Title

Screening emergency department patients for opioid drug use: A qualitative systematic review

Authors

Sahota, Preet Kaur
Shastry, Siri
Mukamel, Dana B
et al.

Publication Date

2018-10-01

DOI

10.1016/j.addbeh.2018.05.022

Peer reviewed

Screening emergency department patients for opioid drug use: A qualitative systematic review ☆

Preet Kaur Sahota[a], Siri Shastry[a], Dana B. Mukamel[b], Linda Murphy[c], Narisu Yang[d], Shahram Lotfipour[a], Bharath Chakravarthy[a],

[a] Department of Emergency Medicine, University of California, Irvine, Orange, CA, USA
[b] Department of Medicine, University of California, Irvine, Irvine, CA, USA
[c] Reference Department, University of California, Libraries, Irvine, CA, USA
[d] Department of Emergency Medicine, University College Hospital, London, UK

Highlights

- There is a lack of validity and reliability data for opioid screening instruments.
- There is no clear evidence to state which instruments are appropriate for use.
- There is a need for reliable and valid opioid screening instruments in EM settings.

A B S T R A C T

Introduction: The opioid drug epidemic is a major public health concern and an economic burden in the United States. The purpose of this systematic review is to assess the reliability and validity of screening instruments used in emergency medicine settings to detect opioid use in patients and to assess psychometric data for each screening instrument. Methods: PubMed/MEDLINE, PsycINFO, Cochrane Database of Systematic Reviews, Cochrane Central Register of Controlled Trials, Web of Science, Cumulative Index to Nursing and Allied Health Literature and ClinicalTrials.gov were searched for articles published up to May 2018. The extracted articles were independently screened for eligibility by two reviewers. We extracted 1555 articles for initial screening and 95 articles were assessed for full-text eligibility. Six articles were extracted from the full-text assessment. Results: Six instruments were identified from the final article list: Screener and Opioid Assessment for Patients with Pain - Revised; Drug Abuse Screening Test; Opioid Risk Tool; Current Opioid Misuse Measure; an Emergency Medicine Providers Clinician Assessment Questionnaire; and an Emergency Provider Impression Data Collection Form. Screening instrument characteristics, and reliability and validity data were extracted from the six studies. A meta-analysis was not conducted due to heterogeneity between the studies. Conclusions: There is a lack of validity and reliability evidence in all six articles; and sensitivity, specificity and predictive values varied between the different instruments. These instruments cannot be validated for use in emergency medicine settings. There is no clear evidence to state which screening instruments are appropriate for use in detecting opioid use disorders in emergency medicine patients. There is a need for brief, reliable, valid and feasible opioid use screening instruments in the emergency medicine setting.

1. Introduction

At the beginning of 1999, the opioid drug epidemic began to disperse across the United States (U.S.) and it persists, despite efforts to end this epidemic from spreading further.

The Centers for Disease Control and Prevention (CDC) report that opioids such as prescription opioids, heroin and fentanyl killed over 42,000 individuals in 2016 (Centers for Disease Control and Prevention, 2018a). Furthermore, in 2016, there were 32,445 deaths involving prescription opioids specifically. This is almost a 10,000 death increase in the span of a year as 22,598 deaths involving prescription opioids were reported in 2015 (Centers for Disease Control and Prevention, 2018b). In 2016, the rate of overdose of prescription opioids in men was 6.2 and the rate in women was 4.3 (Centers for Disease Control and Prevention, 2018c). The number of drug overdose deaths exceeds alcohol use and motor vehicle traffic-related deaths, illustrating the severity and concern of drug overdose in the U.S (Centers for Disease Control and Prevention, 2017a; Centers for Disease Control and Prevention, 2017b).

As defined by the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, the terms substance abuse and dependence have been replaced with the expression "substance use disorders," categorized on a scale from mild to severe. The Substance Abuse and Mental Health Services Administration describes concrete symptoms of opioid use disorder, such as "strong desire[s] for opioids, inability to control or reduce use, continued use despite interference with major obligations or social functions, use of larger amounts over time, development of tolerance, spending a great deal of time to obtain and use opioids" (Substance Abuse and Mental Health Services Administration, 2017)… as well as adverse health outcomes. During periods of attempted withdrawals, individuals may experience changes in mood or behavior, nausea, vomiting, fatigue, fever and insomnia, to name a few (Substance Abuse and Mental Health Services Administration, 2017).

The rise in opioid use disorders reinforces the severity of the opioid drug epidemic, as well as the undue burdens placed on individuals and the American healthcare system (Substance Abuse and Mental Health Services Administration, 2017). A retrospective study reported on the economic burden of prescription opioid use and misuse in 2013. About $78.5 billion of total U.S. economic burden was accredited to prescription opioid misuse. About one-third ($28.9 billion) was spent on healthcare and substance abuse treatments (Florence, Zhou, Luo, & Xu, 2016).

The rate of opioid prescribing in the U.S. plateaued between 2010 and 2012 (Compton, Jones, & Baldwin, 2016). Despite the decline in prescription opioid abuse, there are surges of heroin use and overdose deaths. In 2016, about 475,000 people ages 12 and older were classified as current heroin users. This corresponds to about 0.2% of the population ages 12 or older (Substance Abuse and Mental Health Services Administration, 2018). According to the 2016 National Survey on Drug Use and Health report, "The percentage of young adults [ages 18 to 25 years] in 2016 who were current heroin users (0.3%) was higher than the percentages in 2002 through 2004, and it was similar to the percentages in 2005 through 2015." (Substance Abuse and Mental Health Services Administration, 2018) Nonmedical prescription opioid use is associated as a risk factor for future heroin use, although the transition from nonmedical prescription opioids to heroin is rare and occurs at a steady rate (Compton et al., 2016).

