OXFORD

OPIOIDS & SUBSTANCE USE DISORDERS SECTION

Original Research Article

Re-assessing the Validity of the Opioid Risk Tool in a Tertiary Academic Pain Management Center Population

Meredith R. Clark, DNP,* Robert W. Hurley, MD, PhD,[†] and Meredith C.B. Adams, MD, MS[†]

*Division of Pain Medicine, Department of Anesthesiology, Medical College of Wisconsin, Wauwatosa, Wisconsin; [†]Section of Pain Medicine, Department of Anesthesiology, Wake Forest School of Medicine, Medical Center Drive, Winston-Salem, North Carolina, USA

Correspondence to: Meredith C.B. Adams, MD, MS, Section of Pain Medicine, Department of Anesthesiology, Wake Forest School of Medicine, Medical Center Drive, Winston-Salem, NC 27157. Tel: 336-716-1228; Fax: 336-716-8190; E-mail: meradams@wakehealth.edu.

Abstract

Objective. To analyze the validity of the Opioid Risk Tool (ORT) in a large. diverse population.

Design. A cross-sectional descriptive study.

Setting. Academic tertiary pain management center.

Subjects. A total of 225 consecutive new patients, aged 18 years or older.

Methods. Data collection included demographics, ORT scores, aberrant behaviors, pain intensity scores, opioid type and dose, smoking status, employment, and marital status.

Results. In this population, we were not able to replicate the findings of the initial ORT study. Self-report was no better than chance in predicting those who would have an opioid aberrant behavior. The ORT risk variables did not predict aberrant behaviors in either gender group. There was significant disparity in the scores between self-reported ORT and the ORT supplemented with medical record data (enhanced ORT). Using the enhanced ORT, high-risk patients were 2.5 times more likely to have an aberrant behavior than the low-risk group. The only risk variable associated with aberrant behavior was personal history of prescription drug misuse.

Conclusions. The self-report ORT was not a valid test for the prediction of future aberrant behaviors in this academic pain management population. The original risk categories (low, medium, high) were not supported in the either the self-reported version or the enhanced version; however, the enhanced data were able to differentiate between high- and low-risk patients. Unfortunately, without technological automation, the enhanced ORT suffers from practical limitations. The self-report ORT may not be a valid tool in current pain populations; however, modification into a binary (high/low) score system needs further study.

Key Words. Opioid Risk Tool; Opioid Stratification Tool; Screening Tools; Opioid Misuse; Validity

Introduction

The rising tide of opioid abuse, misuse, and diversion in the community have cascaded into predictions of 500,000 opioid-related deaths in the United States over the next decade [1]. Prescription opioid medications have become the leading drug of abuse, surpassing cocaine and heroin [2]. High prescription rates among

medical providers have continued despite limited evidence demonstrating sustained benefit and evidence suggesting long-term harm of chronic opioid therapy (COT) [3]. Health risks include an increased risk for motor vehicle collisions [4], poor surgical outcomes [5,6], myocardial infarction [7], bone fractures [8], increased health care expenditures [6,9], and death [10]. Such adverse events are directly related to opioid dosages, resulting in a higher risk of morbidity [10,11]. In many states, more community residents died in 2013 as a result of a drug overdose than from motor vehicle crashes, suicide, breast cancer, colon cancer, firearms, influenza, or HIV [12]. This highlights the importance of understanding a patient's related risk factors to ensure safe treatment and monitoring while reducing risk and improving patient functionality and prognosis.

Opioid prescription guidelines have been developed in an effort to improve the quality and appropriateness of care, promote safer practices, and improve treatment outcomes [11,13,14] for the chronic noncancer pain population. Some opioid prescription guidelines endorse the use of an opioid screening tool, such as the Opioid Risk Tool (ORT), to assist in the prediction of opioid-related aberrant behaviors [15–17]. The Centers for Disease Control and Prevention (CDC) identified insufficient evidence for the effectiveness of opioid risk screening tools [14]. Specifically, the ORT research results were found to be inconsistent and the quality of the studies for all opioid risk stratification tools to be insufficient [14]. Despite this, the ORT is used in clinical practice and has been adopted by health systems as an opioid risk stratification tool [11,18].

The ORT screening instrument was developed in 2001 to aid the medical provider's initial assessment of the risk of opioid-related aberrant behaviors or identify putative risk factors for drug abuse, misuse, and diversion in chronic nonmalignant pain patients [19]. This pilot study introduced the ORT as a brief patient reporting tool to help differentiate between high, moderate, and low risk for opioid usage [19]. The ORT has been used as a comparison tool in the validation of other opioid risk or screening tools, but has produced highly variable results. In a small sample of chronic pain patients, the ORT had a significantly lower sensitivity than that of the clinical interview or Screener and Opioid Assessment for Patient with Pain (SOAPP) [20]. Other studies concluded the ORT has low sensitivity, but modest specificity when compared with SOAPP, the Pain Medication Questionnaire (PMQ), the Brief Risk Questionnaire (BRQ) [21], and clinical interviews in a study of patients engaging in aberrant behavior [22]. Although the ORT is widely used, the original pilot study has not been replicated in a sufficient population size to assess its validity and precision (sensitivity and specificity). In this study, we re-evaluate the validity of the patient self-reported ORT pilot study in a large, diverse pain center clinical population. In a separate analysis, patient self-report ORT data were verified and enhanced with available medical records by a medical provider to determine whether this altered the precision of the scale.

Methods

This was a cross-sectional study in which data were collected from 225 consecutive patients aged 18 years or older in their initial visit from November 16, 2015, to September 30, 2016. Copyright permission of the ORT and Institutional Review Board approval were obtained.

Data Collection

All demographic data were collected from a standardized questionnaire and/or the electronic health record (EHR) including age, gender, marital status, insurance source, educational level achieved, employment status, smoking status, and pain intensity using the numerical rating scale (NRS; 1–10). Aberrant behaviors from the original pilot study were used in this study [19]. Data regarding each aberrant behavior observed or reported over a 12-month period subsequent to the patient's initial visit were collected. Three additional aberrant behaviors (self-directed care, verbal threats, and physical threats) were separately documented in the data set over the same 12 months. These behaviors were included based on the pain center policy in accordance with local best practices. Aberrant behaviors were documented in the patient's EHR after being observed directly, reported by the patient or a family member (and verified with the patient) and/or detected by a urine drug inventory (UDI). All comparisons made to the original study data included the original aberrant behaviors, not the additional three aberrant behaviors unless specifically noted and reported in the text.

ORT Procedure

The results of the ORT were available to the clinician after the initial visit and were not used for clinical decision-making in these patients. The ORT score was calculated using two methods. In the traditional method used in the original pilot study, the patient self-reported risk variables using the ORT questionnaire [19]. The second ORT collection method, termed enhanced ORT, was calculated using a combination of self-report ORT data and EHR data. The EHR data were derived from the first author's manual review each subject's medical records from the six months preceding the initial pain center visit for the presence of any ORT variable. In cases of discrepancy, the EHR record was considered the gold standard and was included in the calculation of the enhanced ORT score.

Statistical Analysis

Several statistical methods were used to compare the ORT and aberrant behaviors. An effort was made to replicate the statistics of the original study where appropriate or illustrative. Descriptive statistics were used for demographic and ORT data. The Kruskal-Wallis analysis and Wilcoxon rank-sum tests were used for nonparametric data. The relationship between ORT risk categories (low, moderate, and high) and the presence or absence of aberrant outcomes was evaluated using the Pearson chi-square test of significance. Logistic regression was used to calculate

computed odds ratios and 95% confidence intervals for the presence or absence of aberrant ORT risk categories. A *P* value of <0.05 was considered significant.

