



Concurrent validation of the Clinical Opiate Withdrawal Scale (COWS) and single-item indices against the Clinical Institute Narcotic Assessment (CINA) opioid withdrawal instrument

D. Andrew Tompkins^{a,*}, George E. Bigelow^a, Joseph A. Harrison^a, Rolley E. Johnson^{a,b}, Paul J. Fudala^b, Eric C. Strain^a

^a Johns Hopkins University School of Medicine, Department of Psychiatry and Behavioral Sciences, Behavioral Pharmacology Research Unit, 5510 Nathan Shock Drive, Baltimore, MD 21224, USA

^b Reckitt Benckiser Pharmaceuticals Inc., Richmond, VA 23235, USA

ARTICLE INFO

Article history:

Received 10 April 2009

Received in revised form 25 June 2009

Accepted 1 July 2009

Available online 3 August 2009

Keywords:

Opioid withdrawal

Opioid dependence

Clinical Opiate Withdrawal Scale

Clinical Institute Narcotic Assessment

Naloxone

Precipitated withdrawal

ABSTRACT

Introduction: The Clinical Opiate Withdrawal Scale (COWS) is an 11-item clinician-administered scale assessing opioid withdrawal. Though commonly used in clinical practice, it has not been systematically validated. The present study validated the COWS in comparison to the validated Clinical Institute Narcotic Assessment (CINA) scale.

Method: Opioid-dependent volunteers were enrolled in a residential trial and stabilized on morphine 30 mg given subcutaneously four times daily. Subjects then underwent double-blind, randomized challenges of intramuscularly administered placebo and naloxone (0.4 mg) on separate days, during which the COWS, CINA, and visual analog scale (VAS) assessments were concurrently obtained. Subjects completing both challenges were included ($N = 46$). Correlations between mean peak COWS and CINA scores as well as self-report VAS questions were calculated.

Results: Mean peak COWS and CINA scores of 7.6 and 24.4, respectively, occurred on average 30 min post-injection of naloxone. Mean COWS and CINA scores 30 min after placebo injection were 1.3 and 18.9, respectively. The Pearson's correlation coefficient for peak COWS and CINA scores during the naloxone challenge session was 0.85 ($p < 0.001$). Peak COWS scores also correlated well with peak VAS self-report scores of bad drug effect ($r = 0.57$, $p < 0.001$) and feeling sick ($r = 0.57$, $p < 0.001$), providing additional evidence of concurrent validity. Placebo was not associated with any significant elevation of COWS, CINA, or VAS scores, indicating discriminant validity. Cronbach's alpha for the COWS was 0.78, indicating good internal consistency (reliability).

Discussion: COWS, CINA, and certain VAS items are all valid measurement tools for acute opiate withdrawal.

© 2009 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

The opiate withdrawal syndrome, a constellation of characteristic signs and symptoms, has been called “one of the most stereotyped syndromes in clinical medicine” (Isbell, 1950). The first instrument to quantitatively measure withdrawal was developed by Kolb and Himmelsbach in the mid-1930s (Kolb and Himmelsbach, 1938). That scale was based on clinical observations and was weighted heavily towards physical signs of withdrawal, such as systolic blood pressure changes, mydriasis, fever, and respiratory rate changes. In the 1960s, the Opiate Withdrawal Experience Scale, a subset of self-report questions from the Addiction Research

Center Inventory (ARCI), was used to quantify the subjective symptoms of withdrawal (Haertzen and Meketon, 1968). However, this scale was time consuming for subjects to complete, even with the derived short form Opiate Withdrawal Questionnaire (Haertzen et al., 1970).

Following development of those initial instruments, multiple other subjective and objective scales have been developed and used (Handelsman et al., 1987; Judson et al., 1980; Wang et al., 1974; Bradley et al., 1987; Gossop, 1990). Methods for using these scales have sought to improve on sensitivity and specificity for detecting withdrawal by controlling the level of physical dependence, the time point within the withdrawal syndrome when the assessment is made, and the possibility of feigned responses. In 1988, Peachey and Lei reported on the reliability and validity of the Clinical Institute Narcotic Assessment (CINA), one of the first scales to include both opiate withdrawal signs and symptoms (Peachey and

* Corresponding author. Tel.: +1 410 550 5953; fax: +1 410 550 0030.
E-mail address: dtompki1@jhmi.edu (D.A. Tompkins).

Lei, 1988). This scale was validated using a naloxone challenge in heroin-dependent subjects and the peak score was found to predict the clinically determined maintenance methadone dose used to treat these patients. However, the CINA required nursing support to measure heart rate and blood pressure and contained items which could be easily feigned. As well, there was no fixed upper limit to the scale given the variable contribution of blood pressure and pulse ratings.

