



Do we need urine drug screens in opioid addiction treatment: An observational study on self-report versus urine drug screens

Alannah McEvoy^a, Myanca Rodrigues^b, Brittany B Dennis^{c,d}, Jacqueline Hudson^a, David C. Marsh^e, Andrew Worster^b, Lehana Thabane^b, Zainab Samaan^{a,b,f,*}

^a Department of Psychiatry and Behavioural Neurosciences, McMaster University, 100 West 5th St., Hamilton, ON L8N 3K7, Canada

^b Department of Health Research Methods, Evidence, and Impact, McMaster University, 1280 Main St. W, Hamilton, ON L8S 4L8, Canada

^c Department of Medicine, Division of Social Medicine, University of British Columbia, Vancouver, BC V5Z1M9, Canada

^d British Columbia Centre on Substance Use, Vancouver, BC V6Z 2A9, Canada

^e NOSM University, 935 Ramsey Lake Rd., Sudbury, ON P3E 2C6, Canada

^f Department of Psychiatry, Queen's University, Kingston, ON K7L 4X3, Canada

ARTICLE INFO

Keywords:

Medication assisted treatment
Opioid use disorder
Sensitivity
Specificity
Urine drug screens

ABSTRACT

Background: The substance use crisis continues to progress. Medication for Opioid Use Disorder (MOUD) are prescribed to reduce opioid use and related harms; however, many individuals continue to use substances while on treatment. The objective of this study was to describe the temporal and demographic trends of the agreement between self-reported and urine tested substances.

Methods: The current study is a retrospective secondary analysis of three phases of a prospective cohort study (Pilot 2011, Genetics of opioid addiction (GENOA) 2013–2017, and Pharmacogenetics of opioid substitution treatment (POST)) 2018–2022) spanning 2011–2022. We compared the self-reported substance use data for opioids, benzodiazepines, amphetamine/methamphetamine (AMP/MET), and cocaine with urine drug results. We compared the positive predictive value (PPV), false omission rate (FOR), sensitivity, and specificity between (i) different drugs; (ii) by sex, and (iii) age group at enrollment in each phase of the study using self-reported substance use at baseline and retrospective electronic health record data on urine drug screenings collected over the same time period.

Results: Overall, the average PPV and FOR for any drug across all phases was 80.7 % and 37.9 %, respectively. Sensitivity and specificity were highest for cocaine and lowest for benzodiazepines. We found no specific trend by sex. Lastly, we found a higher sensitivity for opioids and AMP/MET in those under 25 years of age compared to other age groups. PPV increased over time for benzodiazepines, AMP/MET and cocaine and FOR was higher during the pilot and POST phases than the GENOA phase.

Conclusion: Our study highlights the unique challenges associated with ascertaining substance use behaviour for individuals receiving MOUD, indicating many patients will accurately report substance use while others do not. It is therefore important to consider the context of the patient, and the type of the co-substance used to select patient-centred testing as indicated. Therefore, the answer to the question of do we need urine drug screen is yes in some cases.

1. Introduction

The substance use crisis continues to progress, marked by rising mortality rates from overdose, with an increasing trend in Canada of opioid-related deaths (The Government of Canada, 2022). Opioid Use Disorder (OUD) is a chronic remitting condition that often requires long-term care. To reduce opioid-related death and harm, Medication for

Opioid Use Disorder (MOUD) have been employed to reduce opioid cravings, withdrawal symptoms, and use (Coffa & Snyder, 2019; Morin et al., 2017). Despite the efforts to reduce opioid use, it is not uncommon for patients to continue to use illicit opioids while receiving treatment, and in addition to opioid use, individuals with OUD are likely to use multiple substances (Compton et al., 2021; Mattick et al., 2002; Oviedo-Joekes et al., 2009).

* Corresponding author at: Department of Psychiatry, Queen's University, 752 King St. W, Kingston, ON K7L 4X3, Canada.
E-mail address: z.samaan@queensu.ca (Z. Samaan).

Urine drug screening (UDS) is a common practice in drug treatment programs to evaluate adherence with MOUD and monitor illicit drug use and is considered the gold standard of assessing substance use (Owen et al., 2012; Pergolizzi et al., 2010; Wilcox et al., 2013). However, UDS are not without their caveats including high cost, varying and narrow windows for detection based on the substance, inaccuracy, inaccessibility in certain areas for patients and a lack of information on the quantity of use (Moeller et al., 2017, 2008; Morin et al., 2017; Wilcox et al., 2013). Further, clinics continue to practice direct observation of urine sampling and patients have described UDS methods as embarrassing, degrading, and invasive, while some, but not all participants, have noticed a gradual process of acceptance of this practice (Strike & Rufo, 2010).

The validity of using self-reported drug use, as a measure for ongoing drug use, has been previously discussed in the literature, with mixed results reported. A study investigating the validity of self-reported drug use of individuals receiving MOUD found that individuals underreported cocaine and opioid use despite high agreement amongst other drugs, similar to a study of women with co-occurring post-traumatic stress and substance use disorders (Ruglass et al., 2023; Wilcox et al., 2013). A study on cocaine use treatment found that participants were more motivated to report drug use accurately at the beginning of the study with declining motivation as it progressed (Schuler et al., 2009). A more recent systematic review of the agreement between self-reported substance use and biological test screening (including urine) reported agreement between these methods of checking substance use, however it was not exclusively focused on patients with an opioid use disorder (Bharat et al., 2023).

Finally, in a study of individuals not currently in a formal drug treatment program, UDS was found to be an important assessment tool when treating substance use disorders as 40 % of participants denied drug use despite corresponding positive UDS (McDonnell et al., 2016). Despite the number of studies investigated the agreement between self-report and urine drug tests, the findings remain inconsistent and considering the ongoing opioid crisis and the need for safe and effective management of opioid use disorder, further investigations related to the reliance on urine drug screens is needed.

Demographic trends may affect the agreement between self-reported drug use and UDS. It was reported that participants who were older, female, had a disability, in legal and social programs, and had more severe drug problems were more likely to underreport drug use (McDonnell et al., 2016). In contradiction, another study found that females were more likely to accurately self-report drug use compared to men (Schuler et al., 2009). Further possible trends in underreporting substance use include being pregnant, employed, being of African American descent and being younger in age (Wilcox et al., 2013).

Thus, the objective of this study was to complete a retrospective secondary analysis to describe both the temporal and demographic trends comparing self-reported drug use and UDS, for three phases of observational cohort for OUD treatment conducted during 2011–2022. Understanding the accurate disclosure of self-report drug use is important to ensure we minimise unnecessary tests and burden to patients.

2. Material and methods

This study is reported in accordance with the STROBE guidelines (Supplementary File 1).

