Development of Short-Form Versions of the Screener and Opioid Assessment for Patients with Pain-Revised (SOAPP-R)

A Proof-of-Principle Study

Finkelman, M.D.; Smits, N.; Kulich, R.J.; Zacharoff, K.L.; Magnuson, B.E.; Chang, H.; Dong, J.; Butler, S.F.

DOI
10.1093/pm/pnw210

Publication date
2016

Document Version
Final published version

Published in
Pain Medicine

Pain Medicine 2017; 18: 1292–1302
doi: 10.1093/pm/pnw210

OXFORD

Original Research Article

Development of Short-Form Versions of the Screener and Opioid Assessment for Patients with Pain-Revised (SOAPP-R): A Proof-of-Principle Study

Matthew D. Finkelman, PhD,* Niels Smits, PhD,† Ronald J. Kulich, PhD,‡,§ Kevin L. Zacharoff, MD,¶ Britta E. Magnuson, DMD,‖ Hong Chang, PhD,‖‖ Jinghui Dong, PhD, and Stephen F. Butler, PhD¶**

*Department of Public Health and Community Service, Tufts University School of Dental Medicine, Boston, Massachusetts, USA; ‡Craniofacial Pain and Headache Center, Tufts University School of Dental Medicine, Boston, Massachusetts, USA; ‖Department of Diagnostic Sciences, Tufts University School of Dental Medicine, Boston, Massachusetts, USA; †Research Institute of Child Development and Education, University of Amsterdam, Amsterdam, The Netherlands; §Department of Anesthesia, Critical Care and Pain Medicine, Massachusetts General Hospital, Boston, Massachusetts, USA; ¶Inflexxion, Inc, Newton, Massachusetts, USA; ‖‖Tufts University School of Medicine, Boston, Massachusetts, USA; **Sackler School of Graduate Biomedical Sciences, Boston, Massachusetts, USA

Correspondence to: Matthew D. Finkelman, PhD, Department of Public Health and Community Service, Tufts University School of Dental Medicine, 1 Kneeland Street, Boston, MA 02111, USA. Tel: 617-636-3449; Fax: 617-636-6511; E-mail: matthew.finkelman@tufts.edu.

Funding sources: Data used in this study were collected as part of a NIDA grant (Grant no: DA015617, P.I. Butler). Data for this study were provided by Inflexxion, Inc. Research reported in this publication was supported by the National Institute on Drug Abuse of the National Institutes of Health under Award Number R03DA036683.

Disclosure and conflicts of interest: KLZ and SFB are employees of Inflexxion, Inc. Inflexxion holds the copyright for the Screener and Opioid Assessment for Patients with Pain–Revised (SOAPP-R).

Disclosure: The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Abstract

Background. The Screener and Opioid Assessment for Patients with Pain–Revised (SOAPP-R) is a 24-item questionnaire designed to assess risk of aberrant medication-related behaviors in chronic pain patients. The introduction of short forms of the SOAPP-R may save time and increase utilization by practitioners.

Objective. To develop and evaluate candidate SOAPP-R short forms.

Design. Retrospective study.

Setting. Pain centers.

Subjects. Four hundred and twenty-eight patients with chronic noncancer pain.

Methods. Subjects had previously been administered the full-length version of the SOAPP-R and been categorized as positive or negative for aberrant medication-related behaviors via the Aberrant Drug Behavior Index (ADBI). Short forms of the SOAPP-R were developed using lasso logistic regression. Sensitivity, specificity, and area under the curve (AUC) of all forms were calculated with respect to the ADBI using the complete data set, training-test analysis, and 10-fold cross-validation. The coefficient alpha of each form was also calculated. An external set of 12 pain practitioners reviewed the forms for content.

Results. In the complete data set analysis, a form of 12 items exhibited sensitivity, specificity, and AUC greater than or equal to those of the full-length SOAPP-R (which were 0.74, 0.67, and 0.76, respectively). The short form had a coefficient alpha of 0.76. In the training-test analysis and 10-fold cross-validation, it exhibited an AUC value within 0.01 of that of the full-length SOAPP-R. The majority of external practitioners reported a preference for this short form.

Conclusions. The 12-item version of the SOAPP-R has potential as a short risk screener and should be tested prospectively.

Key Words. Chronic Pain; Opioids; SOAPP-R; Substance Abuse; Short Form; Risk Stratification

Introduction

Chronic pain is a condition that affects as many as 100 million Americans, with access to adequate care remaining an ongoing public health concern [1–5]. Opioids have been considered an option for the patient; these medications have advantages in certain situations [6,7], but they can lead to possible negative effects [8–11]. Patients may exhibit drug-seeking behaviors, such as visiting multiple providers for prescriptions or misusing the medications by taking nonprescribed doses at more frequent intervals [11,12]. When prescribing opioids, providers are now being asked to carefully consider the patient's risk, while the tools to assist the clinician in conducting a comprehensive risk assessment remain lacking [13].