Wesson and colleagues, therefore, developed the Clinical Opiate Withdrawal Scale (COWS). This scale was designed to be administered quickly, was intended to improve upon existing measurement tools, and was first published in a training manual for buprenorphine treatment (Wesson et al., 1999). The COWS consisted of an 11-item rating system that could be completed within 2 min by a trained observer and could track opioid withdrawal as differentiated from opioid toxicity through serial measurements. Total scores ranged from 0 to 47, and withdrawal was classified as mild (5–12), moderate (13–24), moderately severe (25–36), or severe (>36). These category scores were not derived using standard statistical techniques but were based upon the authors' clinical expertise (Wesson and Ling, 2003). Because of its clinical utility, its association with buprenorphine maintenance, and ease of application, the COWS has become widely used for assessing opiate withdrawal (Center for Substance Abuse Treatment, 2004). Although the scale was modeled after items on previously validated scales, the COWS itself has never been systematically validated (Wesson and Ling, 2003). The present project assessed the validity of the COWS in comparison to a previously validated instrument, the CINA, using a double-blind, placebo-controlled naloxone challenge in opioid-dependent individuals. As well, comparisons between the COWS, CINA, and single-item subjective ratings (Visual Analogue Scales) were done to examine the validity and possible utility of using one overall item to assess opioid withdrawal.

2. Methods

2.1. Participants

Forty-six out-of-treatment opioid-dependent volunteers participated while residing on a supervised research unit at the Johns Hopkins Behavioral Pharmacology Research Unit (BPRU). Participants were recruited for a clinical trial that will be reported on separately; the trial was registered at www.clinicaltrials.gov, identifier NCT00637000. The analyses in this paper were done as part of the confirmation of opioid physical dependence required for the subsequent opioid clinical pharmacology study. In order to be enrolled, participants had to meet DSM-IV-TR (American Psychiatric Association, 2000) criteria for opioid dependence and be between the ages of 18 and 65, willing to stay on the residential research unit for up to 12 days in order to complete the full clinical trial, and on adequate birth control (if female). Exclusionary factors were clinically significant medical or psychiatric diagnoses (i.e. schizophrenia or active suicidal ideation); engaged in opioid agonist, partial agonist or antagonist treatment immediately prior to admission; pregnant or lactating; physically dependent on alcohol or sedative hypnotics; and poor oral health (i.e. active aphthous stomatitis, active oral herpes, tongue or mouth piercing, or requiring immediate dental attention). This last condition was included because the main clinical trial involved sublingual drug administration.

The Johns Hopkins Institutional Review Board approved this study and all participants provided written, informed consent. Subjects in the present analysis were primarily male (74%), Caucasian (61%) and had a mean age of 41.7 years. The primary opioid abused by subjects was either heroin (mean use 6 years, SD 9.5) or prescription opioids (mean use 4.7 years, SD 4) prior to study entry, and all subjects had been using opioids daily (96%) or near daily (at least 16 days; 4%) in the 30 days before study entry. Forty-nine subjects initially enrolled; the present report includes the 46 who completed both the placebo and naloxone challenges. Two volunteers withdrew for non-study-related personal reasons after one challenge session, and one participant withdrew after experiencing a panic attack during the naloxone challenge session. Additionally, two participants had their naloxone sessions stopped after 30 min for excessive withdrawal symptoms.

2.2. Morphine stabilization phase and description of challenge sessions

Participants were screened on an outpatient basis and then admitted to the research unit where they were stabilized on 30 mg of subcutaneously administered morphine given four times daily (120 mg/day) for 2–8 days prior to the challenge

Table 1

Comparison of COWS and CINA item content and scoring.

Sign or symptom	Subjective vs. objective ^a	Possible scores	
		COWS ^b	CINA ^b
Anxiety or irritability	S	0–2, 4	–
Temperature changes	S	–	0–2
GI upset, including abdominal pain ^c	S/O	0–3, 5	0, 2, 4, 6 ^c
Restlessness	S/O	0, 1, 3, 5	0–2 ^c
Bone or joint aches	S/O	0–2, 4	0–2
Sweating	S/O	0–4	0–3
Runny nose or tearing ^c	O	0–2, 4	0–2 ^c
Tremor	O	0–2, 4	0–3
Gooseflesh	O	0, 3, 5	0–3
Yawning	O	0–2, 4	0–2
Pupil size	O	0–2, 5	–
Pulse rate	O	0–2, 4	Pulse/10
Systolic blood pressure (SBP)	O	–	SBP/10
Maximum possible score		48	30 + pulse/10 and SBP/10

^a Indicates whether item is a subjective symptom (S) or an objective sign (O). Those listed as (S/O) indicate that an item score includes assessment of both signs and symptoms, with lower scores for subjective symptom report and higher scores for objective signs.

^b Entries are possible scores on each item for each instrument. If no score is shown, the scale does not include that item.