2.1. Setting and participants

Data for the retrospective secondary analysis were collected in three phases of prospective cohort study conducted in Ontario, Canada. Each phase was conducted as a separate study. The objective of the cohort study was to examine the biological and social factors affecting treatment outcomes for patients with OUD. The study was conducted in collaboration with the Canadian Addiction Treatment Centre (CATC)

and McMaster University and recruited participants who met criteria for OUD, were receiving a MOUD and were 18 years or older in Pilot and GENetics of Opioid Addiction (GENOA) phases and 16 years or older in Pharmacogenetics of Opioid Substitution Treatment response (POST). Briefly, the GENOA Pilot study phase (n = 260), henceforth referred to as the pilot phase, was conducted at four clinical sites between June and December 2011, the GENOA study phase (n = 1390) was conducted at 20 clinical sites between January 2013 and April 2017 and the POST phase (n = 2625) was conducted at 54 CATC sites between May 2018 to December 2022. Details on each phase and final numbers included in the retrospective secondary analyses after exclusions can be found in [Supplementary File 2](#) and [Fig. 1](#).

The study was conducted in accordance with the Declaration of Helsinki Ethical Principles for Medical Research Involving Human Subjects. Ethics approval was obtained by the Hamilton Integrated Research Ethics Board (pilot and GENOA: 11–056; POST: 4556), and all participants provided verbal and written informed consent.

2.2. Data sources and measures

Self-reported drug use was collected at study enrollment from the Mini International Neuropsychiatric Interview (M.I.N.I.) for pilot participants and the Maudsley Addition Profile (MAP) for GENOA and POST participants. The M.I.N.I. collected self-reported drug use for the past 12 months while the MAP collected self-reported drug use over the past 30 days. The M.I.N.I. asks participants to self-report if they have ever used a substance illicitly more than once within the past 12 months for stimulants, cocaine, opiates, hallucinogenic, dissociative drugs, inhalants, cannabis, sedatives, hypnotics, anxiolytics, benzodiazepines, tobacco and any other drugs used but not captured (Sheehan et al., 1998). The MAP asks participants to self-report how many days in the past 30 days they had used a substance for alcohol, heroin, illicit methadone, illicit benzodiazepines, cocaine powder, crack cocaine, amphetamines, cannabis and any other drugs used but not captured, including opioids (Marsden et al., 2002).

UDS were collected as per usual clinical care at the CATC, which averages to approximately once per week. Urine samples were supervised, tested, analyzed, and reported as the number of positive screens for each patient for the drug detected in the test using the FaStep Assay (Trimed Supply Network Ltd, Concord, Ontario, Canada) (Trimed Supply Network Ltd., n.d.). Differences exist between which drugs were tested using UDS in each study phase which is reported in [Supplementary File 2](#). The pilot phase collected UDS on opiates, oxycodone, amphetamines, THC, ecstasy, barbiturates, benzodiazepines, cocaine, and ethanol. The GENOA phase collected UDS on opioids, benzodiazepines, THC, amphetamines, cocaine, methamphetamines, barbiturates, and ketamine. The POST phase collected UDS on opioids, benzodiazepines, THC, amphetamines, cocaine, and methamphetamines.

Only UDS data (from participants electronic medical records as part of routine care) collected at study enrollment was included in the current study as the time-frame overlapped with the self-reported drug use data.

2.3. Sample size

Participants were included in this study only if they were unique individuals within each study phase, (if they participated in different study phases unintentionally, they were counted once in each study phase). Within each study phase participants were excluded if duplicate or if they were missing UDS. As the three phases of the study recruited participants from overlapping clinics over a 11- year period and, it is plausible that between-study phase duplicates exist. Duplicates between phases were not removed. All participants were receiving and taking their respective MOUD, as evidenced by their UDS results. The total sample size included in this study is n = 4181.

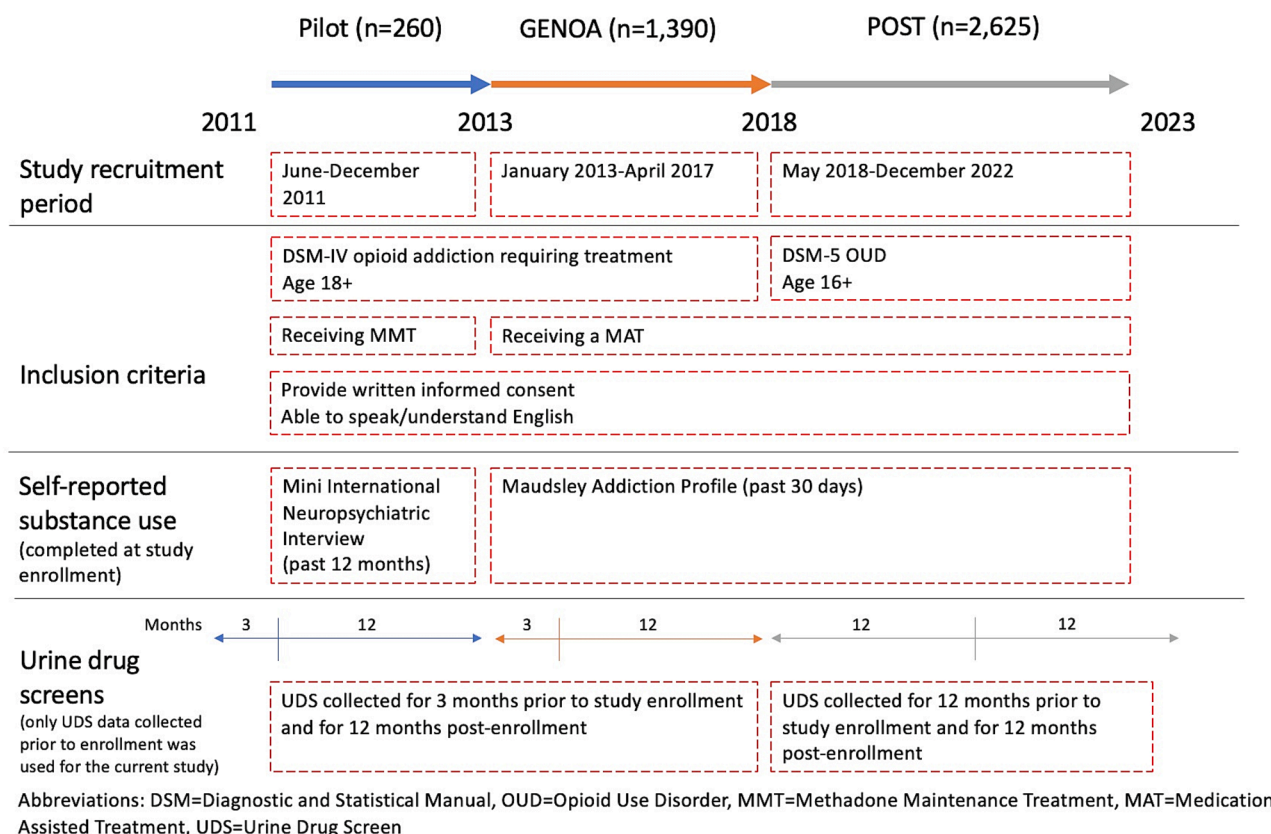


Fig. 1. Study Timeline.

2.4. Statistical methods

Descriptive statistics were reported for each study phase, by sex, to describe the demographic and clinical characteristics. Continuous variables were expressed as means with standard deviations, while categorical variables were expressed as counts.