As of 2008, the U.S. Preventive Services Task Force (USPSTF) determined "that the current evidence is insufficient to assess the balance of benefits and harms of screening adolescents, adults, and pregnant women for illicit drug use." (United States Preventive Services Task

Force, 2017) The USPSTF conducted a systematic review to identify validated screening instruments for the detection of drug misuse in ambulatory general medical settings and found "the evidence is not sufficient, however, to establish the positive predictive value of these tests when used in a general medical patient population with a predictably lower prevalence of drug use/misuse. The available evidence does not permit one to determine the overall clinical utility of these instruments when applied in a busy primary care practice setting, and especially in screening pregnant women for drug use." (Lanier & Ko, 2008; United States Preventive Services Task Force, 2017).

There is an urgent need to identify possible screening instruments for illicit drug use, specifically opioid use disorders, in various patient populations. Screening instruments provide healthcare providers with information in order to disseminate resources to patients who are at risk for substance use disorders. Screening instruments are available to almost all patient populations, in several clinical settings, and there are different forms of screening tools available, including questionnaires/instruments (Center for Substance Abuse Treatment, 2009; Substance Abuse and Mental Health Services Administration-Health Resources and Services Administration Center for Integrated Health Solutions, 2017).

Screening instruments have been validated for other health concerns, conditions and disorders, including alcohol use disorders (AUDs) and intimate partner violence, in emergency medicine (EM) settings (Feldhaus et al., 1997; Jones, 2011). These previous studies illustrate that it is possible to screen for multiple health conditions in EM settings. In particular, EM settings require time-sensitive screening instruments due to the fast-paced nature of the clinical environment and patient volume.

Furthermore, previous literature has depicted associations between emergency department use and drug overdose, as one study found a strong association between ED visits and the risk of subsequent prescription drug overdose deaths (Brady et al., 2015). Additionally the CDC found that from July 2016 to September 2017, 142,557 ED visits (15.7 per 10,000 visits), in 45 states, were suspected opioid overdoses. Rates increased in demographic groups and in five U.S. regions. In 16 states, 119,198 ED visits (26.7 per 10,000 visits) were suspected opioid involved overdoses (Vivolo-Kantor et al., 2018). As a result, the CDC includes an implication for public health practice statement: "Educating ED physicians and staff members about appropriate services for immediate care and treatment and implementing a post-overdose protocol that includes naloxone provision and linking persons into treatment could assist EDs with preventing overdose." (Vivolo-Kantor et al., 2018) From these results, it is evidence that EM settings play a significant role in the public health response to the opioid epidemic and because pain related cases are common in the EM setting, it is necessary for EM providers to monitor patients for possible opioid use disorders (Cordell et al., 2002).

The objective of this qualitative systematic review is to analyze existing literature and provide comprehensive psychometric evidence concerning the use of screening instruments in EM settings to detect opioid use in patients. We believe that shorter, highly-reliable and validated screening instruments will provide accurate data regarding opioid use, and serve as the best options for screening instruments to use in EM settings. Due to contamination, mislabeling and lost sampling of toxicology screening tests, we will not analyze invasive screening tests such as blood, urine and saliva sampling (Beck et al., 2014; Lanier & Ko, 2008; O'Neal & Poklis, 1998). We aim to provide reliability

and validity psychometric evidence for screening instruments that can detect opioid use in fast-paced EM settings to contribute to the existing literature in this field. We hope to provide EM physicians and clinicians with information concerning which screening instruments they can utilize to screen EM patients for possible opioid use disorder patterns and provide immediate interventions and educational programs for these patients.

2. Materials and methods

2.1. Literature search strategies

We systematically searched the following databases: PubMed/MEDLINE, PsycINFO, Cochrane Database of Systematic Reviews, Web of Science, Cumulative Index to Nursing and Allied Health Literature (CINAHL) and Cochrane Central Register of Controlled Trials (CENTRAL) from their inception dates to May 2018. We developed the optimal search strategy in PubMed and applied filters to restrict the final search results. The search strategy was moderately modified for use in the other selected databases. The search strategies for each database are depicted in Appendix A (Appendix A.).

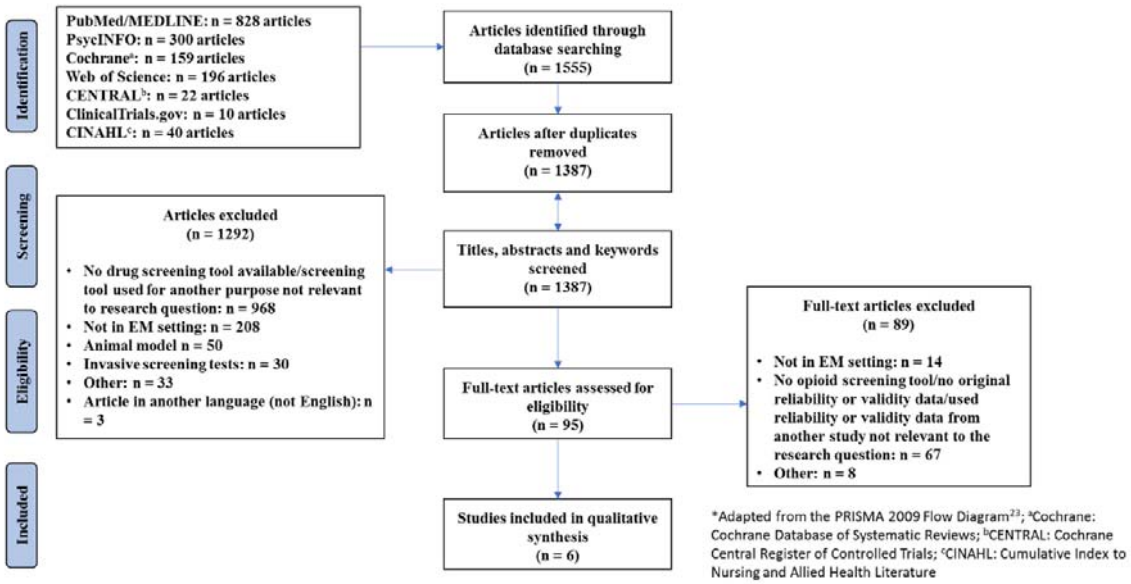

Fig. 1. Systematic literature review methodology flow diagram.