Consistent with the pilot study, the receiver operating characteristic (ROC) analysis was used to calculate the area under the curve (c-statistic) and to validate the ORT. The ROC was used for a dichotomous prediction using the binary classifier of the presence of aberrant behavior (yes/no). The ROC was completed with genders combined and split in order to remain consistent with the analysis of the original study. For interpretation of the c-statistic, $c = 0.5$ suggests no discrimination (i.e., no better than chance), $0.7 \leq c < 0.8$ is considered acceptable, $0.8 \leq c < 0.9$ is considered excellent, and $c \geq 0.9$ is considered outstanding discrimination between those exhibiting aberrant behaviors from those who do not [19].

Additional logistic regression models were used to assess the effect of individual ORT risk variables (e.g., personal history of alcohol abuse) on the outcome. An unadjusted binary univariate analysis was performed to determine whether there was a relationship between each ORT categorical risk factor and the presence of aberrant behavior. Those ORT categorical risk factors with a $P < 0.25$ were carried forward in the multivariable binary logistic regression. $P < 0.05$ was considered significant for the multivariable binary logistic regression.

To examine the accuracy of patient self-reporting, Cohen's Kappa was included to measure the relative observed agreement and the hypothetical probability of chance agreement among the patient self-report ORT and the enhanced version of the ORT. For interpretation of Cohen's Kappa, 0.41–0.60 was considered a moderate agreement, 0.61–0.80 was a substantial agreement, and 0.81–0.99 was almost perfect agreement [23]. Bland and Altman introduced the Bland-Altman (B&A) plot to describe agreement between two quantitative measurements [24]. This analysis was included to evaluate bias between the two measurement techniques of the self-report (gold standard) and enhanced ORT. This analysis allows the identification of systematic differences between the techniques [24]. The mean difference represents the bias, and the standard deviation of the differences measures the random deviation from the mean. If this mean difference of the two measurement techniques differs significantly from zero using a one-sample *t* test, this indicates the presence of measurement technique bias [25].

The data and analysis plan of the ORT and enhanced ORT were determined prior to data collection. All data were analyzed using Statistical Package for the Social Services, v. 24.0 (SPSS Inc., Chicago, IL, USA).

Results

Opioid Risk Tool Score

Using the self-report ORT, differences were found in the number of subjects among the three risk groups (low,

medium, and high) developed in the original study. Age, employment status, educational level, and smoking status were significantly different among levels, but no other demographic variable including gender differed (Table 1). However, there was a gender difference when the risk groups were collapsed and the total self-report ORT score was used ($P = 0.036$). In the enhanced ORT data set, a difference was found between the risk groups for employment status ($P < 0.001$). Similar to the self-report ORT, the enhanced ORT score did not show gender differences by ORT risk group ($P = 0.121$), but genders did differ when the total ORT score was used ($P = 0.004$).

The discrepancy between the two collection methods of ORT was examined. The frequency of individual ORT risk variables was not similar between the self-report and enhanced ORT data (Table 2). Significant discrepancy was found in personal history of illegal drug use ($P = 0.0151$), personal history of prescription drug use ($P = 0.0012$), and personal history of depression ($P = 0.0003$).

Aberrant Behavior

No gender differences were present in total aberrant behaviors ($P = 0.220$) or the binary classifier of presence or absence of an aberrant behavior ($P = 0.325$). The most common aberrant behaviors were attempts to obtain opioids from other providers and reports of self-increased dosage in both genders. When examining the data set including the three aberrant behaviors not included in the original pilot study, the most common aberrant behavior was "self-directed care" (e.g., refusing all offered nonopioid pain management medications and interventions).

Aberrant Behavior and ORT

Twenty-eight percent of patients in the low-risk category displayed aberrant behaviors, and 46% of these subjects had an ORT score of zero. Twenty-nine percent of subjects in the medium and 29% in the high-risk category had a documented aberrant behavior (Table 3). When using the enhanced ORT score, 21% of patients in the low-risk category displayed aberrant behaviors and 49% of these subjects had an ORT score of zero. Thirty-five percent and 44% percent of subjects in the medium- and high-risk categories had a documented aberrant behavior, respectively. Risk categories of the self-report ORT were not significantly different using the binary classifier of presence/absence of aberrant behavior ($p = 0.671$).The self-report data with the 3 additional aberrant behaviors included, and the enhanced ORT data with or without these additional aberrant behaviors were not significant ($p = 0.780$, $p = 0.810$, and $p = 0.074$; respectively). Based on the ORT risk categories, high-risk patients were no more likely to have an aberrant behavior than either low-risk, or low and moderate–risk patients combined, with (Table 4) or without the

Table 1 Self-reported ORT demographics

Characteristics	Low Risk, No. (%)	Moderate Risk, No. (%)	High Risk, No. (%)	*P* Value
Gender				0.444
Female (N = 130)	106 (81.5)	17 (13)	7 (5.5)	
Male (N = 95)	73 (76.8)	12 (12.6)	10 (10.4)	
Age (54.8 ± 15.7), y	55.1 ± 16.4	52.0 ± 14.7	51.2 ± 11.8	0.000
OME (78.9 ± 183.3)	81.2 ± 201.0	61.1 ± 93.9	123.4 ± 237.3	0.484
Ethnicity				0.441
African American (N = 44)	34 (77.2)	6 (13.6)	4 (9.2)	
Caucasian (N = 165)	132 (80.0)	21 (12.7)	12 (7.3)	
Hispanic (N = 10)	9 (90.0)	1 (10.0)	0 (0)	
Asian (N = 1)	1 (100)	0 (0)	0 (0)	
Native American (N = 3)	2 (66.7)	0 (0)	1 (33.3)	
Unidentified (N = 2)	1 (50.0)	1 (50.0)	0 (0)	
Marital status				0.146
Divorced (N = 36)	25 (69.5)	6 (16.7)	5 (13.8)	
Married (N = 106)	86 (81.1)	16 (15.1)	4 (3.8)	
Single (N = 70)	57 (81.4)	6 (8.6)	7 (10.0)	
Widowed (N = 11)	10 (90.1)	1 (9.9)	0 (0.0)	
Unidentified (N = 2)	1 (50.0)	0 (0.0)	1 (50.0)	
Insurance status				0.421
Commercial (N = 59)	52 (88.2)	5 (8.4)	2 (3.4)	
Medicaid (N = 52)	35 (67.3)	10 (19.2)	7 (13.5)	
Medicare (N = 100)	81 (81.0)	13 (13.0)	6 (6.0)	
Workman's comp (N = 9)	7 (77.8)	1 (11.1)	1 (11.1)	
Unidentified (N = 5)	4 (80.0)	0 (0)	1 (20.0)	
Employment status				0.000
Employed (N = 68)	59 (86.8)	7 (10.3)	2 (2.9)	
Retired (N = 52)	46 (88.4)	5 (9.6)	1 (2.0)	
Student (N = 1)	1 (100)	0 (0)	0 (0.0)	
Unemployed (N = 87)	59 (67.8)	16 (18.4)	12 (13.8)	
Unidentified (N = 17)	14 (82.4)	1 (5.9)	2 (11.7)	
Educational level				0.000
College (N = 94)	75 (79.8)	13 (13.8)	6 (6.4)	
High school (N = 88)	72 (81.8)	11 (12.5)	5 (5.7)	
Grade (N = 13)	7 (53.8)	3 (23.1)	3 (23.1)	
Unidentified (N = 30)	25 (83.3)	2 (6.7)	3 (10)	
Smoking status				0.001
Former smoker (N = 25)	18 (72.0)	5 (20.0)	2 (8.0)	
Never smoked (N = 138)	117 (84.8)	14 (10.1)	7 (5.1)	
Smoker (N = 50)	33 (66.0)	9 (18.0)	8 (16.0)	
Unidentified (N = 12)	11 (91.7)	1 (8.3)	0 (0)	
Pain levels				
Rest (5.6 ± 2.6)	5.6 ± 2.6	5.5 ± 2.5	6.1 ± 2.3	0.376
Movement (8.0 ± 2.0)	7.9 ± 2.0	8.6 ± 1.6	8.3 ± 1.8	0.625
Average (7.2 ± 2.0)	7.2 ± 2.0	7.5 ± 1.5	6.8 ± 1.7	0.812