Self-reported drug use and UDS data were collected at study enrollment, during overlapping time periods. Self-reported drug use and UDS data were converted to binary data points wherein any self-reported drug use or positive UDS for a substance was reported as a positive self-reported drug or positive UDS, respectively, and no reported drug use or only negative UDS were reported as negative self-reported drug use or negative UDS, respectively. Positive predictive value (PPV), the proportion of individuals who self-reported drug use and had a positive UDS for all individuals who self-reported drug use, false omission rate (FOR), the proportion of individuals who did not self-report drug use and had a positive UDS of all those who self-reported no drug use, sensitivity (which is the ability of a diagnostic test to detect those who truly have the condition) and specificity (which is the ability of a diagnostic test to detect those who truly do not have the condition) were calculated for opioids, benzodiazepines, amphetamines and methamphetamines (AMP/MET) and cocaine across all study phases separately. PPV was calculated by dividing the true positives (TP; those who reported drug use and had a positive UDS) by the sum of the true positives and false positives (FP; those who reported drug use, but UDS was negative) ($TP/[TP + FP]$) (Trevethan, 2017). FOR was calculated by dividing the false negatives (FN) by the sum of the false negatives (those who reported no drug use, but UDS was positive) and true negatives (those who reported no drug use and UDS was negative) (TN): ($FN/[FN + TN]$) (Akobeng, 2007; Kost, 2021). Sensitivity was calculated by dividing the TP by the sum of the TP and FN and multiplying by 100: ($TP/[TP + FN]$)*100 (Trevethan, 2017). Specificity was calculated by dividing the TN by the sum of the TN and FP and multiplying by 100:

($TN/[TN + FP]$)*100 (Trevethan, 2017). The overall PPV and FOR was taken for each substance by averaging the PPVs and FORs for each study phase. Sensitivity and specificity were calculated by using UDS results as the true condition, or disease status, and self-reported data as the test. PPV, FOR, sensitivity and specificity were also conducted separately for men and women to test for sex differences in self-reported drug use as well as across age groups including those < 25 years of age, adults between 25–54 years of age and adults over the age 55 years, commonly defined age groups in OUD (Krebs et al., 2021; Rieb et al., 2020).

Participants were excluded from the analyses if 1) they did not have => three urine drug screen tests results, or 2) they had a prescription for the drug of interest because the urine drug test is expected to be positive for this drug.

All analyses were conducted using Stata MP version 17 (StataCorp, 2021).

3. Results

Altogether 4181 participants were included in this analysis, 248 from the pilot, 1357 from the GENOA and 2576 from the POST phase. Fig. 2 illustrates the study flow diagram. Supplementary Figures S1-3 illustrate the study flow diagram and participants excluded for each analysis. Briefly, 239 participants were included for opioids, 212 participants were included for benzodiazepines, 216 participants were included for amphetamines and 248 participants were included for cocaine in the pilot phase. For the GENOA phase, 1240 participants were included for opioids, 1060 participants were included for benzodiazepines, 575 participants were included for amphetamines and 1255 participants were included for cocaine. For the POST phase, 2375 participants were included for opioids, 2068 participants were included for benzodiazepines, 2165 participants were included for amphetamines and 2437 participants were included for cocaine. Participant characteristics by study phase are summarized in Table 1.

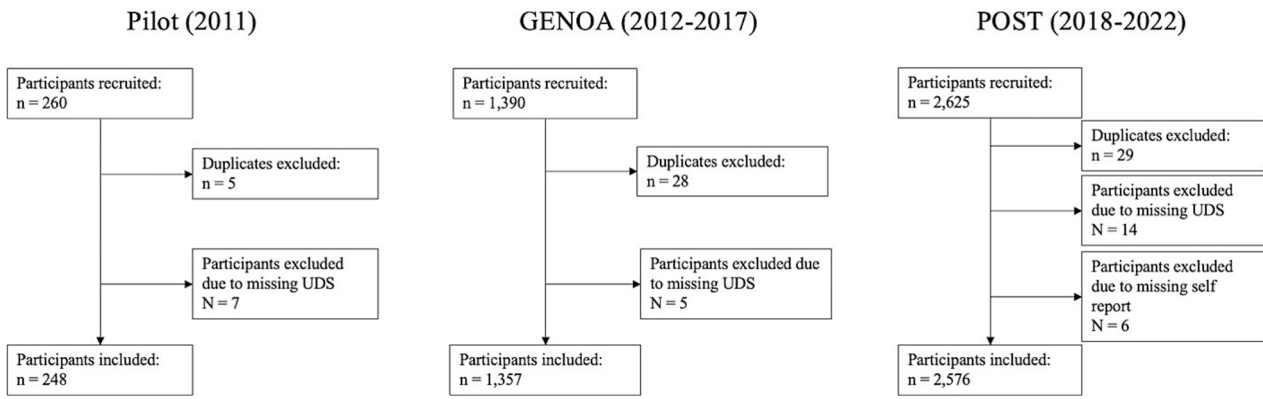


Fig. 2. Study Flow Diagram.

Table 1 Participant demographics.

