank Ms. Jenna Glidden for her assistance with screening, data extraction, and quality assessment. We also wish to acknowledge Ms. Jill Boruff for her peer revision of the search strategies.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.heliyon.2023.e14630>.

References

- [1] O. Rogeberg, R. Elvik, The effects of cannabis intoxication on motor vehicle collision revisited and revised, *Addiction* 111 (8) (2016) 1348–1359.
- [2] M. Li, J. Brady, C. DiMaggio, A. Lusardi, K. Tzong, G. Li, Marijuana use and motor vehicle crashes, *Epidemiol. Rev.* 34 (2012) 65–72.
- [3] R.L. Hartman, M.A. Huestis, Cannabis effects on driving skills, *Clin. Chem.* 59 (3) (2013) 478–492.
- [4] M. Asbridge, J. Hayden, J. Cartwright, Acute cannabis consumption and motor vehicle collision risk: systematic review of observational studies and meta-analysis, *BMJ* 344 (2012) e536.
- [5] E. AH, Marijuana use and driving in Washington State: risk perceptions and behaviors before and after implementation of retail sales, *Traffic Inj. Prev.* 20 (1) (2019) 23–29.
- [6] F.J. Couper, B.L. Peterson, The prevalence of marijuana in suspected impaired driving cases in Washington state, *J. Anal. Toxicol.* 38 (8) (2014) 569–574.
- [7] S.B. Windle, M.J. Eisenberg, P. Reynier, et al., Association between legalization of recreational cannabis and fatal motor vehicle collisions in the United States: an ecologic study, *CMAJ Open* 9 (1) (2021) E233–E241.
- [8] M. Kruse, M. Perez, M. Blatt, et al., Marijuana Legalization and Rates of Crashing under the Influence of Tetrahydrocannabinol and Alcohol, *The American Surgeon*, 2021, 0003134821995053.
- [9] T.J. Lane, W. Hall, Traffic fatalities within US states that have legalized recreational cannabis sales and their neighbours, *Addiction* 114 (5) (2019) 847–856.
- [10] G. Bergen, A. Pitan, S. Qu, et al., Publicized sobriety checkpoint programs: a community guide systematic review, *Am. J. Prev. Med.* 46 (5) (2014) 529–539.
- [11] R.A. Shults, R.W. Elder, D.A. Sleet, et al., Reviews of evidence regarding interventions to reduce alcohol-impaired driving, *Am. J. Prev. Med.* 21 (4 Suppl) (2001) 66–88.
- [12] A. Razaghizad, S.B. Windle, G. Gore, et al., Interventions to prevent drugged driving: a systematic review, *Am. J. Prev. Med.* 61 (2) (2021) 267–280.
- [13] S.C.D. Dobri, A.H. Moslehi, T.C. Davies, Are oral fluid testing devices effective for the roadside detection of recent cannabis use? A systematic review, *Publ. Health* 171 (2019) 57–65.
- [14] J.N. Scherer, T.R. Fiorentin, B.T. Borille, et al., Reliability of point-of-collection testing devices for drugs of abuse in oral fluid: a systematic review and meta-analysis, *J. Pharmaceut. Biomed. Anal.* 143 (2017) 77–85.
- [15] M.B. Robertson, A. Li, Y. Yuan, et al., Correlation between oral fluid and blood THC concentration: a systematic review and discussion of policy implications, *Accid. Anal. Prev.* 173 (2022), 106694.
- [16] O.E. Johnson, G.M. Miskelly, J.D. Rindelaub, Testing for cannabis intoxication: current issues and latest advancements, *WIREs Forensic Sci.* 4 (4) (2022), e1450.
- [17] B.C. Ginsburg, Strengths and limitations of two cannabis-impaired driving detection methods: a review of the literature, *Am. J. Drug Alcohol Abuse* 45 (6) (2019) 610–622.
- [18] M.D.F. McInnes, D. Moher, B.D. Thombs, et al., Preferred reporting Items for a systematic review and meta-analysis of diagnostic test accuracy studies: the PRISMA-DTA statement, *JAMA* 319 (4) (2018) 388–396.