^c The CINA contains two separate scores for these items, whereas the COWS has only one.

sessions (mean 4.4 days, SD 1.3). After stabilization, participants received intramuscularly administered injections of placebo and 0.4 mg naloxone in a randomized, double-blind fashion in two sessions separated by at least 24 h. Withdrawal assessments and drug effect rating scales, vital signs, and pupil measurements were collected every 15 min, starting 30 min pre- and through 150 min post-injection, except for the time of drug administration (time 0). Trained research assistants, who were present during the entire session, collected the data and administered the scales.

2.3. Measurements

Withdrawal measurements consisted of the CINA, COWS, and VAS self-report items.

2.3.1. CINA and COWS

The item content and scoring of the CINA and COWS are summarized in Table 1. There is substantial overlap in content, but each scale also includes items absent from the other. The CINA consists of 13 items – 1 purely subjective symptom item, 7 purely objective sign items, and 5 items that included subjective and objective components. The COWS consists of 11 items – 1 purely subjective symptom item, 6 objective sign items, and 4 items that included subjective and objective components. Item scoring options were specified differently for the two scales, but each scale summed the scores of its items to produce a total score. The COWS provided instructions for categorical ratings of pupil size and pulse, including an option for a zero score. On the CINA, the heart rate and blood pressure items ensured a minimum score of approximately 20 even in the absence of any withdrawal.

2.3.2. VAS

Visual analog scales (VASs) were single-item questions that assess subjective drug effects at the time of scale completion (Preston et al., 1988). Ratings were completed on a computer; using a mouse, the subject positioned an arrow along a 100-point line marked at either end with “none” and “extremely.” VAS items in the present study were: “Do you feel any DRUG EFFECT?”, “Does the drug have any GOOD EFFECTS?”, “Does the drug have any BAD EFFECTS?”, “How HIGH are you?”, “Does this drug make you feel SICK?”, and “Do you LIKE the drug?”

2.3.3. Pupil diameter

Pupil diameter was assessed with a digital pupillometer (Neuroptics, Inc.) in constant room lighting. The measurements provided by the pupillometer were also used for the pupil score in the COWS, which required observers to categorize the pupil diameter (Table 1).

2.4. Statistical analysis

Mean scores and standard deviations were calculated for each time point in the naloxone challenge sessions using SASTM software, Ver. 9.1 (SAS Institute, 2003).

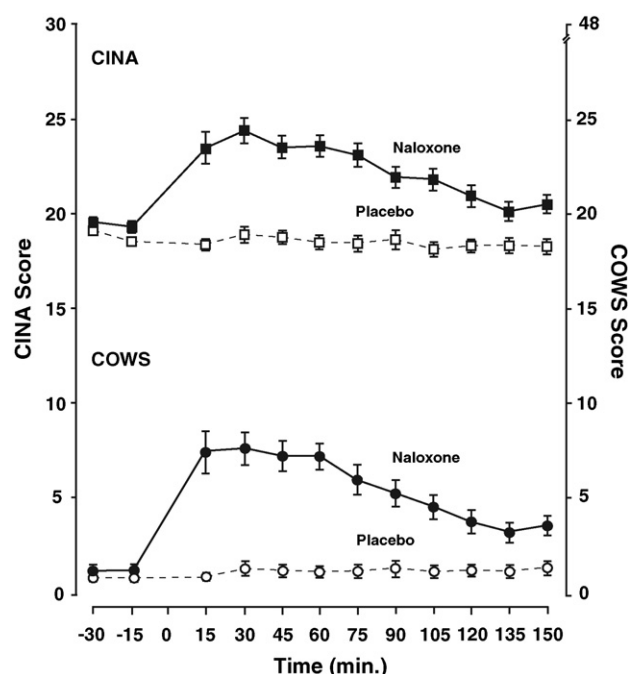


Fig. 1. Mean COWS and CINA scores (\pm SEM) vs. time. The Clinical Opiate Withdrawal Scale (COWS) and Clinical Institute Narcotic Assessment (CINA) had peak scores in subjects on average 30 min post-naloxone injection (\pm SEM). The two scales show a similar time course for withdrawal signs and symptoms.

Repeated measures regressions were used to assess significant differences on the separate opioid withdrawal measurements, using drug (naloxone vs. placebo), time, and drug-by-time interaction terms. All the rest of the statistical calculations used SPSS 16.0 (SPSS Inc., Chicago, IL). Concurrent validity was assessed using Pearson's correlation coefficients calculated between peak CINA, COWS, VAS items, and pupil diameter during the naloxone challenge session. Correlations of time to peak on the different measures were similarly calculated. Internal consistency reliability of the 11 COWS items was assessed with Cronbach's alpha statistic. Lastly, inter-item correlation matrices were created to describe the association of individual COWS and CINA items to each other and to the total scale score.

3. Results

3.1. CINA vs. COWS

Overall, the COWS and CINA scales were very similar in terms of both the magnitude and time course of their withdrawal score changes in the naloxone challenge session, demonstrating the concurrent validity of the COWS. Fig. 1 shows the mean scores and standard errors (SEM) of the COWS and CINA graphed vs. time.