Comparison of the ORT categorical risk groups: Kruskal-Wallis test for age, MME, and pain: chi-square test for gender, ethnicity, insurance source, marital status, education, employment status, and smoking status. Fisher exact test was substituted if assumptions were not met.

OME = oral morphine mg equivalent; ORT = Opioid Risk Tool.

additional three aberrant behaviors. Using the enhanced ORT collection method, high-risk patients were 2.5 times as likely to have an aberrant behavior than the low. This was not significant when low and moderate were combined and compared with the high-risk group; however, the moderate-risk group combined with the high-risk group was 1.9 times as likely to have an aberrant behavior than the low-risk group.

Table 2 Discrepancy between self-report and enhanced ORT collection methods

	Self-Reported ORT, No. (%)	Enhanced ORT, No. (%)	P Value (95% CI)
ORT section 1			
Family history of alcohol abuse	51 (22.67)	59 (26.22)	P = 0.3815 (−4.699 to 11.752)
Family history of illegal abuse	23 (10.22)	29 (12.89)	P = 0.3762 (−3.577 to 8.926)
Family history of prescription abuse	9 (4.00)	13 (5.78)	P = 0.3819 (−2.586 to 6.236)
ORT section 2			
Personal history of alcohol abuse	22 (9.78)	34 (15.11)	P = 0.0871 (−1.082 to 11.758)
Personal history of illegal drug abuse	11 (4.89)	25 (11.11)	P = 0.0151 (0.930 to 11.645)
Personal history of prescription drug abuse	5 (2.22)	21 (9.33)	P = 0.0012 (2.603 to 11.927)
ORT section 3			
Age, y	61 (27.11)	72 (32.00)	P = 0.2562 (−3.846 to 13.550)
ORT section 4			
Personal history of sexual abuse	10 (2.71)	14 (6.22)	P = 0.0718 (−0.625 to 7.853)
Personal history of psychological history	18 (8.00)	25 (11.11)	P = 0.2624 (−2.659 to 8.917)
Personal history of depression	85 (37.78)	123 (54.67)	P = 0.0003 (7.389 to 26.074)

Chi-square analysis.
CI = confidence interval; ORT = Opioid Risk Tool.

Table 3 Webster and Webster, 2005, self-reported ORT, and enhanced ORT by No. of subjects with one or more aberrant behaviors

		Original Pilot Study*		Current Study Self-Reported		Current Study Enhanced Reported	
	Total ORT Score	Males, n/N	Females, n/N	Males, n/N	Females, n/N	Males, n/N	Females, n/N
Low							
	0	NA	NA	12/32	11/46	9/26	9/39
	1	0/3	0/3	8/23	8/36	3/17	4/28
	2	0/0	0/7	3/5	4/16	4/6	6/22
	3	1/3	0/2	3/13	2/9	2/9	0/4
Category total	0–3	1/6	0/12	25/73	25/107	18/58	19/93
Moderate							
	4	4/22	2/17	3/5	0/7	2/5	1/8
	5	4/8	4/28	2/2	1/4	3/5	3/8
	6	6/10	6/12	2/4	0/1	3/5	0/5
	7	4/10	5/16	1/1	0/4	3/6	1/5
Category total	4–7	18/50	17/73	7/12	1/16	11/21	5/26
High							
	8	4/4	2/3	0/0	1/3	1/2	1/4
	9	6/6	4/5	0/1	1/1	1/1	0/1
	10	1/2	4/5	2/3	1/2	3/5	1/2
	11–18	9/9	10/10	0/6	0/1	2/8	3/4
Category total	8–18	20/21	11/23	2/10	3/7	7/16	5/11
Total (%)	0–18	39/77 (50)	37/108 (34)	36/95 (36)	29/130 (22)	36/95 (38)	29/130 (22)

n = number of subjects in the ORT category with one or more aberrant behaviors; N = total number of subjects in the ORT category; ORT = Opioid Risk Tool.
*Data from: Webster LR, Webster RM. Predicting aberrant behaviors in opioid-treated patients: Preliminary validation of the Opioid Risk Tool. Pain Med 2005;6(6):432–42.

Receiver operating curve (ROC) analysis was completed to distinguish between the population with and without aberrant behaviors. ROC analyses were conducted separately for male and female data to remain consistent with the original study. Male and female data were also combined as no gender difference was found between risk categories, the presence of single aberrant behavior, and the total number of aberrant behaviors. For the

Table 4 Comparisons of risk categories by logistic regression

Risk Category Comparisons	Odd Ratios	95% CI
Self-report		
Low vs high	1.033	0.598–1.784
Low+mod vs high	1.021	0.345–3.023
Low vs mod+high	1.239	0.617–2.489
Enhanced		
Low vs high	2.465	1.059–5.737
Low+mod vs high	2.189	0.962–4.978
Low vs mod+high	1.875	1.031–3.412

CI = confidence interval.

binary classifier of presence or absence of an aberrant behavior, the discrimination of the self-report ORT for either males or females was no better than chance ($c = 0.358$, 95% confidence interval [CI] = 0.182–0.533, and $c = 0.573$, 95% CI = 0.409–0.738, respectively) (Figure 1, A and B). Enhanced ORT had similar results with males $c = 0.365$, 95% CI = 0.192–0.538) and females ($c = 0.543$, 95% CI = 0.377–0.709) (Figure 2, A and B). Neither the self-report nor the enhanced ORT had significant discriminative ability when male and female data were combined ($c = 0.531$, 95% CI = 0.406–0.655, and $c = 0.540$, 95% CI = 0.417–0.664, respectively) (Figure 3). The addition of three aberrant behaviors did not improve the discriminative quality of either the self-report or the enhanced ORT when males and females were considered separately (Supplementary Data) or when genders were combined ($c = 0.500$, 95% CI = 0.416–0.584, and $c = 0.427$, 95% CI = 0.314–0.513, respectively) (Supplementary Data).

The intent of the ORT was to predict the risk for future aberrant behaviors. To analyze the relationship between the presence of aberrant behavior in this patient population, a univariate logistic regression with a multivariate regression for factor identification was used. The ORT risk variable, presence of depression, predicted the aberrant behavior in the multivariate regression model using the self-reported ORT data (Table 5). No risk variable predicted the presence of aberrant behavior in the enhanced ORT data set. However, when the three aberrant behaviors not included in the original study were included, the enhanced ORT demonstrated that a personal history of prescription drug abuse was a predictor of aberrant behavior when other factors were taken into account ($P = 0.020$).