- [19] J. McGowan, M. Sampson, D.M. Salzwedel, E. Cogo, V. Foerster, C. Lefebvre, PRESS peer review of electronic search strategies: 2015 guideline statement, *J. Clin. Epidemiol.* 75 (2016) 40–46.
- [20] A. Agresti, B.A. Coull, Approximate is better than "exact" for interval estimation of binomial proportions, *Am. Statistician* 52 (2) (1998) 119–126.
- [21] P.F. Whiting, A.W. Rutjes, M.E. Westwood, et al., QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies, *Ann. Intern. Med.* 155 (8) (2011) 529–536.
- [22] A. Verstraete, ROSITA Roadside Testing Assessment: Final Report, Ghent: Rosita consortium, 2001.
- [23] M. Gronholm, P. Lillsunde, A comparison between on-site immunoassay drug-testing devices and laboratory results, *Forensic Sci. Int.* 121 (1–2) (2001) 37–46.
- [24] A. Verstraete, E. Raes, Rosita-2 Project: Final Report, Academia Press, Ghent, 2006.
- [25] M. Laloup, M. Del Mar Ramirez Fernandez, M. Wood, G. De Boeck, V. Maes, N. Samyn, Correlation of Delta9-tetrahydrocannabinol concentrations determined by LC-MS-MS in oral fluid and plasma from impaired drivers and evaluation of the on-site Drager DrugTest, *Forensic Sci. Int.* 161 (2–3) (2006) 175–179.
- [26] A. Pehrsson, T. Gunnar, C. Engblom, H. Seppa, A. Jama, P. Lillsunde, 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.* 175 (2–3) (2008) 140–148.
- [27] M. Concheiro, A. de Castro, O. Quintela, A. Cruz, M. Lopez-Rivadulla, Confirmation by LC-MS of drugs in oral fluid obtained from roadside testing, *Forensic Sci. Int.* 170 (2–3) (2007) 156–162.
- [28] D.J. Crouch, J.M. Walsh, L. Cangianelli, O. Quintela, Laboratory evaluation and field application of roadside oral fluid collectors and drug testing devices, *Ther. Drug Monit.* 30 (2) (2008) 188–195.
- [29] T. Blencowe, A. Pehrsson, P. Lillsunde, DRUID Deliverable 3.2.2: Analytical Evaluation of Oral Fluid Screening Devices and Preceding Selection Procedures, 2010.
- [30] T. Blencowe, A. Pehrsson, P. Lillsunde, et al., An analytical evaluation of eight on-site oral fluid drug screening devices using laboratory confirmation results from oral fluid, *Forensic Sci. Int.* 208 (1–3) (2011) 173–179.
- [31] A.S. Goessaert, K. Pijl, J. Veramme, A. Verstraete, Analytical evaluation of a rapid on-site oral fluid drug test, *Anal. Bioanal. Chem.* 396 (7) (2010) 2461–2468.
- [32] A. Pehrsson, T. Blencowe, K. Vimpari, K. Langel, C. Engblom, P. Lillsunde, An evaluation of on-site oral fluid drug screening devices DrugWipe 5+ and Rapid STAT using oral fluid for confirmation analysis, *J. Anal. Toxicol.* 35 (4) (2011) 211–218.

- [33] A. Pehrsson, T. Blencowe, K. Vimpari, A. Impinen, T. Gunnar, P. Lillsunde, Performance evaluation of the DrugWipe R 5/5+ on-site oral fluid screening device, *Int. J. Leg. Med.* 125 (5) (2011) 675–683.
- [34] S. Vanstechelma, C. Isalberti, T. Van der Linden, K. Pil, S.A. Legrand, A.G. Verstraete, Analytical evaluation of four on-site oral fluid drug testing devices, *J. Anal. Toxicol.* 36 (2) (2012) 136–140.