The mean peak COWS (7.6) and CINA (24.4) scores occurred on average at 30 min post-injection, which is within the expected time range of peak withdrawal following intramuscular naloxone (Daftery, 1974; Wang et al., 1974; Judson et al., 1980). Additionally, time to peak (TTP) analysis revealed positive correlation between COWS TTP and CINA TTP ($r=0.66$, $p<0.0001$). Repeated measure regression analysis revealed statistically significant effects on COWS for drug (naloxone vs. placebo) ($F=79.3$, $df=45$, $p<0.0001$), time ($F=15.03$, $df=495$, $p<0.0001$), and drug-by-time interaction ($F=13.82$, $df=476$, $p<0.0001$). There were also significant effects for the above three analyses on the CINA ($F=77.4$, $df=45$, $p<0.0001$), ($F=10.35$, $df=495$, $p<0.0001$), and ($F=10.94$, $df=477$, $p<0.0001$), respectively. Table 2 shows a strong positive correlation between peak COWS and CINA scores ($r=0.85$, $p<0.001$) during the naloxone challenge session.

Table 2 also shows the effect of omitting various physiological measurement items from the CINA and COWS. For the COWS, removing the pupil diameter item (COWS noPUP), heart rate item (COWS noHR), and both of these measures (COWS noPHYS) still leave these modified COWS scores highly correlated with the CINA, indicating that these items may not be needed to detect this level of opioid withdrawal with the COWS. The score on the COWS heart rate item is 0 (<80), 1 (81–100), 2 (101–120), or 4 (>120). In this sample, subjects had little change in heart rate during the naloxone session (peak 7.5 bpm change from baseline); therefore, this item rarely affected the total COWS score which may explain high similarity in correlation coefficients between COWS vs. CINA and COWS without the heart rate vs. CINA. Similarly, correlations between the CINA without heart rate item (CINA noHR), systolic blood pressure score (CINA noBP), or both of these measurements (CINA noPHYS) were highly correlated with the CINA total score, as well as the COWS total score and the various modified versions of the COWS.

3.2. CINA and COWS vs. VAS

Two VAS items, bad effects and sick, showed a similar time course but greater variability in mean score than the CINA (Fig. 2). VAS mean peak scores for bad effects and sick occurred on average somewhat later than the CINA peak, with the peak score occurring at 60 min for bad effects (score=33.2) and 45 min for sick (score=28.1). The VAS time course of a rapid increase in scores after injection and then a gradual decline over 2.5 h was very similar to the time course seen with the CINA and COWS (Fig. 1). Results from repeated measures regression revealed statistically significant effects for drug condition, time, and drug-by-time interaction on these two VAS items ($p<0.0001$ in all cases). Correlation analysis showed moderately good association between peak CINA and bad effects ($r=0.63$, $p<0.001$) as well as sick ($r=0.65$, $p<0.001$)

Table 2

Correlation matrix for select opiate withdrawal assessment tools.

	COWS	COWS noPUP	COWS noHR	COWS noPHYS	CINA	CINA noHR	CINA noBP	CINA noPHYS	Bad effects	Sick
COWS	1	***	***	***	***	***	***	***	***	***
COWS noPUP	0.98	1	***	***	***	***	***	***	***	***
COWS noHR	1	0.98	1	***	***	***	***	***	***	***
COWS noPHYS	0.98	1	0.98	1	***	***	***	***	***	***
CINA	0.85	0.84	0.84	0.82	1	***	***	***	***	***
CINA noHR	0.85	0.85	0.84	0.84	0.98	1	***	***	***	***
CINA noBP	0.88	0.88	0.87	0.86	0.96	0.92	1	***	***	***
CINA noPHYS	0.89	0.90	0.89	0.89	0.93	0.94	0.98	1	***	***
Bad effects	0.57	0.57	0.56	0.56	0.63	0.59	0.64	0.62	1	***
Sick	0.57	0.59	0.56	0.58	0.65	0.64	0.64	0.64	0.88	1

COWS=Clinical Opiate Withdrawal Scale. COWSnoPUP=COWS with pupil diameter categorization score removed. COWSnoHR=COWS with heart rate categorization score removed. COWSnoPHYS=COWS with both pupil diameter and heart rate scores removed. CINA=Clinical Institute Narcotic Assessment. CINAnoHR=CINA with the heart rate score removed. CINAnoBP=CINA with the systolic blood pressure score removed. CINAnoPHYS=CINA with both heart rate and blood pressure scores removed. This matrix shows a strong correlation between COWS and CINA and suggests that certain objective measurements of withdrawal, i.e. heart rate, blood pressure, and pupil size, could be omitted without losing the ability to detect opiate withdrawal. All values are significant at $p<0.001$.