Agreement Between Patient-Completed ORT and Medical Provider–Enhanced ORT

The ORT total score mean +/− standard deviation by self-report and enhanced methodology were 2.29 +/−3.55 and 3.15 +/−4.10, respectively. The medians showed a significant discrepancy between the two collection methods ($P < 0.001$). Cohen's Kappa (κ) is the proportion of agreement between two observed raters. The self-report vs enhanced completion of the ORT was compared and identified 71.4% agreement. Incorporating chance resulted in $\kappa = 0.703$ ($P < 0.001$, SE = 0.033), suggesting substantial agreement. A Bland-Altman analysis was included to evaluate whether discrepancies and/or bias exists between the mean differences of the second method (enhanced) of completion of the ORT when compared with the first method (self-report) (Figure 4). The mean difference between the methods (−0.889 ± 1.94, 95% CI = −1.144 to −0.633) was significantly different in a one-way sample t test ($P < 0.001$). The B&A plot shows that the moderate range of the ORT scale accounts for a large part of the difference.

Discussion

Opioid prescribing guidelines have been developed in an effort to improve the quality and appropriateness of care, promote safer practices, and improve treatment outcomes [11,13,14] for the chronic noncancer pain population. These best-practice clinical guidelines endorse the use of screening tools, such as the ORT, to identify high-risk patients prior to starting opioid therapy [3,14–17,26] despite insufficient evidence of their effectiveness [14]. The ORT [19], introduced in 2005 (with data collection in 2001), has been widely used due to its simplicity, ease of administration, and apparent face validity. The lack of a subsequent study replicating the initial pilot study's findings prior to implementation into clinical practice has remained a significant criticism of the instrument [26].

In this large interdisciplinary academic pain center, we were unable to replicate the findings of the original ORT pilot study using the same ORT collection methodology, aberrant behavior, and 12-month time frame. Although the original study reported a c-statistic in the excellent range for discrimination (males $c = 0.82$ and females $c = 0.85$), the current study never exceeded chance discrimination (males $c = 0.36$ and females $c = 0.57$) [19]. Therefore, in this study population, the self-reported ORT was unable to discriminate between those who would or would not have an aberrant behavior over 12 months at a rate greater than chance.

The original study divided the population into three risk categories (low, medium, and high) based on their responses to the ORT risk variable questions. These divisions were not supported in this study as the likelihood of aberrant behavior did not differ between risk categories using the same time and aberrant behavior criteria of the original study. However, the enhanced data set including the additional three aberrant behaviors did show that the high-risk category had significantly higher odds of having an aberrant behavior than the low-risk group when moderate was excluded or when the moderate category was combined with the
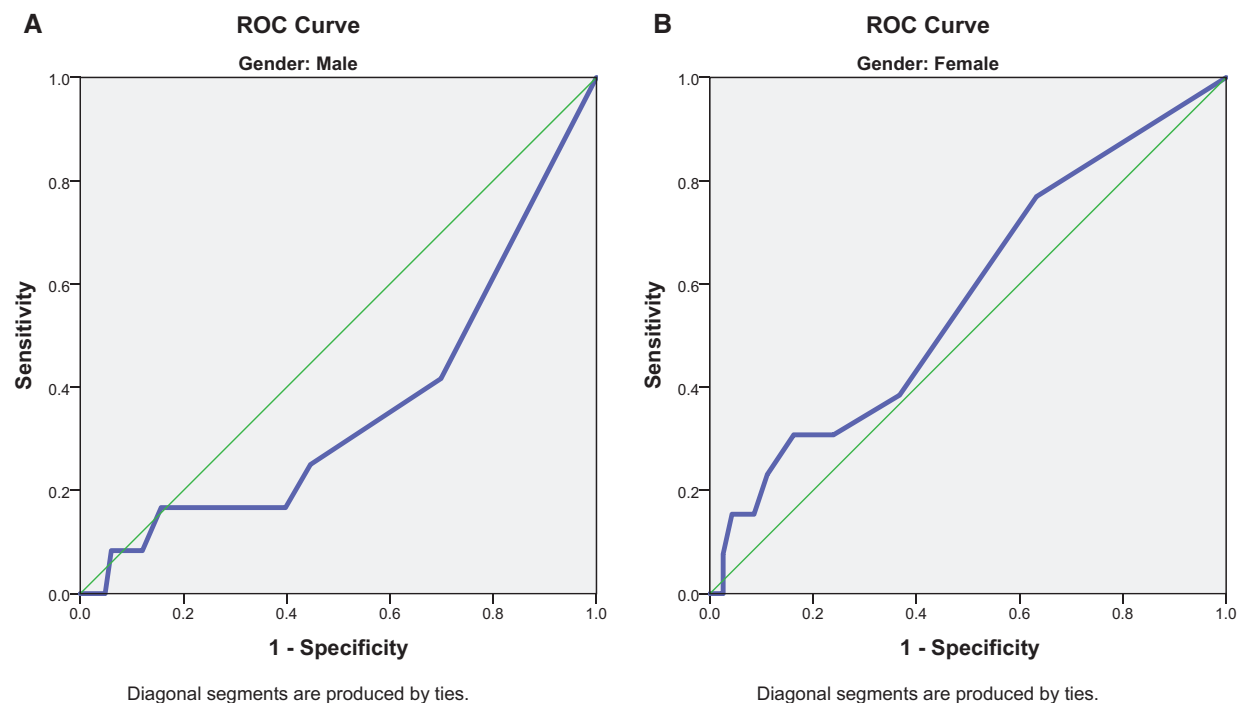
A

ROC Curve

Gender: Male

B

ROC Curve

Gender: Female

Diagonal segments are produced by ties.

Diagonal segments are produced by ties.

Figure1 (**A**) Male gender receiver operating characteristic (ROC) curve–self-report Opioid Risk Tool (ORT). (**B**) Female gender ROC curve–self-report ORT.

high-risk group. The separation of genders was not supported in the current study. Male and female patients had a similar total number of aberrant behaviors, frequency of aberrant behavior, and ORT scores when examined by risk category. Additionally, no single risk variable measured on the self-reported ORT was predictive of the presence of aberrant behavior. There is increasing evidence concerning the frequency and types of aberrant behaviors in relation to the potential relationship to opioid use disorders. However, unlike the original study, the presence of aberrant behavior was not associated with increasing risk category. In the original study, for those with an ORT Total Score of >11, 100% displayed at least one aberrant behavior. When compared with the current study pain population, there were fewer patients in this category and none of these patients displayed any aberrant behavior. In the current study, 29.4% of the high-risk group (eight to 18) had aberrant behaviors, while 70.5% of this group had an aberrant behavior in the original work. Unlike the original work, there was no dose-response relationship between ORT risk category and the aberrant behavior (Table 3). In our sample, the minority of patients were in the high-risk category, whereas the original study skewed toward moderate- to high-risk patients based on the ORT total score. The lower total ORT score and overall rate of aberrant behavior are somewhat surprising given the increase prevalence of prescription opioid misuse and abuse in the United States [27,28].

There are numerous explanations for the differences found between the original pilot study and this study.