2.1.1. Searching additional resources

We screened the references from included studies to find other, potentially relevant studies that might be missed by the online literature searches. We also searched ClinicalTrials.gov to identify ongoing or recently completed trials. A health sciences librarian (LM) assisted with the design of the search strategy.

2.2. Study selection

Initially, two reviewers (PS and SS) independently screened the titles and abstracts for eligible studies using the established inclusion/exclusion criteria. Relevant abstracts were selected for full-text article eligibility and assessment. Differences in opinion were resolved independently by a third reviewer (BC). The flowchart of the study selection process is outlined in Fig. 1.

2.2.1. Inclusion/exclusion criteria

The inclusion criteria contain the following requirements: instruments must be history and/or questionnaire-based screening tools; instruments must take<10 min to perform and score (Mdege & Lang, 2011); no formal training is required to use or interpret the screening tool (as determined previously) (Mdege & Lang, 2011); all screening tools used to screen for the use of opioid drugs in patients; only English language studies (no funding for translational services); all age groups; all study types (including peer-reviewed articles, conference abstracts and all study designs); all locations within and out of the U.S.; no restrictions on publication dates; studies only conducted in EM settings; studies that assess at least one reliability and/or validity psychometric variable of the instrument; and studies that may include sensitivity, specificity, positive predictive value and negative predictive value information (as secondary analyses).

Reliability and validity psychometric data variables of interest are primarily the following: reliability (inter-rater reliability, test-retest reliability, Cronbach's alpha-internal consistency) and validity (face, convergent, divergent, construct, discriminative). Additional reliability and validity psychometric variables will be reported, if available.

The exclusion criteria contain the following requirements: studies which assess invasive screening (urine, plasma, saliva testing), non-English instruments; non-English studies; exclude studies which focus on alcohol screening only; and exclude books.

Fig. 1 details how many articles were excluded by each exclusion criterion by PS. There were few articles screened that were borderline/unclear or did not completely fit the inclusion criteria of the review. We excluded these studies, because they contained reliability and/or validity data established in medical settings outside of EM and/or were determined in a previous study that was cited by the article. PS, SS and BC excluded 89 articles after assessing the full-text of the 95 remaining articles and obtained a final list of six articles for qualitative data analysis (Fig. 1).

2.3. Data extraction and analysis

Our primary outcome variables of interest were any reliability and validity data presented in the six articles chosen for final review. This included test-retest reliability, inter-rater reliability (kappa), internal consistency (Cronbach's alpha) reliability testing, and validity testing (face, convergent, divergent, construct, discriminative). Other, less common, reliability and validity psychometric variables were also considered for analysis, if available in the studies.

We extracted the following secondary outcome variables if the data was available in the literature: study identifiers (i.e. study authors, publication years); screening instrument characteristics (i.e. self-administered/clinician-administered); study reference standards and secondary descriptive statistics (specificity, sensitivity and predictive value

data). All statistics were analyzed descriptively and presented qualitatively due to the complexities of the data points and statistics involved with quantitative review. Differences in opinion, in regards to the independently extracted data, were discussed and resolved between the three reviewers. All data was extracted into an Excel spreadsheet.

2.4. Ethical considerations

This review is a shortened revision of a Master of Science thesis (Sahota, 2017). This review follows Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines with the exception of the risk of bias evaluation (Preferred Reporting Items for Systematic Reviews and Meta-Analyses, 2017).

3. Results

3.1. Screening instrument characteristics

Appendix B contains study identifier information for all six studies (Appendix B.). The studies will be referred to by their "Study Number" to maintain consistency throughout the review. The following six screening instruments were identified for this review: Screener and Opioid Assessment for Patients with Pain - Revised (SOAPP-R); Drug Abuse Screening Test (DAST-20); Opioid Risk Tool (ORT); Current Opioid Misuse Measure (COMM); an Emergency Medicine Providers (EMPs) Clinician Assessment Questionnaire; and an Emergency Provider Impression Data Collection Form (Chalmers, Wilson, Mullinax, & Brennan, 2016; Kim, Hoppe, Kiemele, & Weiner, 2017; Reyes-Gibby, Anderson, & Todd, 2016; Varney et al., 2013; Weiner et al., 2013; Weiner, Horton, Green, & Butler, 2016).

The longest screening instrument, for which data are available, is the SOAPP-R. The SOAPP-R instrument contains 24 items. The six studies do not report the time frame for completion of the instruments; however, information presented in the articles implicitly report that all assessments take approximately 10 min to complete (all instruments are brief screening instruments or questionnaires).

All reported instruments screen for varying use of opioids in patients, and the SOAPP-R and the Emergency Provider Impression Data Collection Form also screen for other illicit/abuse drugs and drug seeking behavior, respectively (Appendix C.).

All six screening instruments screen for opioid drug use; however, the time frames for use differ substantially. The SOAPP-R screens for drug use within seven days prior to screening and over the patient's lifetime. DAST-20 screens for recent drug use, within seven days as reported by Study 1. ORT screens for recent drug use, within seven days as reported by Study 1 and within three months as reported by Study 4. The COMM can screen for drug use within three months as well. The EMPs Clinician Assessment Questionnaire assesses drug use "at least within 90 days prior to visit", as indicated in Study 5 (Varney et al., 2013). The Emergency Provider Impression Data Collection Form screens for drug use within 12 months prior to the administration of the instrument.

Only the SOAPP-R, ORT and COMM instruments are self-administered, as reported in Studies 3 and 4. The EMPs Clinician Assessment Questionnaire and Emergency Provider Impression Data Collection Form were completed by the clinicians

and providers in each study, not by the patients. In Study 5, the patients completed COMM; however, the COMM data was used as a reference standard to assess the EMPs Clinician Assessment Questionnaire.