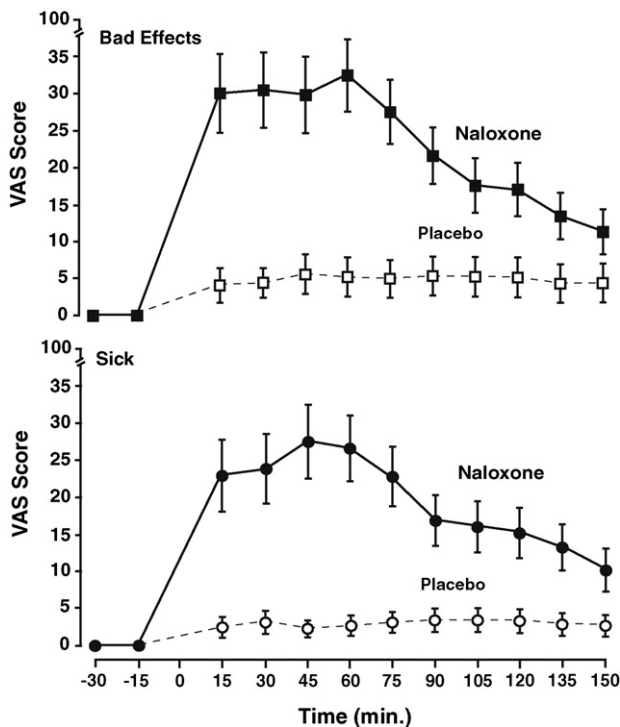


Fig. 2. Mean bad effects and sick effects VAS items (\pm SEM) vs. time. These graphs show the average (\pm SEM) time course of subjective response on two Visual Analogue Scale (VAS) questions: (upper panel) “Does the drug have any BAD EFFECTS?” and (lower panel) “Does this drug make you feel SICK?” The over all time courses for both VAS items follow a similar course as the COWS and CINA mean scores seen in Fig. 1.

(Table 2). Correlations between peak VAS and COWS were slightly lower; however, there was a strong correlation between the two VAS items ($r=0.88$, $p<0.001$). Lastly, no significant correlations were seen between peak CINA or COWS scores and VAS ratings for good effects, drug liking, or high.

3.3. CINA and COWS vs. quantitative pupil measurements

The time course of pupil diameter change showed a mean peak increase (1.03 mm) that occurred 15 min after naloxone injection, followed by gradual return to baseline. Pupil diameter showed <0.27 mm change from baseline in the placebo session. Repeated measures regression using drug condition, time, and drug-by-time showed significance ($p<0.001$) in each analysis. Maximum pupil diameter and peak CINA and COWS scores showed a modest correla-

tion ($r=0.39$, $p=0.01$ and $r=0.36$, $p=0.01$). There was no significant correlation between maximum pupil diameter and bad effects or Sick VAS.

3.4. Internal consistency and inter-item correlations

Analysis of the internal consistency for the eleven COWS items revealed an overall Cronbach's alpha of 0.78, indicating good reliability. As well, there was surprisingly little inter-item correlation between individual COWS items (Table 3). Only combinations of restlessness and anxiety/irritability (0.67) as well as runny nose/tearing and yawning (0.54) were significantly correlated. A similar inter-item correlation matrix for the CINA revealed that the objective physiological measurements did not correlate well with the total CINA score, and similar items showed significant inter-item correlations as with the COWS (Table 4). Finally, the VAS items correlated with the total COWS and CINA scores about as well as did the individual items constituting the scales (Table 2).

3.5. Analysis of atypical subjects

Seven individuals did not differentiate between the effects of placebo vs. naloxone based upon VAS scores of bad effects. Of these individuals, four had opioid withdrawal (COWS scores ≥ 5) with both placebo and naloxone; two had no withdrawal in either session (COWS <5); and one person had mild withdrawal (COWS score of 6 two hours after injection) with naloxone only. There were no significant differences in demographic or history characteristics that explained those who did or did not differentiate naloxone from placebo. When these individuals were removed, no significant changes occurred in correlation, repeated measures regression, or time to peak analyses.

4. Discussion

The accurate and rapid assessment of opioid withdrawal is important in the clinical management of opioid-dependent patients in both inpatient and outpatient settings. As well, U.S. guidelines for opioid treatment require clinical evidence of dependence in patients, which may include the presence of withdrawal (SAMHSA, 2001). Likewise, office-based outpatient treatment requires a medical professional to assess opioid withdrawal when initiating treatment with buprenorphine or buprenorphine/naloxone (Center for Substance Abuse Treatment, 2004). The present analyses provide validation of a short, easy-to-use scale for withdrawal (COWS) as well as quantification of the relationship of that scale to the CINA and single-item VAS indices of withdrawal. Our results demonstrate that the COWS correlates well with the previously validated

Table 3
Inter-item correlations amongst COWS items.