These include the population studied, factors used in the development of original scale, the changing treatment approaches, and changing cultural norms. Although the original study included a limited sample population demographic description of the pain population, it can likely be surmised that the demographics of the general populations of the two communities surrounding the study sites (Salt Lake City, 2001–2002, vs Milwaukee, 2015–2017) are likely to be different. In this study, our population represented the greater community of our urban/suburban location, similar to many large cities in the United States. Our sample was older, lower opioid consuming, and likely more ethnically diverse than the original sample. Although our population reflects a common insurance payer mix of an academic medical center, it was likely over-represented by unemployed patients and may have overestimated the results in the high-risk category (Table 1). Despite this, we did not find higher aberrant behavior rates than the original study.

The original descriptions of ORT risk factors and their assigned weights vs gender were based on the literature at that time, although current literature shows that in comparison with women, men are more likely to experiment with various types of illicit drugs, with greater emergency department visits for illicit drug use and/or overdose deaths [29]. Despite this fact, abuse of prescription drugs is equal between genders. One exception is the female population aged 12–17 years, which has an increased rate of prescription drug abuse.
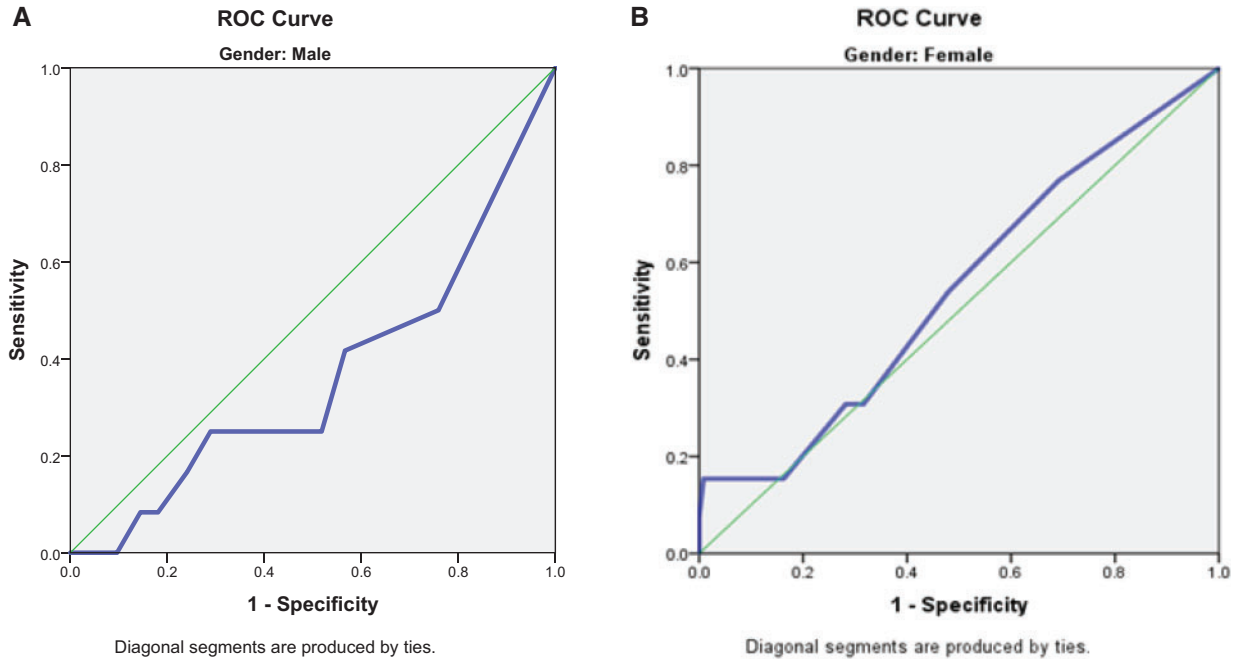
Figure 2 (**A**) Male gender receiver operating characteristic (ROC) curve–enhanced Opioid Risk Tool (ORT). (**B**) Female gender ROC curve–enhanced ORT.
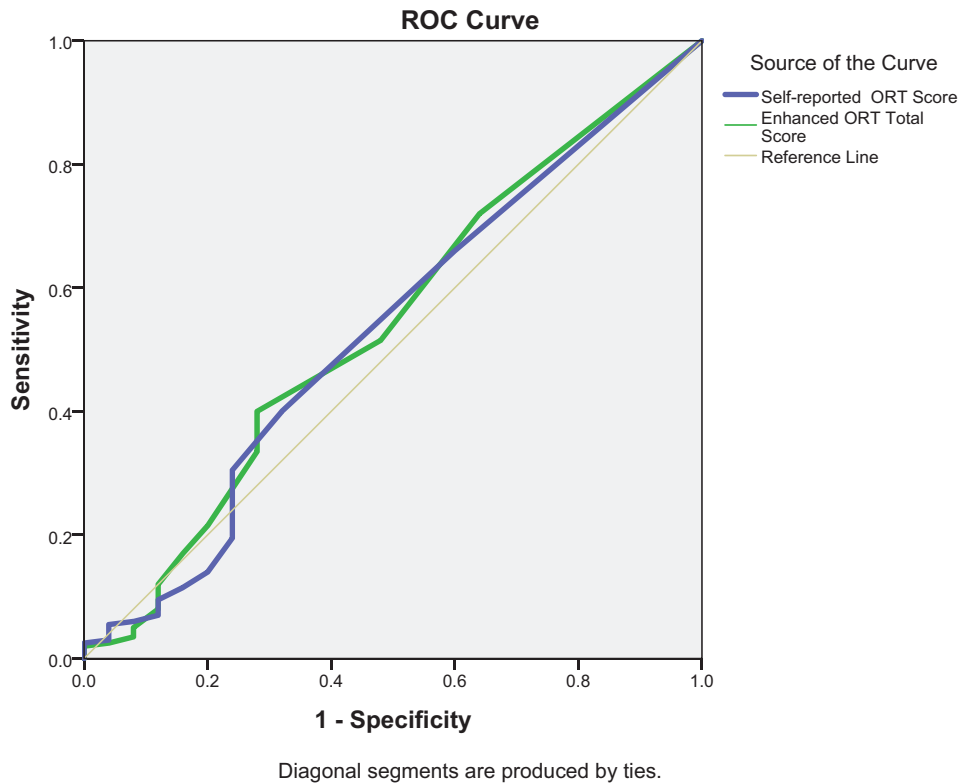

Figure 3 Gender combined receiver operating characteristic–self-report and enhanced Opioid Risk Tool.