Each study in this review, with the exception of Study 2, uses a reference standard for comparison with the instrument of interest. Studies 1, 3 and 6 use prescription drug monitoring program (PDMP) data as the reference standard. The PDMP criteria reference standard contains the following: ≥4 opioid prescriptions and ≥4 different prescribers reported, for the prior 12 months, for patients. Similarly, Study 4 uses nine aberrant patient behaviors as the reference standard. These behaviors include "soliciting opiate prescriptions from 3 or more physicians, forging or selling prescriptions, or an abnormal urine drug screen" (Chalmers et al., 2016). These behaviors were retrieved from patient medical charts and the Controlled Substance Utilization Review and Evaluation System. COMM was used as a reference standard in Study 6 (Appendix D.).

3.2. Study characteristics

Study characteristics are presented in Table 1 (Table 1). The majority of the studies are prospective, cross-sectional or observational studies. Most of the studies survey a convenience sample of patients. Studies 5 and 6 include assessments from EM clinicians and providers. Most studies were conducted within trauma and/or academic EM settings. Study 2 was conducted at a comprehensive cancer center ED and Study 5 was conducted at a military tertiary care hospital ED. All studies were conducted in the U.S. The study populations were mostly composed of adult ED patients, the majority of whom presented to the EM setting with pain symptomatology. There is a wide range of sample sizes reported between the six studies. At the patient level, the sample sizes range from as low as 74 patients (Study 5) to 539 patient encounters (Study 6). At the clinician level, the sample sizes range from 38 clinicians (Study 6) to 70 clinicians (Study 5).

Table 1
Study characteristics.

Study number	Design	Population	Setting (location)/country	Sample size
1	Observational study, convenience sample	Adult ED patients (≥ 18 years old) with a painful condition	Two academic hospitals/USA	n = 121
2	Cross-sectional, single-center	Cancer ED patients	Comprehensive cancer center ED/USA	n = 209
3	Cross-sectional, prospective, convenience sample	ED patients considered for discharge with a prescription for an opioid pain medication	Urban, academic Level 1 trauma center ED/USA	n = 82
4	Prospective observational study; convenience sample	Patients with chief complain of pain ≥ 6 months or with an opiate refill request	Academic ED/USA	n = 154
5	Observational study	ED patients (> 18 years) who received at least one opioid prescription 90 days prior; EM clinicians	Military tertiary care hospital ED/USA	n = 70 (clinicians); n = 74 (patients)
6	Prospective observational study; convenience sample	Emergency providers (attending emergency physicians and nurse practitioners with access to Massachusetts PDMP	Two urban academic Level 1 trauma centers/USA	n = 38 (539 patient encounters)

Abbreviations: ED: emergency department; EM: emergency medicine; PDMP: prescription drug monitoring program; USA: United States of America; N/A: data not available

Table 2
Validity and reliability data.

Study number	Test-retest reliability	Inter-rater reliability	Internal consistency (Cronbach's alpha)	Validity measures	Other reliability measures
1: SOAPP-R	N/A	N/A	N/A	"SOAPP-R has superior test characteristics for detecting aberrant opioid drug-related behavior by PDMP criteria when compared to DAST-20 and ORT." (**face validity**)	N/A
1: DAST-20	N/A	N/A	N/A	N/A	N/A
1: ORT	N/A	N/A	N/A	N/A	N/A
2: SOAPP-R	N/A	N/A	N/A	N/A	Reliability statistic: 0.786
3: SOAPP-R	N/A	N/A	0.91	Unadjusted ROC analysis: AUC of 0.64; adjusted analysis: AUC of 0.81	Item-total correlations: from 0.21 (question 21) to 0.71 (question 10)
4: ORT	N/A	N/A	N/A	Area under the ROC curve: 0.510 (95% CI: 0.416–0.604)	N/A
4: SOAPP-R	N/A	N/A	N/A	Area under the ROC curve: 0.579 (95% CI: 0.487–0.670)	N/A
4: COMM	N/A	N/A	N/A	Area under the ROC curve: 0.574 (95% CI: 0.483–0.665)	N/A
5: EMPs Clinician Assessment Questionnaire	N/A	N/A	N/A	"Clinician assessment is a comparable alternative to the COMM for measuring the potential for ongoing prescription opioid misuse among ED patients." (**face validity**)	N/A
6: Emergency provider impression data collection form	N/A	κ statistic: 0.30 (95% CI: 0.22–0.39)	N/A	"This finding underscores that the results of our study may not be externally valid and that large differences likely exist between different patient populations." (**face validity**)	N/A

Abbreviations: PDMP: prescription drug monitoring program; AUC: area under the curve; ROC: receiver operating characteristic; CI: confidence interval; EMPs: emergency medicine providers; ED: emergency department; κ: kappa; N/A: data not available

3.3. Primary outcome data: validity and reliability

None of the six studies provide test-retest analyses. Only Study 3 reports on the internal consistency of SOAPP-R. Study 2 does not report on any validity testing; Studies 1, 4 and 5 do not report on any reliability testing. Study 2 states: "the reliability of SOAPP-R in our sample was 0.786" (Reyes-Gibby et al., 2016); however, we are not informed as to which reliability test was conducted. Study 3 reports on the item total correlations of individual questions in the SOAPP-R. Study 6 provides a κ (Kappa) statistic to calculate agreement between the Emergency Provider Impression Data Collection Form and PDMP reference standard (we classified the κ statistic as a measure of inter-rater reliability as κ statistics are traditionally used to assess inter-rater reliability; however, it is not explicitly stated in the article that this is an inter-rater reliability assessment). Studies 1, 5 and 6 provide qualitative analyses of their instruments' validity (which we characterized as face validity). These statements are included in Table 2 (Table 2.).

3.4. Secondary outcome data

The data in Table 3 is more extensive than the lack of reliability and validity data in Table 2; there are missing statistics for specificity, sensitivity and predictive value (both positive and negative) only for Study 2 (Table 3.). Specificity, sensitivity and predictive values provide information concerning the usefulness and utility of screening instruments in the chosen study environment. All studies, with the exception of Study 2, report specificity, sensitivity and positive predictive value data. There are no negative predictive values for Studies 1, 2, 4, and 6. We reported the values as percentages and/or decimals, depending on the formatting used in the articles and how the articles explicitly present the values in the text.