- [35] T.R. Arkell, R.C. Kevin, J. Stuart, et al., Detection of DELTA⁹ THC in oral fluid following vaporized cannabis with varied cannabidiol (CBD) content: an evaluation of two point-of-collection testing devices, *Drug Test. Anal.* 11 (10) (2019) 1486–1497.
- [36] D.J. Crouch, J.M. Walsh, R. Flegel, L. Cangianelli, J. Baudys, R. Atkins, An evaluation of selected oral fluid point-of-collection drug-testing devices, *J. Anal. Toxicol.* 29 (4) (2005) 244–248.
- [37] N.A. Desrosiers, D. Lee, D.M. Schwoppe, et al., On-site test for cannabinoids in oral fluid, *Clin. Chem.* 58 (10) (2012) 1418–1425.
- [38] N.A. Desrosiers, G. Milman, D.R. Mendu, et al., Cannabinoids in oral fluid by on-site immunoassay and by GC-MS using two different oral fluid collection devices, *Anal. Bioanal. Chem.* 406 (17) (2014) 4117–4128.
- [39] R.L. Hartman, S. Anizan, M. Jang, et al., Cannabinoid disposition in oral fluid after controlled vaporizer administration with and without alcohol, *Forensic Toxicol.* 33 (2) (2015) 260–278.
- [40] M.N. Newmeyer, M.J. Swortwood, M. Andersson, O.A. Abulseoud, K.B. Scheidweiler, M.A. Huestis, Cannabis edibles: blood and oral fluid cannabinoid pharmacokinetics and evaluation of oral fluid screening devices for predicting DELTA⁹-tetrahydrocannabinol in blood and oral fluid following cannabis brownie administration, *Clin. Chem.* 63 (3) (2017) 647–662.
- [41] R. Risoluti, G. Gullifa, A. Battistini, S. Materazzi, MicroNIR/Chemometrics: a new analytical platform for fast and accurate detection of DELTA⁹-Tetrahydrocannabinol (THC) in oral fluids, *Drug Alcohol Depend.* 205 (2019), 107578.
- [42] M.J. Swortwood, M.N. Newmeyer, O.A. Abulseoud, et al., On-site oral fluid DELTA⁹-tetrahydrocannabinol (THC) screening after controlled smoked, vaporized, and oral cannabis administration, *Forensic Toxicol.* 35 (1) (2017) 133–145.
- [43] S.W. Toennes, K. Schneider, C. Wunder, et al., Influence of ethanol on the pharmacokinetic properties of DELTA⁹-tetrahydrocannabinol in oral fluid, *J. Anal. Toxicol.* 37 (3) (2013) 152–158.
- [44] J.M. Walsh, D.J. Crouch, J.P. Danaceau, L. Cangianelli, L. Liddicoat, R. Adkins, Evaluation of ten oral fluid point-of-collection drug-testing devices, *J. Anal. Toxicol.* 31 (1) (2007) 44–54.
- [45] J.M. Walsh, R. Flegel, D.J. Crouch, L. Cangianelli, J. Baudys, An evaluation of rapid point-of-collection oral fluid drug-testing devices, *J. Anal. Toxicol.* 27 (7) (2003) 429–439.
- [46] C. Wanklyn, D. Burton, E. Enston, et al., Disposable screen printed sensor for the electrochemical detection of delta-9-tetrahydrocannabinol in undiluted saliva, *Chem. Cent. J.* 10 (2016) 1.
- [47] S.M. Wille, V. Di Fazio, S.W. Toennes, J.H. van Wel, J.G. Ramaekers, N. Samyn, Evaluation of DELTA(9) -tetrahydrocannabinol detection using DrugWipe5S(R) screening and oral fluid quantification after Quantisal TM collection for roadside drug detection via a controlled study with chronic cannabis users, *Drug Test. Anal.* 7 (3) (2015) 178–186.
- [48] A. Arroyo, M. Sanchez, E. Barberia, M. Barbal, M.T. Marron, A. Mora, Comparison of the Cozart DDS 801 on-site drug test device and gas chromatography/mass spectrometry (GC/MS) confirmation results of cannabis and cocaine in oral fluid specimens, *Aust. J. Forensic Sci.* 46 (3) (2014) 272–281.