Table 5 Logistic regression

ORT section	Patient-Completed ORT			Medical Provider–Enhanced ORT		
	Univariate Data Analysis P Value, Odds Ratio (95% CI)	Adjusted Multivariate Analysis Odds Ratio (95% CI)	P Value	Univariate Data Analysis P Value, Odds Ratio (95% CI)	Adjusted Multivariate Odds Ratio (95% CI)	P Value
Section 1						
Family H/O alcohol	0.528, 0.847 (0.506–1.418)	1.254 (0.942–1.669)	0.121	0.547, 0.821 (0.49–1.38)		
Family H/O illegal drugs	0.799, 1.071 (0.630–1.822)			0.927, 1.03 (060–1.74)		
Family H/O prescription drugs	0.436, 1.160 (0.798–1.686)			0.426, 1.178 (0.787–1.761)		
Section 2						
Personal H/O alcohol	0.345, 0.721 (0.366–1.422)			0.190, 0.636 (0.323–1.251)	0.636 (0.323–1.251)	0.190
Personal H/O illegal drugs	0.909, 0.907 (0.572–1.643)			0.298, 0762 (0.457–1.272)		
Personal H/O prescription drugs	0.998, 0.000 (0.00–0.00)			0.449, 1.107 (0.850–1.442)		
Section 3						
Age	0.292, 1.601 (0.667–3.843)			0.337, 1.482 (0.619–3.553)		
Section 4						
Sexual abuse history	0.969, 1.016 (0.490–2.104)			0.721, 0.875 (0.422–1.817)		
Section 5						
OCD, ADD, bipolar, and schizophrenia	0.998, 0.000 (0.000–0.000)			0.751, 0.885 (0.414–1.889)		
Depression	0.026, 0.286 (0.095–0.864)	0.286 (0.095–0.864)	0.026	1.0, 1.0 (0.435–2.298)		

ADD = Attention Deficit Disorder; CI = confidence interval; H/O = History of; ORT = Opioid Risk Tool; OCD = Obsessive Compulsive Disorder.
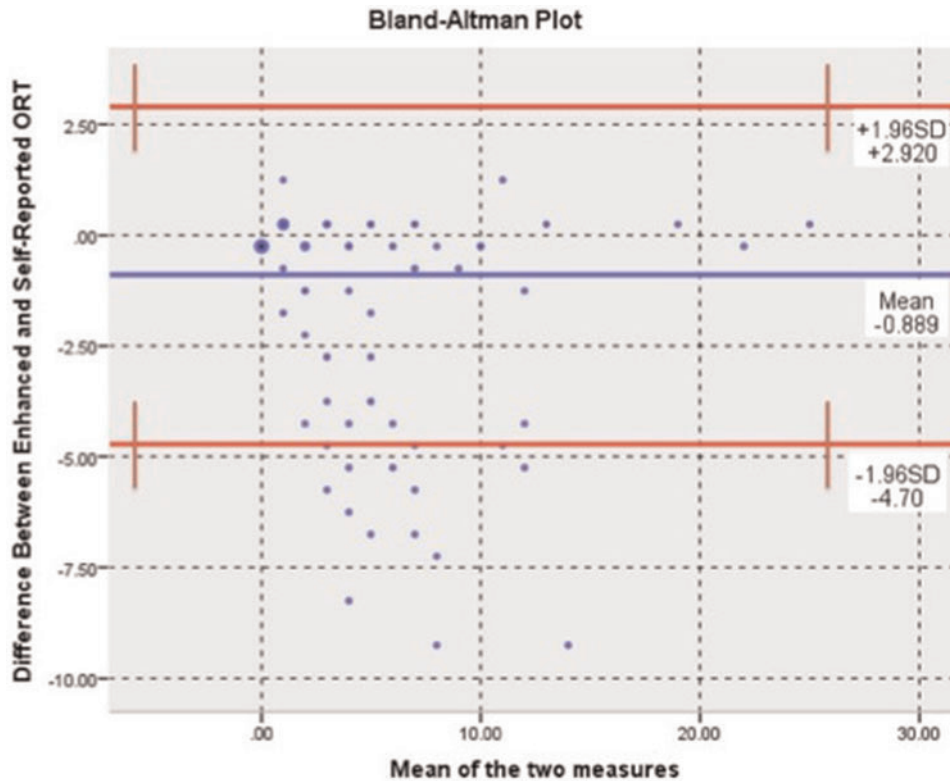
Figure 4 Difference between enhanced and self-reported Opioid Risk Tool.

For most age groups, men have higher rates of use or dependence on illicit drugs and alcohol than women. However, women are equally at risk to become addicted to illicit drugs, alcohol, and prescriptive medications [30]. This gender-based assumption in the development of the instrument may reflect a bias related to the era of its development. The development of the original ORT included numerous assumptions. These included risk factors (e.g., family history of illegal drug use), weighting of risk factors (personal history of substance abuse is more highly weighted than family history of substance abuse), separation of genders in risk factors, and analysis based on the developers' understanding of the literature in 2001. It is unclear in the original pilot study why sexual abuse was not weighted in the total ORT score for male patients. Of the three male patients positive for sexual abuse in our population, each was positive for psychological disease, three out of three were positive for personal substance abuse, and the average oral morphine mg equivalent (OME) was 131. This pattern is also seen in the literature. There is a significant correlation between adverse childhood experiences associated with substance abuse [31]. Although limited research studies have been completed with males, sexual abuse, and association with future substance abuse, a recent epidemiologic survey demonstrated that a significant relationship between childhood sexual abuse and drug abuse disorder is stronger in men than women [32]. Furthermore, the Latino-American men surveyed demonstrated that childhood physical abuse is positively associated with lifetime substance abuse [33]. The lack of weighting of this risk factor in men may influence the resulting analysis.

The original study referred to drug dependence and abuse rates in order to establish the age group (16–45 years) at highest risk. Although illegal drug experimental use is the highest among people in their late teens and twenties [29,34], research also indicates an increase drug abuse among individuals aged 50–60 years [34]. Furthermore, in the age group >65 years, females (7.2%) are three times more likely to abuse prescription medications than men (2.8%) [35] and statistics illustrate that middle-aged adults have the highest prescription painkiller overdose rates [21,30,35,36]. Consistent with this, age was not predictive of aberrant behavior in this study population. The difference between the original study and this one may reflect changes in the age of those who misuse pain medications or who have had aberrant behavior since the development of the ORT.

Some risk factors were discussed but not included in the ORT's development. Post-traumatic stress disorder (PTSD) was mentioned as having a relationship to pre-adolescent sexual abuse and substance abuse disorders, but it does not link it to a risk factor. Childhood trauma, such as physical, emotional, and sexual abuse, has been associated with substance abuse, and the

effect extends into adulthood [37]. Although research has found that sexual abuse is not predictive in adult prescription drug misuse [37], cumulative trauma (combination of physical, emotional, and sexual abuse) was linked to future prescription drug misuse [37–39]. Each addition of adverse childhood experiences increases the odds of the patient having a future drug disorder [31]. This current study found no relationship between preadolescent sexual abuse and the presence of aberrant behavior. Extending the observation to cumulative childhood trauma, or PTSD, may have captured the patients' risks for future misuse, abuse, and diversion of opioid medications.

The pilot study appears to have not used standardized methods to collect aberrant behaviors, such as questionnaires or UDIs, as has been suggested by the literature [26]. This may have resulted in a recall bias in their results. In the present study, the ORT was collected, but clinical action and decision-making were not based on the ORT results. The aberrant behaviors were standardized in the medical record, and UDIs were performed on all patients receiving opioids. This was done to lessen the potential of bias or producing a self-fulfilling prophecy. Despite these differences, similar to the original study, "solicited opioids from other providers," "unauthorized dose escalation," "abnormal urine/blood screen," and "used additional opioid than those prescribed" were the most common aberrant behaviors.

Additional differences in results could relate to the differences in pain treatment philosophy of the clinicians in the two studies. The original study philosophy of treatment was consistent with pain treatment at that time, including titration of opioids to optimal pain relief levels with upper dosage only limited to side effects [19]. In the original study, it was reported that some patients reached several hundred OMEs. The pain treatment approach employed in this pain center might be more reflective of recent research showing decreases in opioid dose prescribing [40]. Studies have shown that increased dosage of prescription opioids is associated with aberrant behavior [10]. Therefore, the higher dosages used in the original study may have contributed to higher incidence of aberrant behavior.