4. Discussion

4.1. Analysis of study characteristics

Four of the six studies reviewed employ the prospective and/or cross-sectional study design. The prospective, cross-sectional design is used for most survey-based studies and allows researchers to interact with patients and determine in a timely manner if opioid use is a concern for these patients (Center for Substance Abuse Treatment, 2009). Opioids are used for pain management in certain populations, such as cancer patients or patients with serious injuries. They are susceptible to opioid use disorders due to prescribed opioid use (NIH Medline Plus, 2017). These patients may be first seen in EM settings; which is why it is vital to screen EM patients for opioid use disorders.

Studies 3, 5 and 6 report small sample sizes (Studies 3 and 5 for patient sample sizes and Studies 5 and 6 for clinician sample sizes). Small sample sizes may not accurately represent the population at large. Larger sample sizes, typically>50–100 patients, are encouraged for these studies to ensure accurate results and reporting for various population groups that filter through EM settings (Lanier & Ko, 2008).

Table 3
Specificity, sensitivity and predictive value data.

Study number	Specificity	Sensitivity	Positive predictive value	Negative predictive value
1: SOAPP-R	85.0% (95% CI: 76.0–91.0%)	33.0% (95% CI: 17.0–54.0%)	39.0% (95% CI: 20.0–61.0%)	N/A
1: DAST-20	79.0% (95% CI: 69.0–86.0%)	10.0% (95% CI: 1.0–37.0%)	8.7% (95% CI: 1.1–28.0%)	N/A
1: ORT	82.0% (95% CI: 72.0–89.0%)	22.0% (95% CI: 9.0–40.0%)	30.0% (95% CI: 13.0–53.0%)	N/A
2: SOAPP-R	N/A	N/A	N/A	N/A
3: SOAPP-R	71.0% (95% CI: 58.8–81.3%)	53.9% (95% CI: 25.2–80.7%)	25.9% (95% CI: 11.2–46.3%)	89.1% (95% CI: 77.7–95.9%)
4: ORT	58.8%	38.4%	0.54	N/A
4: SOAPP-R	50.0%	60.5%	0.60	N/A
4: COMM	52.1%	50.6%	0.57	N/A
5: EMPs Clinician Assessment Questionnaire	0.81	0.33	0.44	0.73
6: Emergency Provider Impression Data Collection Form	72.7% (95% CI: 68.4–77.0%)	63.2% (95% CI: 54.8% to 71.7%)	41.2% (95% CI: 34.4% to 48.2%)	N/A

Abbreviations: CI: confidence interval; EMPs: emergency medicine providers; N/A: data not available

4.2. Validity and reliability: primary outcome data analysis

There is an overwhelming lack of psychometric data for almost all the screening instruments within this review. The internal consistency of SOAPP-R is high (Cronbach's alpha: 0.91) and indicates good reliability of the instrument in screening for opioid detection in EM settings, as indicated in Study 3. High AUC values, as shown for SOAPP-R in Study 3, indicate goodness of fit with the chosen reference standard and discriminative validity. For the adjusted analysis, the AUC value (0.81) for the high-risk SOAPP-R score of 18 was a good fit for the PDMP data reference standard. In contrast, the AUC values for SOAPP-R, ORT and COMM in Study 4 are low (0.579, 0.510 and 0.574 respectively) and do not indicate goodness of fit with the chosen reference standard and indicate poor discriminative validity. This introduces inconsistency between studies in regards to the validity of SOAPP-R.

Study 3 also reports on the item-total correlations of individual questions in the SOAPP-R instrument, indicating that the question items correlate well with each other and ask different questions related to drug use within the instrument. For Study 6, the κ statistic (0.30) indicates only minimal level of agreement between the Emergency Provider Impression Data Collection Form and PDMP reference standard. Within the Discussion section of Study 6, the authors indicate that: “However, κ the statistic (0.30) and relatively low positive predictive value (41.2%) of emergency providers indicates such a large disconnect that one of the 2, either prescription drug monitoring program or physicians, must be incorrect in a substantial proportion of assessments for drug-seeking behavior.” (Weiner et al., 2013) The remaining studies provide qualitative analyses of their respective instrument's validity and it is difficult to ascertain meaningful results without concrete data available to support these statements.

4.3. Secondary outcomes data analysis

Specificity, sensitivity and predictive values validate the use of screening instruments in comparison to the chosen reference standards and provide information concerning the usefulness and utility of the screening instruments in the designated study environment. The SOAPP-R (Study 1), ORT (Study 1), and EMPs Clinician Assessment Questionnaire all report high specificity values (81% - 85%). The DAST-20, SOAPP-R

(Study 3), ORT (Study 4), SOAPP-R (Study 4), COMM and Emergency Provider Impression Data Collection Form report moderate specificity values (50% - 79%).

The sensitivities of the instruments differ across the studies as well. There is poor sensitivity for the SOAPP-R (Study 1), DAST-20, ORT (Study 1), ORT (Study 4) and EMPs Clinician Assessment Questionnaire (10% - 38.4%). The remaining screening instruments (SOAPP-R (Study 3), SOAPP-R (Study 4), COMM and Emergency Provider Impression Data Collection Form) report moderate sensitivity values (50.6% - 63.2%).

Only the ORT (Study 4), SOAPP-R (Study 4) and COMM report moderate positive predictive values (54%, 60% and 57% respectively). The remaining screening instruments report low positive predictive values. The SOAPP-R (Study 3) and EMPs Clinician Assessment Questionnaire report moderate to high negative predictive values (89.1% and 73% respectively).