- [49] S.M. Wille, N. Samyn, M. Ramirez-Fernandez Mdel, G. De Boeck, Evaluation of on-site oral fluid screening using Drugwipe-5(+), RapidSTAT and Drug Test 5000 for the detection of drugs of abuse in drivers, *Forensic Sci. Int.* 198 (1–3) (2010) 2–6.
- [50] H. Gjerde, G.B. Clausen, E. Andreassen, H. Furuhaugen, Evaluation of drager DrugTest 5000 in a naturalistic setting, *J. Anal. Toxicol.* 42 (4) (2018) 248–254.
- [51] F. Musshoff, E.G. Hokamp, U. Bott, B. Madea, Performance evaluation of on-site oral fluid drug screening devices in normal police procedure in Germany, *Forensic Sci. Int.* 238 (2014) 120–124.
- [52] R.P. Compton, USDo Transportation, Field Evaluation of the Los Angeles Police Department Drug Detection Procedure, National Highway Traffic Safety Administration, Washington, D.C, 1986.
- [53] L.D. Edwards, K.L. Smith, T. Savage, Drugged driving in Wisconsin: oral fluid versus blood, *J. Anal. Toxicol.* 41 (6) (2017) 523–529.
- [54] I. Fierro, J.C. Gonzalez-Luque, F.J. Alvarez, The relationship between observed signs of impairment and THC concentration in oral fluid, *Drug Alcohol Depend.* 144 (2014) 231–238.
- [55] R.L. Hartman, J.E. Richman, C.E. Hayes, M.A. Huestis, Drug Recognition Expert (DRE) examination characteristics of cannabis impairment, *Accid. Anal. Prev.* 92 (2016) 219–229.
- [56] D.F. Preusser, R.G. Ulmer, C.W. Preusser, Evaluation of the Impact of the Drug Evaluation and Classification Program on Enforcement and Adjudication, 1992.
- [57] M.N. Walden, Factors influencing field performance: utilizing the Drug Evaluation and Classification (DEC) Program to identify suspected impaired drivers as reported by selected certified police officers in Texas, *Dissertation Abstracts Int. Sec. A: Humanit. Soc. Sci.* 70 (2-A) (2009) 547.
- [58] T.D. Walden, The Drug Evaluation and Classification Program in the State of Texas: A Validation Study, 2005.
- [59] H. Schulze, M. Schumacher, R. Urmeew, et al., Driving under the Influence of Drugs, Alcohol, and Medicines in Europe - Findings from the DRUID Project, 2012.
- [60] J. Liut, U. Bott, B. Madea, M. Krämer, A. Maas, Evaluation of RapidSTAT®, DrugWipe® 6S, DrugScreen® 5TK and DrugScreen® 7TR for on-site drug testing in German police roadside traffic patrol, *Drug Test. Anal.* 14 (8) (2022) 1407–1416.
- [61] F. Grotenhermen, Pharmacokinetics and pharmacodynamics of cannabinoids, *Clin. Pharmacokinet.* 42 (4) (2003) 327–360.
- [62] W.M. Bosker, E.L. Karschner, D. Lee, et al., Psychomotor function in chronic daily Cannabis smokers during sustained abstinence, e53127, *PLoS One* 8 (1) (2013) e53127.
- [63] J.G. Ramaekers, M.R. Moeller, P. van Ruitenbeek, E.L. Theunissen, E. Schneider, G. Kauert, Cognition and motor control as a function of Delta9-THC concentration in serum and oral fluid: limits of impairment, *Drug Alcohol Depend.* 85 (2) (2006) 114–122.
- [64] M.M.G. Leeflang, A.W.S. Rutjes, J.B. Reitsma, L. Hooft, P.M.M. Bossuyt, Variation of a test's sensitivity and specificity with disease prevalence, *CMAJ (Can. Med. Assoc. J.)* 185 (11) (2013) E537–E544.