Qualifications of an aberrant behavior have also changed since the development of the ORT. Aberrant behaviors that were not included in the original study included "self-directed care," defined as refusal to try other nonopioid medications or participate in any other pain reduction strategies. Physical and verbal threats toward clinicians were also included. Physical and verbal abuse has increased toward chronic pain care providers. A survey reported that 51.5% of chronic pain physicians have experienced threats of bodily harm [37]. Specific types of threats were car vandalism, bad reviews on the internet, pulling out a gun to show a physician, threats to the medical board, threats to sue, and threats to light the physician on fire with

gasoline [37]. These behaviors may not have been a significant concern for clinicians at the time the ORT was developed; they have become a concern in the treatment and safety of today's pain population. When three additional aberrant behaviors were included in the analysis, the ORT risk variable "personal history of prescription drug abuse" was predictive of the presence of aberrant behavior.

Enhanced vs Self-Report ORT

Another norm that may have changed and therefore influenced ORT results is the lack of willingness to self-report risk factors. The current study illustrated discriminative ability of the ORT that was no better than chance. Despite significant discrepancy between self-reported ORT results and those of the enhanced ORT in our population, the enhanced ORT using medical records was only modestly better than self-report at predicting aberrant behavior. This is consistent with a small study in which the results of the ORT depended on who completed the scale [41]. The ORT completed by the clinician had a higher predictive rate of aberrant behaviors (57%) when compared with patient self-report (30%). These study authors speculated that patients consciously underreport past substance abuse. Despite these limitations, the odds of aberrant behavior in high- vs low-risk categories were significantly different with the enhanced ORT data. This suggests that there may be value for future studies to re-examine the ORT score as a binary (high/low) variable.

The monitoring of aberrant behaviors also depends on observant health care providers, laboratory testing, and/ or patient self-report [42], all of which have limitations. Health care providers tend to be able identify the most severe aberrant behaviors [42]. Urine drug screens have a limited window of identification of many illicit drugs [42]. Finally, the patient may neglect to self-report aberrant behaviors for fear of termination or reduction of opioid treatment [42,43]. The willingness to self-report aberrant behavior that may result in the discontinuation of opioid medications could be declining with the increased attention to appropriate use of opioids resulting from the current opioid epidemic.

Risk stratification instruments vary in complexity and focus, ranging from three to 42 questions. Instruments differ for self-administered (SOAPP, ORT, SOAPP-R, BRQ, PMQ, and COMM) or interview-administered (DIRE, PDUQ, and ABC) approaches. Different screening instruments focus on past medical history or psychiatric factors, or both in a few cases. Few quality studies have been completed on opioid risk stratification tools. Those that have been completed are lacking in methodology, standard measures, and statistical correlation [26]. Subsequently, the results are inaccurate and inconsistent [14]. A very recent study employed natural language processing techniques (NLP) to derive an her-based ORT score [44]. Using this method, the ORT was found to be highly sensitive and specific for violations of

a clinic's opioid agreement. A stated limitation of this approach is that its retrospective nature limits the included risk variables to those that were found in the EHR. It has also been recognized that clinicians do not recognize and document all ORT risk factors and aberrant behaviors, leading to incomplete patient outcomes [45]. Further investigations into the standardization, accuracy, and consistency of opioid stratification tools are encouraged, such as comparisons with other risk stratification tools or opioid checklists [46].

Limitations

The study had several limitations. Although the objective data today are likely more precise than 16 years ago, with the implementation of the electronic health record (EHR) allowing for thorough health histories to improve the accuracy of objective data, it remains likely that EHR data are not complete in the domains relevant to misuse stratification. Additionally, one team member collected and interpreted the ORT and aberrant behaviors, leading to potential work-up bias in the results, although this was mitigated by examining all patient medical records regardless of initial ORT score or suspicion of underreporting. A secondary review by the co-authors occurred if a question or discrepancy arose. Of note, multiple patients admit to cannabis usage and believe that it should not be categorized as an illegal drug. This may decrease self-reporting of personal illegal drug use history on patient questionnaires. Also, the team member had ample time to review the EHR and collect objective data, unlike a clinician actively seeing patients with a limited time window to review patients' past medical history to obtain an accurate ORT score. This severely limits the value of the enhanced ORT unless deployed as a natural language processing algorithm [44]. Unfortunately, in this population, the enhanced was no better than the self-report ORT. Aberrant behaviors have been described in various degrees from inconsequential to egregious [19]. Weighted or prioritized aberrant behaviors were not completed in either study, meaning that a patient's past positive UDI for illicit drugs was not considered a more severe infraction than a patient who frequently requested an early prescription refill. It is likely that clinician actions would differ based upon the type of aberrant behavior. This question needs further study.

Conclusion

Although the ORT was designed as a fast, simple stratification tool to assess patients' future potential for abuse, misuse, and diversion of opioid medications, in this pain population, the ORT risk assessment was no better than chance when used as a self-report tool or when enhanced by further her-derived data. The present study's results suggest caution when interpreting the ORT results and planning clinical care using ORT findings. Further study on the risk factors and risk factor categorization is warranted before adoption of the ORT in its present form as a clinician aid or in EHR systems.

Supplementary Data

Supplementary Data may be found online at http://painmedicine.oxfordjournals.org.

References

1 Blau M. STAT forecast: Opioids could kill nearly 500,000 Americans in the next decade. STAT: Reporting from the Frontiers of Health and Medicine. 2017.

2 Page MG, Saidi H, Ware MA, Choiniere M. Risk of opioid abuse and biopsychosocial characteristics associated with this risk among chronic pain patients attending a multidisciplinary pain treatment facility. Clin J Pain 2016;32(10):859–69.

3 Chou R, Turner JA, Devine EB, et al. The effectiveness and risks of long-term opioid therapy for chronic pain: A systematic review for a National Institutes of Health Pathways to Prevention Workshop. Ann Intern Med 2015;162(4):276–86.

4 Gomes T, Redelmeier DA, Juurlink DN, et al. Opioid dose and risk of road trauma in Canada: A population-based study. JAMA Intern Med 2013;173(3):196–201.

5 Gulur P, Williams L, Chaudhary S, Koury K, Jaff M. Opioid tolerance–a predictor of increased length of stay and higher readmission rates. Pain Physician 2014;17(4):E503–7.

6 Cron DC, Englesbe MJ, Bolton CJ, et al. Preoperative opioid use is independently associated with increased costs and worse outcomes after major abdominal Surgery. Ann Surg 2017;265(4):695–701.

7 Li L, Setoguchi S, Cabral H, Jick S. Opioid use for noncancer pain and risk of myocardial infarction amongst adults. J Intern Med 2013;273(5):511–26.

8 Li L, Setoguchi S, Cabral H, Jick S. Opioid use for noncancer pain and risk of fracture in adults: A nested case-control study using the general practice research database. Am J Epidemiol 2013;178(4):559–69.

9 Waljee JF, Cron DC, Steiger RM, et al. Effect of preoperative opioid exposure on healthcare utilization and expenditures following elective abdominal surgery. Ann Surg 2017;265(4):715–21.

10 Dunn KM, Saunders KW, Rutter CM, et al. Opioid prescriptions for chronic pain and overdose: A cohort study. Ann Intern Med 2010;152(2):85–92.