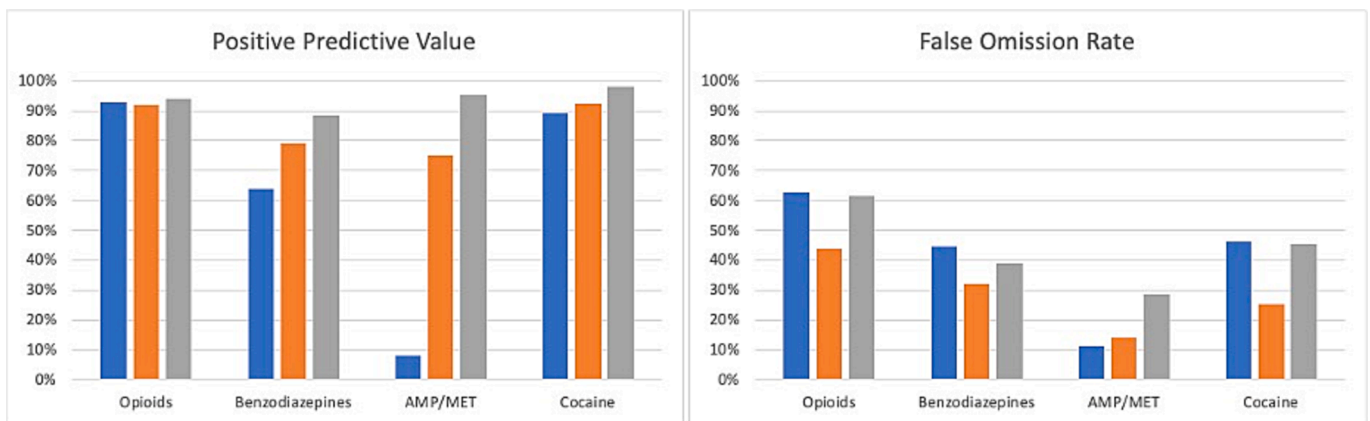
	Pilot (2011)			GENOA (2013–2017)			POST (2018–2022)		
	Total n = 248	Male n = 147	Female n = 101	Total n = 1357	Male n = 757	Female n = 600	Total n = 2576 ^a	Male n = 1419	Female n = 1156
Age (years); mean (SD)	37.0 (10.3)	38.2 (10.8)	35.2 (9.3)	38.5 (11.1)	39.3 (11.4)	37.5 (10.6)	39.5 (10.9)	39.9 (10.9)	39.0 (10.8)
Unemployment; n (%)	175 (70.6)	98 (66.7)	77 (76.2)	887 (65.5) ^b	449 (59.5) ^c	438 (73.1) ^d	1741 (67.6)	864 (60.9)	877 (75.9)
Age at first regular opioid use (years); mean (SD)	23.3 (9.4)	23.4 (9.9)	23.2 (8.6)	25.4 (8.9)	25.5 (9.1)	25.2 (8.6)	24.9 (9.3)	25.0 (9.6)	24.8 (8.9)
Years in current treatment; mean (SD)	3.2 (3.5)	3.5 (3.3)	2.9 (3.7)	3.7 (4.0)	3.7 (4.0)	3.6 (4.0)	4.4 (5.1)	4.3 (4.9)	4.6 (5.3)
MOUD									
Methadone, n (%)	248 (100)	147 (100)	101 (100)	1276 (94.0)	709 (93.7)	567 (94.5)	2042 ^a (79.3)	1133 (80.0)	908 (78.6)
Suboxone, n (%)	0 (0)	0 (0)	0 (0)	80 (5.9)	48 (6.3)	32 (5.3)	529 (20.5)	285 (20.1)	244 (21.1)
Other, n (%)	0 (0)	0 (0)	0 (0)	1 (0.1)	0 (0)	1 (0.2)	5 (0.2)	1 (0.1)	4 (0.4)
Opioids prescription; n (%)	8 (3.2)	5 (3.4)	3 (3.0)	45 (3.3)	24 (3.2)	21 (3.5)	100 (3.9)	60 (4.2)	40 (3.5)
Benzodiazepine prescription; n (%)	36 (14.5)	22 (15.0)	14 (13.9)	222 (16.4)	122 (16.1)	100 (16.7)	411 (16.0)	201 (14.2)	209 (18.8)
Stimulant prescription; n (%)	7 (2.8)	5 (3.4)	2 (2.0)	49 (3.6)	26 (3.4)	23 (3.8)	185 (7.2)	108 (7.6)	77 (6.7)

^a One individual reported being intersex and was therefore not included in the male or female counts.

^b Data available for n = 1354.

^c Data available for n = 755.

^d Data available for n = 599.



Legend: Blue = Pilot (2011), Orange = GENOA (2013-2017), Grey = POST (2018-2022)

Fig. 3. PPV and FOR by Study Phase.

Overall, of the participants in any study phase who self-reported opioid use, 93.0 % had a positive opioid UDS and 56.0 % who self-reported no opioid use had a positive opioid UDS. For benzodiazepine use, 77.1 % of participants who self-reported benzodiazepine use had a corresponding positive UDS and 38.7 % of participants who self-reported no benzodiazepine use had a positive UDS. For AMP/MET use, 59.4 % of participants who self-reported AMP/MET use had a positive UDS and 18.0 % participants who self-reported no AMP/MET use had a positive UDS. Finally, for cocaine use, 93.2 % of participants who self-reported cocaine use had a positive UDS and 39.0 % of participants who self-reported no cocaine use had a positive UDS. Thus, on average the PPV and FOR for any drug across all study phases was 80.7 % and 37.9 %, respectively. Fig. 3 shows the PPV and FOR for each study.

3.1. Descriptive trends across studies

Results for PPV, FOR, Sensitivity and Specificity by drug can be seen in Table 2 and results by age can be seen in Table 3. Trends in the results provided in Table 2 and 3 are shown using bar charts in Figs. 3-7 wherein Fig. 3 depicts the PPV (individuals who self-reported drug use and had a positive UDS) and FOR (individuals who did not self-report drug use and had a positive UDS) for each study, Fig. 4 depicts the PPV by sex, Fig. 5 depicts the FOR by sex, Fig. 6 depicts the PPV by age, and Fig. 7 depicts the FOR by age.

Table 2
PPV, FOR, sensitivity and specificity.

	Pilot			GENOA			POST		
	Total	Males	Females	Total	Males	Females	Total	Males	Females
Opioids									
Positive Predictive Value (%)	92.96	92.68	93.33	91.87	93.49	89.47	94.14	93.54	94.94
False Omission Rate (%)	62.89	61.02	65.79	43.89	43.27	44.59	61.31	60.59	62.16
Sensitivity [% (95 % CI)]	68.39 (61.33, 74.88)	67.86 (58.37, 76.37)	69.14 (57.89, 78.93)	38.24 (34.57, 42.01)	41.58 (36.57, 46.72)	34.00 (28.65, 39.67)	32.30 (30.04, 34.62)	33.93 (30.83, 37.13)	30.24 (26.96, 33.68)
Specificity [% (95 % CI)]	78.26 (63.64, 89.05)	79.31 (60.28, 92.01)	76.47 (50.10, 93.19)	95.89 (93.90, 97.38)	96.36 (93.58, 98.17)	95.35 (92.02, 97.57)	95.51 (93.74, 96.89)	94.83 (92.20, 96.77)	96.34 (93.70, 98.10)
Benzodiazepines									
Positive Predictive Value (%)	63.79	53.13	76.92	79.00	73.21	86.36	88.41	89.42	86.67
False Omission Rate (%)	44.81	48.39	39.34	32.29	31.43	33.33	39.02	36.49	42.17
Sensitivity [% (95 % CI)]	34.91 (25.90, 44.78)	27.42 (16.85, 40.23)	45.45 (30.39, 61.15)	20.31 (16.42, 24.65)	19.90 (14.68, 26.02)	20.77 (14.88, 26.95)	16.33 (13.96, 18.93)	19.46 (16.00, 23.29)	12.68 (9.62, 16.30)
Specificity [% (95 % CI)]	80.19 (71.32, 87.30)	76.19 (63.79, 86.02)	86.05 (72.07, 94.70)	96.87 (95.26, 98.05)	96.00 (93.49, 97.74)	97.98 (95.64, 99.25)	98.39 (97.50, 99.03)	98.38 (97.13, 99.19)	98.40 (96.87, 99.31)
AMP/MET									
Positive Predictive Value (%)	7.89	9.09	6.25	75.00	76.92	72.73	95.22	94.09	96.85
False Omission Rate (%)	11.24	14.42	6.76	14.04	11.68	16.95	28.69	30.10	27.06
Sensitivity [% (95 % CI)]	13.04 (2.78, 33.59)	11.76 (1.46, 36.44)	16.67 (0.42, 64.12)	32.73 (24.08, 42.33)	37.04 (24.29, 51.26)	28.57 (17.30, 42.21)	36.02 (32.75, 39.40)	37.00 (32.63, 41.53)	34.55 (29.62, 39.74)
Specificity [% (95 % CI)]	81.86 (75.69, 87.03)	81.65 (73.09, 88.42)	82.14 (72.26, 89.65)	97.42 (95.54, 98.66)	97.72 (95.10, 99.16)	97.03 (93.65, 98.90)	98.88 (98.15, 99.37)	98.44 (97.22, 99.22)	99.37 (98.39, 99.83)
Cocaine									
Positive Predictive Value (%)	89.29	89.23	89.36	92.37	96.99	87.60	97.89	98.41	97.22
False Omission Rate (%)	46.32	50.00	40.74	25.28	27.12	22.95	45.32	46.61	43.72
Sensitivity [% (95 % CI)]	61.35 (53.42, 68.86)	58.59 (48.24, 68.40)	65.63 (52.70, 77.05)	49.09 (44.59, 53.60)	46.34 (40.27, 52.28)	52.80 (45.88, 59.65)	44.77 (42.20, 47.37)	45.35 (41.90, 48.84)	44.09 (40.19, 48.05)
Specificity [% (95 % CI)]	85.88 (76.64, 92.49)	85.42 (72.24, 93.93)	86.49 (71.23, 95.46)	97.38 (95.98, 98.39)	99.02 (97.50, 99.73)	95.49 (92.78, 97.40)	98.58 (97.62, 99.22)	98.84 (97.50, 99.57)	98.28 (96.64, 99.25)