4.4. Limitations

The most prevalent limitation is the lack of screening instrument data and methodology in majority of the studies. Previously validated screening instruments do have reliability and validity data from studies conducted in primary care and pain management settings; however, these data do not necessarily extrapolate to EM settings (Finkelman et al., 2016; van der Westhuizen, Wyatt, Williams, Stein, & Sorsdahl, 2016; Weiner et al., 2016). Primary care and EM settings differ in their patient populations and delivery of healthcare. Due to the amount of pain-related cases presented to the ED in comparison to other medical specialties, it is vital to validate screening instruments through the use of pilot studies to ensure that the instruments truly serve their purpose in a new environment, with different patient populations (American College of Emergency Physicians, 2017; Centers for Medicare and Medicaid Services, 2017).

We did not rate the quality of each study as there are no pre-established grading scales to evaluate prospective, survey-based studies. We included studies that evaluated opioid use and possible additional illicit drug use. It may be more practical to screen for general illicit drug use in EM settings to capture more substance use disorder patients (including opioid use disorder) for intervention and treatment.

We may not have screened for all relevant studies due to limitations in the number of search terms and synonyms we were able to use, and barriers such as language differences and unpublished studies that do not have reportable data at this time for analysis (Denison et al., 2013).

4.4.1. Prevalence of Bias

For our review, the articles contain the following limitations: missing or unavailable data due to selective reporting and the use of reliability and validity data from previous studies without confirming use of the screening instrument in EM settings. Because we could not locate a systematic review of opioid use screening instruments used in EM settings, we used review protocols of other systematic reviews performed in different medical settings to guide the framework of our research (Lanier & Ko, 2008; Mdege & Lang, 2011; Tiet, Finney, & Moos, 2008).

For the data provided in the six studies, there was a lack of background information concerning calculations of specificity and sensitivity and how qualitative analyses were determined. The values of specificity, sensitivity and predictive values differed between the same instruments in different studies, indicating differences between the studies. These differences may include the study populations, how the EM settings are organized, administrative issues, the use of different reference standards and how physicians disseminate patient care in their respective EDs. There may be regional differences that can explain the results from this review as EDs differ region from region across the U.S.

Due to differences in reference standards and heterogeneity of data between the six studies, it is difficult to conduct any meta-analysis or quantitative review. Publication biases may exist as well as selective reporting in the six studies, which may affect the analyses.

Qualitative analyses are used for variety of reasons. Meta-analysis of data may be difficult to achieve due to quantitative heterogeneity between the articles. Approaches to research may have changed throughout time and to review all articles within the scope of the research topic, qualitative analyses may provide more accurate comparisons than quantitative syntheses. Additionally, qualitative analyses are typically used when reviewing measurement approaches, such as the assessment of screening instruments (University of Stirling, 2017).

4.5. Future directions

Extensive studies that test a majority of the psychometric properties of opioid use screening instruments in EM settings are needed. Self administered instruments can save time for the clinician and provide meaningful information for further treatments and interventions needed for at-risk populations (Center for Substance Abuse Treatment, 2009).

One example of a study conducted in EM settings aims to validate the use of the 31-item, 20-min Prescription Drug Use Questionnaire, Patient Version against the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) criteria for prescription opioid use disorders (Beaudoin, Merchant, & Clark, 2016). The DSM-5 criteria may serve as a better reference standard as this classification is a well-established classification amongst the medical community and "has widespread importance and influence on how disorders are diagnosed, treated, and investigated."(Hasin, O'Brien, Auriacombe, et al., 2013) However, as one study mentions, "We think such a study should be done to accurately make comparisons between DSM-IV and DSM-5 AUD [Alcohol Use Disorder]" (Hasin, Auriacombe, et al., 2013), indicating that further studies must verify the validity of DSM-5 as a gold/criterion standard for psychometric evaluations of substance use disorders.

Although it is impossible to include all reliability and validity testing in one study, clinicians should consider consulting statisticians/experts in the field of psychometrics to determine which tests would provide useful reliability and validity data to the medical community. As per this review, there is a need for studies that test test-retest, interrater and internal consistency reliability; and construct, convergent, divergent and discriminative validity (Mdege & Lang, 2011; Tiet et al., 2008). Clinicians should consider including most, if not all, of these reliability and validity tests in order to

provide more information concerning the screening instruments in circulation within the medical field and provide more credibility for the instruments that are used as reference standards and for various settings in which they have not yet been validated in (Sullivan, 2011).

5. Conclusions

The lack of validity and reliability data hinder the selection of appropriate screening instruments for use to detect opioid use disorders in EM settings. Of the screening instruments presented within this systematic review, the SOAPP-R provides the most amount of data and promise for use in EM settings; however, the lack of studies and inconsistencies of instrument data across studies indicate that none of the screening instruments presented in this study are suitable for all EM settings. Additionally, there is a lack of information concerning the utility of the screening instruments for use by the clinician to ensure timeliness within the fast-paced EM setting.

As a result of this review, we know that there is a lack of opioid screening instrument validation studies within EM settings. If this knowledge gap is filled, we can retrieve useful data to conduct screening in fast-paced settings and implement innovative instruments to ensure that all patients at-risk for opioid use disorders are captured before adverse health outcomes and further drug dependency manifest. Due to the rise in the rate of heroin overdose as a result of the opioid epidemic, it is vital to implement validated screening instruments in the EM setting and capture as many patients as possible for education, intervention and treatment.

Role of funding sources

There were no funding sources for this research.

Contributors

Preet Kaur Sahota conducted the study as a Master of Science Thesis and conducted the literature searches, served as a reviewer, conducted all literature data analyses and wrote the complete first draft of the manuscript. Dr. Bharath Chakravarthy served as the third reviewer for the review, and contributed to and approved the final manuscript. Dr. Siri Shastry was the second reviewer for the review and conducted the literature searches; Dr. Shastry approved the final manuscript. Dr. Dana B. Mukamel contributed to the study design, and contributed to and approved the final manuscript. Linda Murphy created and finalized the literature search strategy and terms, and contributed to and approved the final manuscript. Dr. Narisu Yang contributed to the study design and approved the final manuscript. Dr. Shahram Lotfipour contributed to the study design, and contributed to and approved the final manuscript. I verify that all authors have approved the final manuscript draft.

Conflict of interest

The authors have no conflicts of interest to report.