11 Edlund MJ, Martin BC, Russo JE, et al. The role of opioid prescription in incident opioid abuse and dependence among individuals with chronic noncancer pain: The role of opioid prescription. Clin J Pain 2014;30(7):557–64.

12 Miller J. Drug Overdoses Are a Leading Cause of Injury Deaths in Wisconsin. Madison, WI: Wisconsin Department of Health Services; 2015.

13 Manchikanti L, Abdi S, Atluri S, et al; American Society of Interventional Pain Physicians. American Society of Interventional Pain Physicians (ASIPP) guidelines for responsible opioid prescribing in chronic non-cancer pain: Part I–evidence assessment. Pain Physician 2012;15(suppl 3):S1–65.

14 Dowell D, Haegerich TM, Chou R. CDC guideline for prescribing opioids for chronic pain—United States, 2016. MMWR Recomm Rep 2016;65(1):1–49.

15 Lasser KE, Shanahan C, Parker V, et al. A multi-component intervention to improve primary care provider adherence to chronic opioid therapy guidelines and reduce opioid misuse: A cluster randomized controlled trial protocol. J Subst Abuse Treat 2016;60:01–9.

16 Chou R, Fanciullo GJ, Fine PG, et al; American Pain Society-American Academy of Pain Medicine Opioids Guidelines Panel. Clinical guidelines for the use of chronic opioid therapy in chronic noncancer pain. J Pain 2009;10(2):13–30.

17 APS-AAPM. APS-AAPM clinical guidelines for the use of chronic opioid therapy in chronic noncancer pain. Available at: http://www.americanpainsociety. org/uploads/education/guidelines/chronic-opioid-therapy-cncp.pdf. 2009. http://www.drugabuse. gov/(Accessed January 6, 2017).

18 Vargas-Schaffer G, Cogan J. Attitudes toward opioids and risk of misuse/abuse in patients with chronic noncancer pain receiving long-term opioid therapy. Pain Med 2017;19(2):319–27.

19 Webster LR, Webster RM. Predicting aberrant behaviors in opioid-treated patients: Preliminary validation of the opioid risk tool. Pain Med 2005;6(6):32–42.

20 Moore TM, Jones T, Browder JH, Daffron S, Passik SD. A comparison of common screening methods for predicting aberrant drug-related behavior among patients receiving opioids for chronic pain management. Pain Med 2009;10(8):426–33.

21 Jones T, Lookatch S, Moore T. Validation of a new risk assessment tool: The brief risk questionnaire. J Opioid Manag 2015;11(2):71–83.

22 Jones T, Moore T, Levy JL, et al. A comparison of various risk screening methods in predicting discharge from opioid treatment. Clin J Pain 2012; 28(2):3–100.

23 Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33 (1):59–74.

24 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1(8476):307–10.

25 Giavarina D. Understanding Bland Altman analysis. Biochem Med (Zagreb) 2015;25(2):141–51.

26 Chou R, Fanciullo GJ, Fine PG, et al. Opioids for chronic noncancer pain: Prediction and identification of aberrant drug-related behaviors: A review of the evidence for an American Pain Society and American Academy of Pain Medicine clinical practice guideline. J Pain 2009;10(2):31–46.

27 Centers for Disease Control and Prevention. Vital signs: Overdoses of prescription opioid pain relievers and other drugs among women–United States, 1999-2010. MMWR Morb Mortal Wkly Rep 2013; 62(26):537–42.

28 Guy GP Jr, Zhang K, Bohm MK, et al. Vital signs: Changes in opioid prescribing in the United States, 2006-2015. MMWR Morb Mortal Wkly Rep 2017;66 (26):97–704.

29 United States. Substance Abuse and Mental Health Services Administration. Office of Applied Studies. National Survey on Drug Use and Health, US Substance Abuse and Mental Health Services Administration Office of Applied S, Center for Behavioral Health S, Quality. Results from the National Survey on Drug Use and Health. 2009. http://www.drugabuse.gov/(Accessed January 7, 2017)

30 NIDA. Substance Use in Women. 2016. https:// www.drugabuse.gov/publications/research-reports/ substance-use-in-women(Accessed January 7, 2017).

31 LeTendre ML, Reed MB. The effect of adverse childhood experience on clinical diagnosis of a substance use disorder: Results of a nationally representative study. Subst Use Misuse 2017;52(6):89–97.

32 Choi NG, DiNitto DM, Marti CN, Choi BY. Association of adverse childhood experiences with lifetime mental and substance use disorders among men and women aged 50+ years. Int Psychogeriatr 2017;29(3):59–72.

33 Ai AL, Lee J, Solis A, Yap C. Childhood abuse, religious involvement, and substance abuse among Latino-American men in the United States. Int J Behav Med 2016;23(6):64–75.

34 NIDA. National Institute on Drug Abuse, National Institute on Drug Abuse. Advancing Addiction Science. 2000. http://www.drugabuse.gov/ (Accessed January 7, 2017)

35 Smith K. Gender Differences in Primary Substance of Abuse Across Age Groups. Rockville, MD: The CBHSQ Report; 2013.

36 Center for disease control and prevention: Prescription pain killer overdoses in the us. Available at: https://www.cdc.gov/vitalsigns/ PainkillerOverdoses/index.html. 2016(Accessed July 29, 2017).

37 Quinn K, Boone L, Scheidell JD, et al. The relationships of childhood trauma and adulthood prescription pain reliever misuse and injection drug use. Drug Alcohol Depend 2016;169:90–8.

38 Daigre C, Rodriguez-Cintas L, Tarifa N, et al. History of sexual, emotional or physical abuse and psychiatric comorbidity in substance-dependent patients. Psychiatry Res 2015;229(3):43–9.

39 Mandavia A, Robinson GG, Bradley B, Ressler KJ, Powers A. Exposure to childhood abuse and later substance use: Indirect effects of emotion dysregulation and exposure to trauma. J Trauma Stress 2016;29(5):22–9.

40 Garg RK, Fulton-Kehoe D, Turner JA, et al. Changes in opioid prescribing for Washington workers' compensation claimants after implementation of an opioid dosing guideline for chronic noncancer pain: 2004 to 2010. J Pain 2013;14(12):620–8.

41 Jones T, Passik SD. A comparison of methods of administering the opioid risk tool. J Opioid Manag 2011;7(5):47–51.

42 Nikulina V, Guarino H, Acosta MC, et al. Patient vs provider reports of aberrant medication-taking behavior among opioid-treated patients with chronic pain who report misusing opioid medication. Pain 2016;157(8):791–8.

43 Fleming MF, Balousek SL, Klessig CL, Mundt MP, Brown DD. Substance use disorders in a primary care sample receiving daily opioid therapy. J Pain 2007;8(7):73–82.

44 Haller IV, Renier CM, Juusola M, et al. Enhancing risk assessment in patients receiving chronic opioid analgesic therapy using natural language processing. Pain Med 2017;18(10):952–60.

45 Setnik B, Roland CL, Pixton GC, Sommerville KW. Prescription opioid abuse and misuse: Gap between primary-care investigator assessment and actual extent of these behaviors among patients with chronic pain. Postgrad Med 2017;129(1):5–11.

46 Pilkonis PA, Yu L, Dodds NE, et al. An item bank for abuse of prescription pain medication from the patient-reported outcomes measurement information system (PROMIS). Pain Med 2017;18(8):516–27.