Item	1	2	3	4	5	6	7	8	9	10	11	12
1. Pulse	1.00	****	****	****	****	****	****	****	****	****	****	****
2. GI upset	0.05 [*]	1.00	****	****	****	****	****	****	****	****	****	****
3. Sweating	0.06 [*]	0.27	1.00	****	****	****	****	****	****	****	****	****
4. Tremor	0.11	0.14	0.10	1.00	****	****	****	****	****	****	****	****
5. Restlessness	-0.03 [*]	0.33	0.35	0.13	1.00	****	****	****	****	****	****	****
6. Yawning	0.09 [*]	0.28	0.34	0.16	0.40	1.00	****	****	****	****	****	****
7. Pupil size	0.27	0.10	0.19	0.10	0.15	0.33	1.00	****	****	****	****	****
8. Anxiety/irritability	0.06 [*]	0.42	0.31	0.14	0.67	0.40	0.17	1.00	****	****	****	****
9. Bone/joint aches	0.10	0.22	0.18	0.11	0.18	0.08 [*]	0.09 [*]	0.33	1.00	****	****	****
10. Gooseflesh	0.12	0.29	0.40	0.01 [*]	0.33	0.47	0.24	0.46	0.17	1.00	****	****
11. Runny nose/tearing	0.12	0.28	0.24	0.10	0.41	0.54	0.32	0.47	0.23	0.39	1.00	****
12. Total COWS score	0.24	0.52	0.55	0.26	0.66	0.75	0.53	0.72	0.35	0.69	0.69	1.00

Both the COWS and CINA inter-item correlation matrices were based on 534 observations from the 46 subjects over the entire naloxone challenge session. The overall pattern of correlations did not change appreciably when based on just the $N=46$ individual observations at peak opioid withdrawal.

^{*} $p \geq 0.05$. All others $p < 0.05$. Bolded items indicate strong correlation ($r \geq 0.5$).

Table 4
Inter-item correlations amongst CINA items.

Item	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Abdominal changes	1.00	****	****	****	****	****	****	****	****	****	****	****	****	****
2. Temperature changes	0.27	1.00	****	****	****	****	****	****	****	****	****	****	****	****
3. Nausea/vomiting	0.35	0.33	1.00	****	****	****	****	****	****	****	****	****	****	****
4. Muscle aches	0.19	0.21	0.22	1.00	****	****	****	****	****	****	****	****	****	****
5. Gooseflesh	0.19	0.49	0.24	0.16	1.00	****	****	****	****	****	****	****	****	****
6. Nasal congestion	0.18	0.36	0.17	0.14	0.36	1.00	****	****	****	****	****	****	****	****
7. Restlessness	0.05 [*]	0.28	0.28	0.14	0.35	0.26	1.00	****	****	****	****	****	****	****
8. Tremor	0.10	0.11	0.13	0.11	0.10	−0.03 [*]	0.07 [*]	1.00	****	****	****	****	****	****
9. Lacrimation	0.14	0.40	0.25	0.21	0.50	0.60	0.36	0.08 [*]	1.00	****	****	****	****	****
10. Sweating	0.19	0.31	0.24	0.15	0.29	0.21	0.20	0.11	0.22	1.00	****	****	****	****
11. Yawning	0.13	0.41	0.23	0.15	0.45	0.54	0.31	0.08 [*]	0.65	0.23	1.00	****	****	****
12. HR/10	0.11	0.19	0.01 [*]	0.09	0.20	0.20	−0.04 [*]	0.10	0.15	0.01 [*]	0.19	1.00	****	****
13. SBP/10	0.09	0.07 [*]	−0.02 [*]	−0.11	0.05 [*]	−0.07 [*]	0.04 [*]	0.11	0.10	0.05 [*]	0.03 [*]	−0.13	1.00	****
14. CINA Total	0.43	0.64	0.52	0.35	0.63	0.55	0.49	0.28	0.68	0.43	0.64	0.37	0.35	1.00

The items with highest correlation with total CINA score are temperature changes, gooseflesh, lacrimation, and yawning, as compared to restlessness, tremor, lacrimation, and nausea/vomiting in the original CINA validation study (Peachey and Lei, 1988).

^{*} $p \geq 0.05$. All others are $p < 0.05$. Bolded items indicate strong correlation ($r \geq 0.5$).

CINA scale in the context of a standardized naloxone challenge in opioid-dependent persons. The time course of withdrawal as measured by the COWS was congruent with the pharmacologic properties of naloxone. Finally, the overlap in content of the two scales (Table 1) supports the content validity and face validity of the COWS.