Acknowledgements

The authors have no acknowledgements to report.

Disclosure statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.addbeh.2018.05.022.

References

American College of Emergency Physicians. Urgent care fact sheet. (2017). http://newsroom.acep.org/fact_sheets?item=30033 (Accessed July 15, 2017).

Beaudoin, F. L., Merchant, R. C., & Clark, M. A. (2016). Prevalence and detection of prescription opioid misuse and prescription opioid use disorder among emergency department patients 50 years of age and older: Performance of the prescription drug use questionnaire, patient version. The American Journal of Geriatric Psychiatry, 24(8), 627–636. http://dx.doi.org/10.1016/j.jagp.2016.03.010.

Beck, O., Carlsson, S., Tusic, M., Olsson, R., Franzen, L., & Hulten, P. (2014). Laboratory and clinical evaluation of on-site urine drug testing. Scandinavian Journal of Clinical and Laboratory Investigation, 74(8), 681–686. http://dx.doi.org/10.3109/00365513.2014.939995.

Brady, J. E., DiMaggio, C. J., Keyes, K. M., Doyle, J. J., Richardson, L. D., & Li, G. (2015). Emergency department utilization and subsequent prescription drug overdose death. Annals of Epidemiology, 25(8), 613–619.e2. http://dx.doi.org/10.1016/j.annepidem.2015.03.018.

Center for Substance Abuse Treatment (2009). Substance abuse treatment: Addressing the specific needs of women. Treatment Improvement Protocol (TIP) Series, No. 51. Chapter 4: Screening and AssessmentRockville (MD): Substance Abuse and Mental Health Services Administration (US) https://www.ncbi.nlm.nih.gov/books/NBK83253/ (Accessed June 22, 2017).

Centers for Disease Control and Prevention (2017a). National center for health statistics: Alcohol use. https://www.cdc.gov/nchs/fastats/alcohol.htm (Accessed July 19, 2017).

Centers for Disease Control and Prevention (2017b). National center for health statistics: Accidents or unintentional injuries. https://www.cdc.gov/nchs/fastats/accidentalinjury.htm (Accessed July 19, 2017).

Centers for Disease Control and Prevention (2018a). Opioid overdose | drug overdose. https://www.cdc.gov/drugoverdose/index.html (Accessed May 12, 2018).

Centers for Disease Control and Prevention (2018b). Opioid data analysis | drug overdose. https://www.cdc.gov/drugoverdose/data/analysis.html (Accessed May 12, 2018).

Centers for Disease Control and Prevention (2018c). Prescription opioid overdose data | Drug overdose. https://www.cdc.gov/drugoverdose/data/overdose.html (Accessed May 12, 2018).

Centers for Medicare and Medicaid Services (2017). Differences between your provider's office and the emergency department. https://marketplace.cms.gov/outreach-and education/downloads/c2c-emergency-department.pdf (Accessed July 15, 2017).

Chalmers, C. E., Wilson, M. P., Mullinax, S. Z., & Brennan, J. J. (2016). 122 an Evaluation of screening tools to predict opioid misuse in the emergency department. Annals of Emergency Medicine, 68(4), S48. http://dx.doi.org/10.1016/j.annemergmed.2016.08.134.

Compton, W. M., Jones, C. M., & Baldwin, G. T. (2016). Relationship between nonmedical prescription-opioid use and heroin use. New England Journal of Medicine. 374(2), 154–163. http://dx.doi.org/10.1056/NEJMra1508490.

Cordell, W. H., Keene, K. K., Giles, B. K., Jones, J. B., Jones, J. H., & Brizendine, E. J. (2002). The high prevalence of pain in emergency medical care. The American Journal of Emergency Medicine, 20(3), 165–169. http://dx.doi.org/10.1053/ajem.2002.32643.

Denison, H. J., Dodds, R. M., Ntani, G., Cooper, R., Cooper, C., Sayer, A. A., & Baird, J. (2013). How to get started with a systematic review in epidemiology: An introductory guide for early career researchers. Arch Public Health. 71(1), 21. http://dx.doi.org/10.1186/0778-7367-71-21.

van der Westhuizen, C., Wyatt, G., Williams, J. K., Stein, D. J., & Sorsdahl, K. (2016). Validation of the alcohol, smoking and substance involvement screening test in a low and middle-income country cross-sectional emergency Centre study. Drug and Alcohol Review, 35(6), 702–709. http://dx.doi.org/10.1111/dar.12424.

Feldhaus, K. M., Koziol-McLain, J., Amsbury, H. L., Norton, L. M., Lowenstein, S. R., & Abbott, J. T. (1997). Accuracy of 3 brief screening questions for detectingpartner violence in the emergency department. Journal of the American

Medical Association, 277(17), 1357–1361. http://dx.doi.org/10.1001/jama.1997.03540410035027.

Finkelman, M. D., Kulich, R. J., Butler, S. F., Jackson, W. C., Friedman, F. D., Smits, N., & Weiner, S. G. (2016). An investigation of completion times on the screener and opioid assessment for patients with pain - revised (SOAPP-R). Journal of Pain Research, 9, 1163–1171.

Florence, C. S., Zhou, C., Luo, F., & Xu, L. (2016). The economic burden of prescription opioid overdose, abuse, and dependence in the United States, 2013. Medical Care, 54(10), 901–906. http://dx.doi.org/10.1097/MLR.0000000000000625.

Hasin, D. S., Auriacombe, M., Borges, G., Bucholz, K., Budney, A., Crowley, T., ... Wall, M. M. (2013). The DSM-5 field trials and reliability of alcohol use disorder. The American Journal of Psychiatry, 170(4), 442–443. http://dx.doi.org/10.1176/appi.ajp.2013. 13010032.

Hasin, D. S., O'Brien, C. P., Auriacombe, M., Borges, G., Bucholz, K., Budney, A., ... Grant, B. F. (2013). DSM-5 criteria for substance use disorders: Recommendations and rationale. The American Journal of Psychiatry, 170(8), 834–851. http://dx.doi.org/10.1176/appi.ajp.2013.12060782.