CI = Confidence interval.

PPV was similar for all phases for opioids and cocaine while increase in PPV over time was seen for benzodiazepines and AMP/MET. An increase in FOR was seen for AMP/MET only (see Fig. 3). FOR was higher in the pilot and POST phase than during GENOA for opioids, benzodiazepines and cocaine. Specificity was consistently higher in the POST and GENOA phase when compared to the pilot study phase and sensitivity and specificity were highest for cocaine and lowest for benzodiazepines (Table 3). There was an increase in PPV for benzodiazepines and AMP/MET in men. In women only AMP/MET PPV was increased over time. No difference was seen in opioids or cocaine PPV by sex over time (see Fig. 4). FOR was lower in GENOA for all drugs in men while in women FOR was lower in GENOA for all except AMP/MET (see Fig. 5). PPV increased over time for benzodiazepines and AMP/MET for those under 55 years of age while only benzodiazepines PPV increased over-time for those 55 years and older (see Fig. 6). FOR increased for AMP/MET over time for all age groups (see Fig. 7). Lastly, sensitivity was found to be higher for opioids and AMP/MET in participants < 25 years of age.

4. Discussion

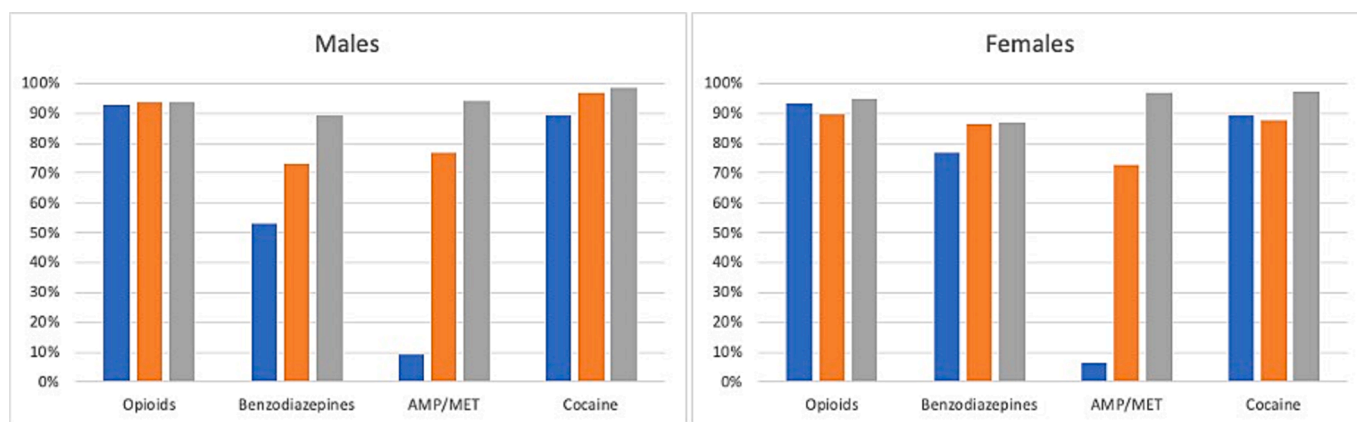
Overall, we found that 80.7 % of participants self-reported drug use and had a corresponding positive UDS for that substance (PPV) and 37.9 % of participants who said they had no current drug use had a corresponding positive UDS (FOR). Further, we found that the FOR was

Table 3
PPV, FOR, sensitivity and specificity by age.

	Pilot			GENOA			POST		
	<25 years	25–54 years	55+ years	<25 years	25–54 years	55 + years	<25 years	25–54 years	55 + years
Opioid									
Positive Predictive Value (%)	100	92.04	83.33	97.56	90.67	94.12	82.22	94.96	97.62
False Omission Rate (%)	0	63.95	75.00	33.82	46.03	34.86	61.62	61.45	60.19
Sensitivity [% (95 % CI)]	100 (85.18, 100*)	65.41 (57.47, 72.77)	45.45 (16.75, 76.62)	63.49 (50.40, 75.27)	36.23 (32.26, 40.36)	29.63 (17.98, 43.61)	37.76 (28.16, 48.12)	32.87 (30.39, 35.43)	24.40 (18.12, 31.62)
Specificity [% (95 % CI)]	100 (29.24, 100*)	77.50 (61.55, 89.16)	66.67 (9.43, 99.16)	97.83 (88.47, 99.94)	95.25 (92.83, 97.04)	98.61 (92.50, 99.96)	82.61 (68.58, 92.18)	96.02 (94.14, 97.43)	98.82 (93.62, 99.97)
Benzodiazepine									
Positive Predictive Value (%)	25.00	70.83	50.00	86.67	78.38	72.73	77.78	89.23	93.75
False Omission Rate (%)	52.63	45.24	22.22	31.25	31.03	43.30	36.84	39.45	36.95
Sensitivity [% (95 % CI)]	16.67 (2.09, 48.41)	37.36 (27.44, 48.13)	33.33 (0.84, 90.57)	34.21 (19.63, 51.35)	19.27 (14.97, 24.19)	16.00 (7.17, 29.11)	25.00 (14.39, 38.37)	15.63 (13.09, 18.45)	16.67 (9.64, 26.00)
Specificity [% (95 % CI)]	60.00 (32.29, 83.66)	83.13 (73.32, 90.46)	87.50 (47.35, 99.68)	96.49 (87.89, 99.57)	97.12 (95.37, 98.35)	94.83 (85.62, 98.92)	94.74 (87.07, 98.55)	98.56 (97.60, 99.21)	99.22 (95.76, 99.98)
AMP/MET									
Positive Predictive Value (%)	12.50	6.67	–	83.33	71.05	100	100	94.49	100
False Omission Rate (%)	7.14	12.50	0	8.11	14.85	11.86	30.69	29.47	22.61
Sensitivity [% (95 % CI)]	50.00 (1.26, 98.74)	9.52 (1.16, 30.38)	N/A	62.50 (24.49, 91.48)	29.67 (20.55, 40.16)	36.36 (10.93, 69.21)	44.64 (31.34, 58.53)	36.45 (32.89, 40.13)	24.64 (15.05, 36.49)
Specificity [% (95 % CI)]	65.00 (40.78, 84.61)	82.61 (75.86, 88.12)	100 (73.54, 100*)	97.14 (85.08, 99.93)	97.09 (94.85, 98.54)	100 (93.15, 100)	100 (94.87, 100*)	98.62 (97.73, 99.23)	100 (97.95, 100*)
Cocaine									
Positive Predictive Value (%)	93.33	88.54	100	88.24	92.54	94.12	97.62	97.88	98.28
False Omission Rate (%)	50.00	45.95	46.15	29.55	25.86	18.10	47.52	47.22	31.31
Sensitivity [% (95 % CI)]	70.00 (45.72, 88.11)	62.50 (53.79, 70.65)	14.29 (0.36, 57.87)	36.59 (22.12, 53.06)	50.84 (45.92, 55.75)	43.24 (27.10, 60.51)	46.07 (35.44, 56.96)	44.56 (41.77, 47.38)	45.97 (36.99, 55.15)
Specificity [% (95 % CI)]	85.71 (42.13, 99.64)	84.51 (73.97, 92.00)	100.00 (59.04, 100*)	96.88 (89.16, 99.62)	97.18 (95.52, 98.35)	98.96 (94.33, 99.97)	98.15 (90.11, 99.95)	98.46 (97.33, 99.20)	99.32 (96.29, 99.98)