Internal consistency of the COWS was high, demonstrating the scale was reliable in measuring the construct of opioid withdrawal. Inter-item correlations indicated little item overlap, providing evidence of content validity (measuring a broad range of symptoms). There was a high degree of consistency across opioid withdrawal measures in terms of identifying and tracking the syndrome over time, demonstrating concurrent validity of these measures. The time course for COWS and CINA were remarkably consistent. The similarity to CINA time course was somewhat less for the two VAS items, but the overall trend of both measures was the same. As well, the variance in the mean scores was relatively minor, except for VAS, which as single items with a larger scale range would be expected to have greater variance. This larger variance of the single-item VAS scores was probably also related to subjects' understanding of the items, personality effects on expressing discomfort, or possibly demographic and history characteristics. Nevertheless, these single-item questions may have utility in following the progress of withdrawal distress and guiding its medical management. Given the strong correlation between CINA and COWS seen in Table 3, both scales are well suited for assessing and tracking opioid withdrawal. Modifying these scales to omit objective physiological indices may not affect each scale's utility in the discrimination of the level of opioid withdrawal or guiding its medical treatment. Therefore, non-medical staff could aid in the assessment of withdrawal, and time of medical staff could be freed up for other needs. However, the physiological measures do provide objective indices to supplement the otherwise subjective responses and could thereby assist the clinician in determining false positive withdrawal responses. The choice of assessment instrument should be determined by site-specific needs, including the probability of feigned responses and the desire for objective indices.

This study has several limitations. First, relatively mild opioid withdrawal was produced, most likely due to the combination of a low naloxone dose (0.4 mg) and a modest level of morphine physical dependence (120 mg/day). However, the recognition of more severe forms of opioid withdrawal is less ambiguous for most clinicians, and the more critical need is to have scales that are sensitive enough to distinguish mild but clinically significant withdrawal. The most useful aspect of an opioid withdrawal scale is in differentiating the presence vs. absence of withdrawal, which the COWS does. The original COWS authors (Wesson and Ling, 2003) had

specifically recommended that a validation be done for the low-end of the scale, which this study accomplished. Second, we did not assess external reliability of these measurements; we did not have multiple raters of the same sessions to include inter-rater reliability and did not perform multiple naloxone challenges to calculate test–retest reliability. This could be done in the future. Third, the modest correlation of COWS and CINA with pupil diameter is puzzling, given that mydriasis is a classic sign in opioid withdrawal (Himmelsbach, 1941). However, the measurement tool in this study may have affected this measure. The digital pupillometer decreased the ambient light reaching the eye by surrounding the eye before determining the pupil diameter; the intensity of lighting influences pupillary response to opioids (Weinhold and Bigelow, 1993). Fourth, this study included seven individuals who failed to distinguish placebo from naloxone. This did not change the overall results significantly but it does highlight two important points: prior literature has shown placebo can precipitate mild withdrawal in heroin-dependent individuals (Kanof et al., 1991) and not all opioid-dependent individuals respond to a naloxone challenge with signs or symptoms of withdrawal (Blachly, 1973; Wang et al., 1982). Finally, while these data document the sensitivity of these indices to opioid withdrawal, they do not address their specificity i.e. the extent to which they may be affected by factors other than opioid withdrawal.

Even with these limitations, the validation of the COWS and correlations with the other opioid withdrawal measurement tools provide useful information for future clinical evaluation of this syndrome. The COWS and CINA followed comparable trajectories for the time course of opioid withdrawal. The VAS bad effects and Sick single-item assessments followed a parallel time course for withdrawal, suggesting these easily administered scales might be useful in certain settings for identifying and following opioid withdrawal. Clinicians who are not worried about feigned responses might simply use these questions to screen quickly for withdrawal and treat where appropriate. In other settings, the COWS or CINA could be used for the identification of withdrawal (with or without objective sign measurement) and for monitoring response to treatment interventions. Having easy and reliable quantification has distinct advantages when following withdrawal and setting up treatment protocols based upon these findings.

In summary, this study shows that the COWS is a valid instrument with sufficient sensitivity to detect mild opiate withdrawal. It would therefore be expected to detect moderate to severe withdrawal. The COWS, as well as the VAS items reported here, have potential uses in inpatient and outpatient treatment, in detoxification, and in research protocols. Their brevity and ease of use make them good choices for use in all these settings.

Role of funding source

Reckitt Benckiser Pharmaceuticals Inc. provided support and medications for use in this study. As well, this company's representatives assisted in the preparation of the protocol and review of the manuscript. NIDA also provided funding but had no other role. Grants were R01 DA08045, K24 DA023186, and T32 DA07209.

Contributors

All authors have contributed to and approved the final manuscript. Dr. Tompkins wrote the first draft of the manuscript, contributed to interpreting data analysis and assisted in editing. Dr. Bigelow assisted in protocol development, interpreting data analyses, writing, and editing paper. Mr. Harrison conducted research, assisted in protocol development and manuscript editing. Drs. Johnson and Fudala assisted in protocol design and editing the manuscript. Dr. Strain assisted in protocol design, study execution, supervision, writing, and editing the manuscript.