Jones, L. A. (2011). Systematic review of alcohol screening tools for use in the emergency department. Emergency Medicine Journal. 28, 182–191. http://dx.doi.org/10.1136/emj.2009.085324.

Kim, K., Hoppe, J. A., Kiemele, E. R., & Weiner, S. G. (2017). A comparison of three screening tools for aberrant opioid drug-related behavior in the emergency department. SAEM Annual Meeting Abstracts. Number 624. 24Acad Emerg Med. S7-S281 https://doi.org/10.1111/acem.13203.

Lanier, D., & Ko, S. (2008). Screening in primary care settings for illicit drug use: Assessment of screening instruments — A supplemental evidence update for the U.S. preventive services task force. Evidence synthesis no. 58, part 2. AHRQ publication no. 08-05108-EF-2. Rockville, Maryland: Agency for Healthcare Research and Quality.

Mdege, N. D., & Lang, J. (2011). Screening instruments for detecting illicit drug use/ abuse that could be useful in general hospital wards: A systematic review. Addictive Behaviors, 36(12), 1111–1119. http://dx.doi.org/10.1016/j.addbeh.2011.07.007.

NIH Medline Plus. Opioids and chronic pain. (2017). https://medlineplus.gov/magazine/ issues/spring11/articles/spring11pg9.html (Accessed June 22, 2017).

O'Neal, C. L., & Poklis, A. (1998). An evaluation of the role of ROC plots in the

Prediction of heroin use from total codeine and total morphine concentrations in urine. Journal of Analytical Toxicology, 22(6), 487–492.

Preferred Reporting Items for Systematic Reviews and Meta-Analyses. Welcome to the preferred reporting items for systematic reviews and meta-analyses (PRISMA). (2017). http://www.prisma-statement.org (Accessed July 04, 2017).

Reyes-Gibby, C. C., Anderson, K. O., & Todd, K. H. (2016). Risk for opioid misuse Among emergency department cancer patients. Academic Emergency Medicine, 23(2), 151–158. http://dx.doi.org/10.1111/acem.12861.

Sahota, P. K. (2017). An analysis of opioid-risk screening instruments for use in the Emergency department: A qualitative systematic review. Doctoral dissertation,Irvine: University of California.

Substance Abuse and Mental Health Services Administration. Substance use disorders. (2017). https://www.samhsa.gov/disorders/substance-use (Accessed May 22, 2017).

Substance Abuse and Mental Health Services Administration. Key substance use and mental health indicators in the United States: Results from the 2016 National Survey on Drug Use and Health. (2018). https://www.samhsa.gov/data/sites/default/files/NSDUH-FFR1-2016/NSDUH-FFR1-2016.htm#illicit9 (Accessed May 12, 2018).

Substance Abuse and Mental Health Services Administration-Health Resources and Services Administration Center for Integrated Health Solutions. Screening tools. (2017). http://www.integration.samhsa.gov/clinical-practice/screening-tools (Accessed May 20, 2017).

Sullivan, G. M. (2011). A primer on the validity of assessment instruments. Journal of Graduate Medical Education. 3(2), 119–120. http://dx.doi.org/10.4300/JGME-D-11-00075.1.

Tiet, Q. Q., Finney, J. W., & Moos, R. H. (2008). Screening psychiatric patients for illicit drug use disorders and problems. Clinical Psychology Review, 28(4), 578–591. http://dx.doi.org/10.1016/j.cpr.2007.08.002.

United States Preventive Services Task Force (2017). Drug use, illicit: Screening. https://www.uspreventiveservicestaskforce.org/Page/Document/UpdateSummaryFinal/drug-use-illicit-screening, Accessed date: 22 May 2017.

University of Stirling. What is a systematic literature review and how do i do one? https://www.stir.ac.uk/media/schools/management/documents/centregradresearch/How%20to%20do%20a%20systematic%20literature%20review%20and%20meta analysis.pdf. 2017 (Accessed July 04, 2017).

Varney, S. M., Bebarta, V. S., Ganem, V. J., Carey, K. R., Ramos, R. G., & Zarzabal, L. A. (2013). Ongoing opioid misuse identification in the emergency department using the COMM instrument compared to clinician judgment. Annals of Emergency Medicine, 62(4), S44–S45. http://dx.doi.org/10.1016/j.annemergmed.2013.07.404.

Vivolo-Kantor, A. M., Seth, P., Gladden, R. M., Mattson, C. L., Baldwin, G. T., ... Colettea, M. A. (2018). Vital signs: Trends in emergency department visits for suspected opioid overdoses — United States, july 2016–september 2017. MMWR. Morbidity and Mortality Weekly Report, 67, 279–285. http://dx.doi.org/10.15585/mmwr.mm6709e1.

Weiner, S. G., Griggs, C. A., Mitchell, P. M., Langlois, B. K., Friedman, F. D., Moore, R. L., Lin, S. C., ... Nelson, K. P., Feldman, J. A. (2013). Clinician impression versus prescription drug monitoring program criteria in the assessment of drug-seeking behavior in the emergency department. Annals of Emergency Medicine, 62(4), 281–289. http://dx.doi.org/10.1016/j.annemergmed.2013.05.025.

Weiner, S. G., Horton, L. C., Green, T. C., & Butler, S. F. (2016). A comparison of an Opioid abuse screening tool and prescription drug monitoring data in the emergency department. Drug and Alcohol Dependence, 159, 152–157. http://dx.doi.org/10.1016/j.drugalcdep.2015.12.007.

☆ Prior Presentations: Master of Science (MS) Thesis Defense Presentation.

* Corresponding author. E-mail addresses: psahota@uci.edu (P.K. Sahota), spshastr@uci.edu (S. Shastry), dmukamel@hs.uci.edu (D.B. Mukamel), lmurphy@uci.edu (L. Murphy), narisu.yang10@imperial.ac.uk (N. Yang), shl@uci.edu (S. Lotfipour), bchakrav@uci.edu (B. Chakravarthy).