CI = Confidence interval.

* One sided, 97.5 % CI.



Legend: Blue = Pilot (2011), Orange = GENOA (2013-2017), Grey = POST (2018-2022)

Fig. 4. PPV by Sex.

highest for opioids, followed by cocaine, benzodiazepines and AMP/MET. The PPV of 80.7 % indicates that 8 out of 10 individuals were accurately reporting substance use that was also detected in urine drug screens. While the FOR (37.9 %) was poor in keeping with previous

studies (Bharat et al., 2023) which reported FOR of 0.30, for cocaine use in RCTs, it may be argued that FOR in this clinical setting was influenced by perceived consequences for reporting substance use, such as an inability to increase dose or no “take-home” dose privileges, although

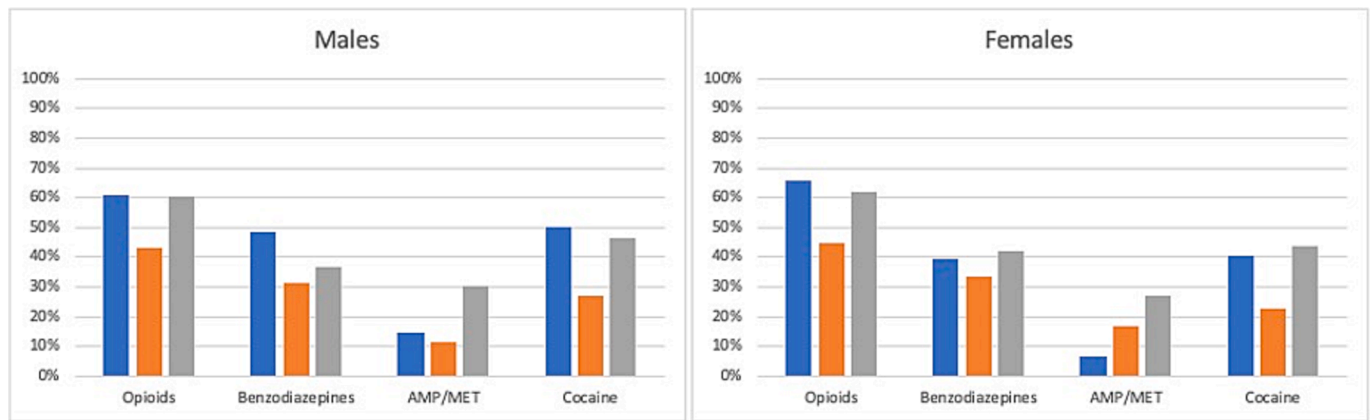
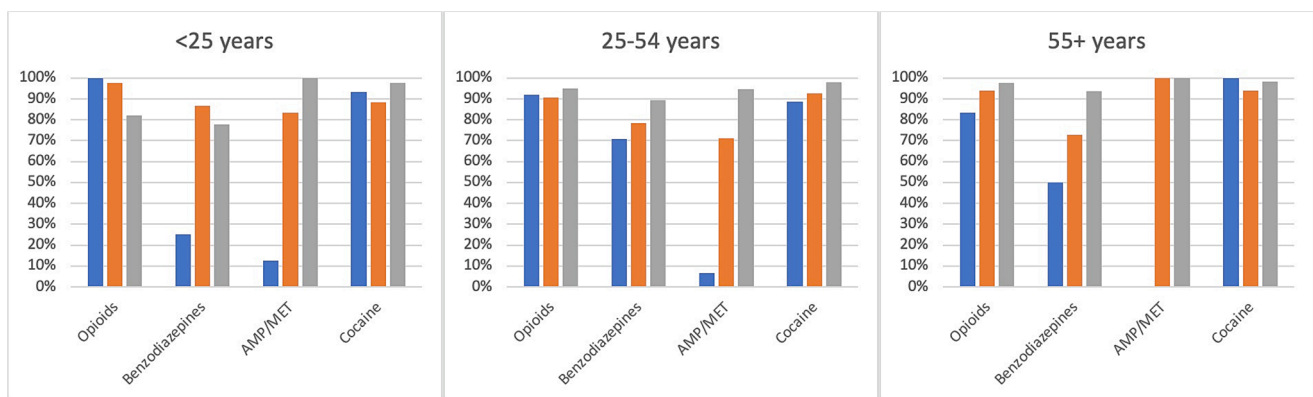
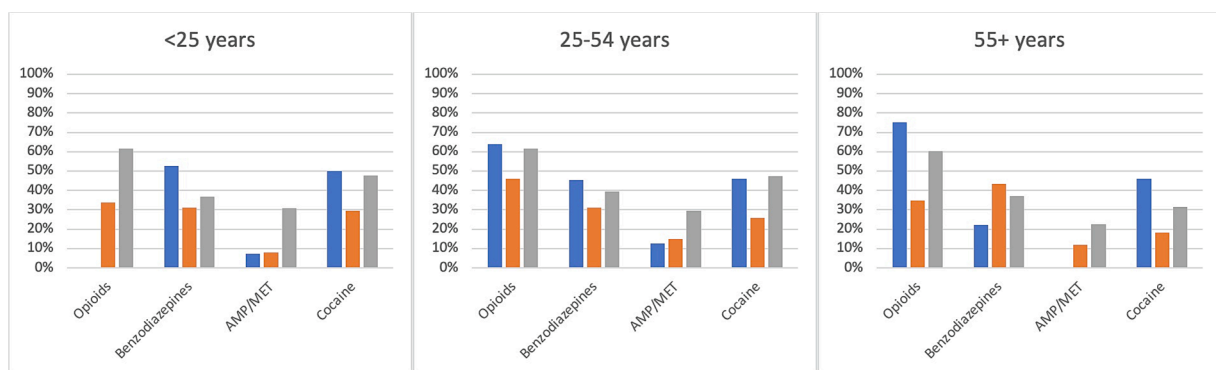


Fig. 5. FOR by Sex.