Conflict of interest

Drs. Tompkins and Bigelow as well as Mr. Harrison have no conflict of interest to report. Drs. Johnson and Fudala are employees of Reckitt Benckiser Pharmaceuticals Inc., which is a maker of buprenorphine and provided the funding and medications for the clinical trial. Dr. Strain also is a paid consultant to Reckitt Benckiser Pharmaceuticals. The terms of this arrangement are being managed by the Johns Hopkins University in accordance with its conflict of interest policies.

Acknowledgements

We appreciate Linda Felch with her help analyzing the data and the many research assistants who ran the study sessions and assisted with data collection. We would especially like to thank Lillian Salinas who helped create the figures and tables.

References

American Psychiatric Association, 2000. *Diagnostic and Statistical Manual of Mental Disorders*, Text Revision, fourth ed. American Psychiatric Association, Washington, DC.

- Blachly, P.H., 1973. Naloxone for diagnosis in methadone programs. *JAMA* 224, 334–335.
- Bradley, B.P., Gossop, M., Phillips, G.T., Legarda, J.J., 1987. The development of an opiate withdrawal scale (OWS). *Br. J. Addict.* 82, 1139–1142.
- Center for Substance Abuse Treatment, 2004. *Clinical Guidelines for the Use of Buprenorphine in the Treatment of Opioid Addiction*. Substance Abuse and Mental Health Services Administration, Rockville, MD.
- Daftery, A.V., 1974. Letter: naloxone challenge in propoxyphene dependence. *N. Engl. J. Med.* 291, 979–980.
- Gossop, M., 1990. The development of a short opiate withdrawal scale (SOWS). *Addict. Behav.* 15, 487–490.
- Haertzen, C.A., Meketon, M.J., 1968. Opiate withdrawal as measured by the addiction research center inventory (ARCI). *Dis. Nerv. Syst.* 29, 450–455.
- Haertzen, C.A., Meketon, M.J., Hooks Jr., N.T., 1970. Subjective experiences produced by the withdrawal of opiates. *Br. J. Addict. Alcohol Other Drugs* 65, 245–255.
- Handelsman, L., Cochrane, K.J., Aronson, M.J., Ness, R., Rubinstein, K.J., Kanof, P.D., 1987. Two new rating scales for opiate withdrawal. *Am. J. Drug Alcohol Abuse* 13, 293–308.
- Himmelsbach, C.K., 1941. The morphine abstinence syndrome, its nature and treatment. *Ann. Intern. Med.* 15, 829–839.
- Isbell, H., 1950. Manifestations and treatment of addiction to narcotic drugs and barbiturates. *Med. Clin. North Am.* 34, 425–438.
- Judson, B.A., Himmelberger, D.U., Goldstein, A., 1980. The naloxone test for opiate dependence. *Clin. Pharmacol. Ther.* 27, 492–501.
- Kanof, P.D., Aronson, M.J., Ness, R., Cochrane, K.J., Horvath, T.B., Handelsman, L., 1991. Levels of opioid physical dependence in heroin addicts. *Drug Alcohol Depend.* 27, 253–262.
- Kolb, L., Himmelsbach, C.K., 1938. Clinical studies of drug addiction. III. A critical review of the withdrawal treatments with method of evaluating abstinence syndromes. *Am. J. Psychiat.* 94, 759–799.
- Peachey, J.E., Lei, H., 1988. Assessment of opioid dependence with naloxone. *Br. J. Addict.* 83, 193–201.
- Preston, K.L., Bigelow, G.E., Liebson, I.A., 1988. Butorphanol-precipitated withdrawal in opioid-dependent human volunteers. *J. Pharmacol. Exp. Ther.* 246, 441–448.
- SAS Institute, 2003. *SAS Software*. SAS Institute, Cary, NC.
- Substance Abuse and Mental Health Services Administration, 2001. 42 CFR Part 8: Opioid Drugs in Maintenance and Detoxification Treatment of Opiate Addiction; Final Rule. *Federal Register* 66, 4076–4102.
- Wang, R.I., Kochar, C., Hasegawa, A.T., Roh, B.L., 1982. Initial methadone dose in treating opiate addiction. *Int. J. Addict.* 17, 357–363.
- Wang, R.I., Wiesen, R.L., Lamid, S., Roh, B.L., 1974. Rating the presence and severity of opiate dependence. *Clin. Pharmacol. Ther.* 16, 653–658.
- Weinhold, L.L., Bigelow, G.E., 1993. Opioid miosis: effects of lighting intensity and monocular and binocular exposure. *Drug Alcohol Depend.* 31, 177–181.
- Committee on Treatment of Opioid Dependence, 1999. In: Wesson, D.R., Ling, W., Jara, G. (Eds.), *Buprenorphine in Pharmacotherapy of Opioid Addiction: Implementation in Office-Based Medical Practice; Translating the Experience of Clinical Trials into Clinical Practice*. California Society of Addiction Medicine, San Francisco, CA.
- Wesson, D.R., Ling, W., 2003. The clinical opiate withdrawal scale (COWS). *J. Psychoactive Drugs* 35, 253–259.