Legend: Blue = Pilot (2011), Orange = GENOA (2013-2017), Grey = POST (2018-2022)
 *PPV for AMP/MET 55+ was not able to be calculated

Fig. 6. PPV by age.



Legend: Blue = Pilot (2011), Orange = GENOA (2013-2017), Grey = POST (2018-2022)
 *FOR for opioids <25 years and AMP/MET 55+ was 0

Fig. 7. FOR by age.

the clinical services in Canada are focused on harm reduction, meaning participants would not be removed from treatment due to ongoing substance use, where the study was conducted (Morin et al., 2022).

Although there were some observations of different increase and decrease in the agreements between self-report and urine drug screen for the different drugs, study phases, age and sex that were explored we cannot conclude a specific trend by sex or age for PPV or FOR. However, we did find an increase in PPV (the proportion of self-reporting drug use and having a positive UDS) over time during the different phases of the

study and a higher FOR (the proportion of self-reporting no drug use and having a positive UDS) for the pilot and POST phase compared to the GENOA phase. We also found that specificity was consistently higher in the POST and GENOA phase when compared to the pilot study phase, meaning that the probability of self-reporting no drug use and having a negative UDS was highest in the latest study phase than the earlier phase. In general, we found the sensitivity and specificity were highest for cocaine and lowest for benzodiazepines. We found no specific trend in sex difference for any drug. For trends in age, we found a higher

probability of self-reporting drug use and having a positive UDS (sensitivity) for opioids and AMP/MET in participants < 25 years old.

Interestingly, the rate of FOR for opioids, cocaine and benzodiazepines were each over 38 %. This is clinically important as the concurrent use of drugs may further increase the risk of death when initiating treatment (The College of Physicians and Surgeons of Ontario, 2023) while negative UDS encourages take-home doses of MOUD and facilitate patients' self-management (Morin et al., 2022). Ongoing opioid and benzodiazepine use may impact self-management and increase the risk of overdose and/or respiratory depression when opioids and benzodiazepines are taken concurrently (Centre for Addiction and Mental Health, 2021). It is understandable however that, patients may be motivated to under-report illicit opioid and benzodiazepine use or decline UDS to increase/keep their take home medications. Similar to our study, others found the under-reporting cocaine use while on buprenorphine/naloxone treatment (Wilcox et al., 2013).

It is important to highlight the increase in PPV from the pilot phase (2011) to the POST phase (2018–2022). FOR was higher in the pilot and POST phases than during GENOA (2013–2017). In general, we see an increase over time in participants self-reporting drug use and having a corresponding positive UDS, suggesting an increase in accurate disclosure of substance use. However, the FOR was higher for the pilot (2011) and POST (2018–2022) phases compared to the GENOA (2013–2017) for all substances aside from AMP/MET, which increase from the pilot phase to the POST phase. As FOR accounts for when participants disclose no drug use and have a positive UDS, the observed change over time is concerning due to the impact additional substance use can have on serious adverse events such as overdose or death (Centre for Addiction and Mental Health, 2021). We may argue that the GENOA time in this context was unique, the opioid crisis was declared after the pilot study, it is possible that awareness was increased in that period, while FOR increased again in POST study time which coincided with increased toxic drug supply on the street and COVID-19 pandemic which was reported to be associated with increase substance use and associated harms (The Government of Canada, 2022).

Lastly, in general we found that the probability of having self-reported no drug use and having a negative UDS (specificity) for the pilot phase was lower than GENOA and POST. The PPV for AMP/MET use in the pilot study was considerably lower 7.9% compared to GENOA 75.0 % and POST 95.2 %. This may be related to the increase in AMP/MET use in the recent years compared to the pilot phase time (Canadian Centre on Substance Use and Addiction, 2020).

Our findings contribute to the body of knowledge and debate related to the use of routine drug screen and reliability of self-reported substance use. A recent review concluded a satisfactory level of agreement between self-report and biological testing for substance use. However, the included studies were related to various drugs as well as controlled studies and with plans to enhance accuracy for self-reporting, for example by letting people know they will be tested (Bharat et al., 2023). In previous studies women constituted 32 % of the population while our study included 44.4 % women demonstrating the changes in opioid use by women overtime. Our study also differs from the studies included in the systematic review in many aspects including a large sample (n = 4181) of pragmatic design, in a clinical setting with consequences for substance use, and all participants were receiving treatment for opioid use disorder. Most importantly, our study differs in that urine drug screens, despite past reports including the recent systematic review, remain a required clinical test in opioid use disorder treatment settings adding to the cost and intensity of treatment among other.

We acknowledge that there may be a value for the use of UDS as reported in some studies related to treatment retention (McEachern et al., 2019; Morin et al., 2022). Balancing the benefits and the risks of any test/intervention in clinical practice should take into consideration patients' safety, preference, efficacy, and timeliness to provide optimal care (McEachern et al., 2019; Morin et al., 2022).

5. Limitations

Self-reported substance use was collected using two separate tools (the MINI for pilot and MAP for GENOA and POST) which use different timeframes to derive patients self-reported substance use (previous 12 months for the MINI and previous 30 days for the MAP).

While UDS were checked on average once a week, it is possible that participants could have used a substance days before providing a urine sample leaving small traces of the drug in their urine that was inadequate to produce a positive test result rendering the UDS negative. Further, participants may report using substances in the past year (Pilot) or last month (GENOA, POST) but stopped using the substance prior to UDS leading to discordant findings between self-report and UDS.

Additionally, it is possible that participants obtaining illicit substances were not aware of the various substances potentially mixed with the substance of choice leading to a positive UDS for an additional substance that they did not self-report. Further, individual variability in metabolism of various substances can impact the detection rate of UDS and therefore UDSs do not always accurately reflect patients' actual substance use (Moeller et al., 2017; Owen et al., 2012).

6. Conclusion

In summary, our study identified an increase in the proportion of participants over time between 2011 and 2022 who accurately self-reported drug use and had a positive UDS, while a proportion also denied use while screening positive on urine test. We further identified an increase in this trend over time in the different phases of the study from 2011 to 2022. It is therefore an important clinical consideration to take the context of the individual (e.g. ongoing substance use or continued self-reporting of substance use) and the type of substances used (e.g. opioids and benzodiazepines which may lead to respiratory depression) when treating patients with OUD. The indications of UDS should be patient-centred, considering the patient's context and preferences, as well as clinical judgement, and not routinely done.

CRediT authorship contribution statement

Alannah McEvoy: Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Conceptualization. **Myanca Rodrigues:** Writing – review & editing, Methodology, Formal analysis, Conceptualization. **Brittany B Dennis:** Writing – review & editing. **Jacqueline Hudson:** Writing – review & editing, Project administration, Investigation. **David C. Marsh:** Writing – review & editing. **Andrew Worster:** Writing – review & editing. **Lehana Thabane:** Writing – review & editing. **Zainab Samaan:** Writing – review & editing, Supervision, Project administration, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.abrep.2024.100575>.

Data availability

The data that has been used is confidential.

References

- Akobeng, A. K. (2007). Understanding diagnostic tests 1: Sensitivity, specificity and predictive values. *Acta Paediatrica*, *96*(3), 338–341.
- Bharat, C., Webb, P., Wilkinson, Z., McKetin, R., Grebely, J., Farrell, M., Holland, A., Hickman, M., Tran, L. T., & Clark, B. (2023). Agreement between self-reported illicit drug use and biological samples: A systematic review and meta-analysis. *Addiction*, *118*(9), 1624–1648.
- Canadian Centre on Substance Use and Addiction (2020). Methamphetamine.
- Centre for Addiction and Mental Health (2021). Opioid Agonist Therapy: A Synthesis of Canadian Guidelines for Treating Opioid Use Disorder. <https://www.camh.ca/-/media/files/professionals/canadian-opioid-use-disorder-guideline2021-pdf.pdf>.
- Coffa, D., & Snyder, H. (2019). Opioid use disorder: Medical treatment options. *American Family Physician*, *100*(7), 416–425.
- Compton, W. M., Valentino, R. J., & DuPont, R. L. (2021). Polysubstance use in the US opioid crisis. *Molecular Psychiatry*, *26*(1), 41–50.
- Kost, G. J. (2021). The impact of increasing disease prevalence, false omissions, and diagnostic uncertainty on coronavirus disease 2019 (COVID-19) test performance. *Archives of Pathology & Laboratory Medicine*, *145*(7), 797–813.
- Krebs, E., Min, J. E., Zhou, H., Davison, C., McGowan, G., & Nosyk, B. (2021). The cascade of care for opioid use disorder among youth in British Columbia, 2018. *Journal of Substance Abuse Treatment*, *130*, Article 108404.
- Marsden, J., Gossop, M., Stewart, D., Best, D., Farrell, M., Lehmann, P., Edwards, C., & Strang, J. (2002). The Maudsley Addiction Profile (MAP): A brief instrument for assessing treatment outcome. *Addiction*, *93*(12). <https://doi.org/10.1046/j.1360-0443.1998.9312185711.x>.
- Mattick, R. P., Breen, C., Kimber, J., & Davoli, M. (2002). Methadone maintenance therapy versus no opioid replacement therapy for opioid dependence. *Cochrane Database of Systematic Reviews (Online)*, *4*. <https://www.cochranelibrary.com/central/doi/10.1002/central/CN-01756756/full>.
- McDonnell, M. G., Graves, M. C., West, I. I., Ries, R. K., Donovan, D. M., Bumgardner, K., Krupski, A., Dunn, C., Maynard, C., & Atkins, D. C. (2016). Utility of point of care urine drug tests in the treatment of primary care patients with drug use disorders. *Journal of Addiction Medicine*, *10*(3), 196.
- McEachern, J., Adaye-White, L., Priest, K. C., Moss, E., Gorfinkel, L., Wood, E., Cullen, W., & Klimas, J. (2019). Lacking evidence for the association between frequent urine drug screening and health outcomes of persons on opioid agonist therapy. *International Journal of Drug Policy*, *64*, 30–33.
- Moeller, K. E., Kissack, J. C., Atayee, R. S., & Lee, K. C. (2017). Clinical interpretation of urine drug tests: What clinicians need to know about urine drug screens. *Mayo Clinic Proceedings*, *92*(5), 774–796.
- Moeller, K. E., Lee, K. C., & Kissack, J. C. (2008). Urine drug screening: Practical guide for clinicians. *Mayo Clinic Proceedings*, *83*(1), 66–76.
- Morin, K. A., Dabous, J. R., Vojtesek, F., & Marsh, D. (2022). Evaluating the association between urine drug screening frequency and retention in opioid agonist treatment in Ontario, Canada: A retrospective cohort study. *BMJ Open*, *12*(10), e060857.
- Morin, K. A., Eibl, J. K., Franklyn, A. M., & Marsh, D. C. (2017). The opioid crisis: Past, present and future policy climate in Ontario, Canada. *Substance Abuse Treatment, Prevention, and Policy*, *12*(1), 1–7.
- Oviedo-Joekes, E., Brissette, S., Marsh, D. C., Lauzon, P., Guh, D., Anis, A., & Schechter, M. T. (2009). Diacetylmorphine versus Methadone for the Treatment of Opioid Addiction. *New England Journal of Medicine*, *361*(8), 777–786. <https://doi.org/10.1056/NEJMoa0810635>
- Owen, G. T., Burton, A. W., & Passik, S. (2012). Urine drug testing: Current recommendations and best practices. *Pain Physician*, *15*(3S), ES119.
- Pergolizzi, J., Pappagallo, M., Stauffer, J., Gharibo, C., Fortner, N., De Jesus, M. N., Brennan, M. J., Richmond, C., Hussey, D., & (IDCSG), I. D. C. S. G. (2010). The role of urine drug testing for patients on opioid therapy. *Pain Practice*, *10*(6), 497–507.
- Rieb, L. M., Samaan, Z., Furlan, A. D., Rabheru, K., Feldman, S., Hung, L., Budd, G., & Coleman, D. (2020). Canadian guidelines on opioid use disorder among older adults. *Canadian Geriatrics Journal*, *23*(1), 123.
- Ruglass, L. M., Shevorykin, A., Zhao, Y., Killen, T. K., Bauer, A. G., Morgan-López, A. A., Back, S. E., Fitzpatrick, S., López-Castro, T., & Norman, S. B. (2023). Self-report and urine drug screen concordance among women with co-occurring PTSD and substance use disorders participating in a clinical trial: Impact of drug type and participant characteristics. *Drug and Alcohol Dependence*.
- Schuler, M. S., Lechner, W. V., Carter, R. E., & Malcolm, R. (2009). Temporal and gender trends in concordance of urine drug screens and self-reported use in cocaine treatment studies. *Journal of Addiction Medicine*, *3*(4), 211.
- Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., & Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (MINI): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry*, *59*(20), 22–33.
- StataCorp. (2021). *Stata Statistical Software: Release 17*. College Station, TX: StataCorp LLC.
- Strike, C., & Rufo, C. (2010). Embarrassing, degrading, or beneficial: Patient and staff perspectives on urine drug testing in methadone maintenance treatment. *Journal of Substance Use*, *15*(5), 303–312.
- The College of Physicians and Surgeons of Ontario (2023). Advice to the profession: Prescribing Drugs.
- The Government of Canada (2022). Opioid- and Stimulant-related Harms in Canada.
- Trevethan, R. (2017). Sensitivity, specificity, and predictive values: Foundations, pliabilitys, and pitfalls in research and practice. *Frontiers in Public Health*, *5*, 307.
- Trimed Supply Network Ltd. (n.d.). FaStep Assay. <https://www.trimed-inc.com/wp-content/uploads/2018/04/Fastep-Package-Insert.pdf>.
- Wilcox, C. E., Bogenschutz, M. P., Nakazawa, M., & Woody, G. (2013). Concordance between self-report and urine drug screen data in adolescent opioid dependent clinical trial participants. *Addictive Behaviors*, *38*(10), 2568–2574.