Validated questionnaires can assist in the screening of patients for possible future aberrant medication-related behaviors, with the original Screener and Opioid Assessment for Patients with Pain (SOAPP) [14–16] being among the most well known. As the SOAPP's items were conceptually derived [14], a more empirically based questionnaire was needed, leading to the development of the 24-item Screener and Opioid Assessment for Patients with Pain–Revised (SOAPP-R) [17]. The latter questionnaire has been validated [17] and cross-validated [18].

One important practical concern shared by the SOAPP and the SOAPP-R, as well as by other questionnaires, is their length. Questionnaires with many items take time not only for the patients themselves but also add to the administrative burden incurred by providers. With increasingly limited funds and restricted provider time, there is a critical need for efficiency in the health care setting [19]. To save time and increase utilization by providers, researchers have studied different approaches to reducing test length. In the current context, a static short form has been introduced [20], and a "proof of principle" has been established for a computerized variable-length version of the SOAPP-R that reduces its average number of items administered [21]. However, each of these approaches has a drawback. First, the existing static short form is based on the original SOAPP, not the more rigorously developed SOAPP-R. Only two of the short form's five items were selected for the SOAPP-R. Regarding the variable-length version of the SOAPP-R, this version requires that testing be conducted by computer, which limits its utility. Indeed, in some assessment settings, the infrastructure for computerized testing may not be present, in which case a short instrument that can be administered via paper-and-pencil is more practical.

The primary aim of this research was to develop candidate SOAPP-R short forms and evaluate them statistically in comparison with the full-length SOAPP-R. A secondary aim was to obtain feedback on the content of the different forms from an external set of evaluators. Comparisons between forms were made in order to recommend a single short form to be cross-validated in further study.

Methods

The Institutional Review Board at Tufts Medical Center and Tufts University Health Sciences Campus granted exempt status for this research project.

Subjects

The data set used in this retrospective study had previously been employed to develop a computerized variable-length version of the SOAPP-R [21]. The data set included item responses from 428 subjects who had completed the full-length version of the SOAPP-R, either as part of the SOAPP-R's original validation study (N = 207) or its cross-validation study (N = 221). Each patient in the original validation study was being treated at a pain clinic in Massachusetts, Ohio, or Pennsylvania. Each patient in the cross-validation study was being treated at a pain management center in Indiana, Massachusetts, New Hampshire, Pennsylvania, or Ohio. All patients had been prescribed opioids for chronic noncancer pain at their time of recruitment. Patients signed an informed consent form and were notified that study outcomes would not become part of their clinical record. Procedures of the studies were approved by participating centers' human subjects committees.

The SOAPP-R

Table 1 shows the 24 self-report items comprising the SOAPP-R. Each of these items is scored on a five-point scale (0 = "never," 1 = "seldom," 2 = "sometimes," 3 = "often," 4 = "very often"). A respondent's total score on the SOAPP-R is calculated by summing the scores of the individual items, after which the total score may be compared with a prespecified cut-off value. A positive result on the SOAPP-R (indicating greater risk) is obtained if the total score meets or exceeds the cut-off

Table 1 Results of the item analysis (N = 428); an "X" indicates that a given item is included in a given short form

Item	Group	% never	% seldom	% sometimes	% often	% very often	AUC for ADBI	SOAPP -R-7	SOAPP -R-9	SOAPP -R-10	SOAPP -R-11	SOAPP -R-12
1. How often do you have mood swings?	ADBI negative	8	31	37	15	8	0.62					
	ADBI positive	1	20	40	26	12						
2. How often have you felt a need for higher doses of medication to treat your pain?	ADBI negative	14	26	40	16	5	0.64				X	X
	ADBI positive	5	16	39	30	10						
3. How often have you felt impatient with your doctors?	ADBI negative	30	34	26	7	3	0.63		X	X	X	X
	ADBI positive	12	35	32	14	6						
4. How often have you felt that things are just so overwhelming that you can't handle them?	ADBI negative	29	32	25	11	4	0.66	X	X	X	X	X
	ADBI positive	10	28	31	21	10						
5. How often is there tension in the home?	ADBI negative	26	39	26	6	4	0.67	X	X	X	X	X
	ADBI positive	7	34	33	17	9						
6. How often have you counted pain pills to see how many are remaining?	ADBI negative	37	36	19	6	1	0.62					
	ADBI positive	26	28	25	15	6						
7. How often have you been concerned that people will judge you for taking pain medication?	ADBI negative	46	22	21	9	2	0.61					
	ADBI positive	30	21	23	19	6						
8. How often do you feel bored?	ADBI negative	29	33	25	10	3	0.61					
	ADBI positive	16	28	34	19	4						
9. How often have you taken more pain medication than you were supposed to?	ADBI negative	54	33	10	2	1	0.68	X	X	X	X	X
	ADBI positive	26	39	23	9	3						
10. How often have you worried about being left alone?	ADBI negative	59	23	12	4	2	0.59					
	ADBI positive	46	21	22	8	4						
11. How often have you felt a craving for medication?	ADBI negative	61	22	14	2	1	0.61					
	ADBI positive	43	27	19	8	3						
12. How often have others expressed concern over your use of medication?	ADBI negative	57	28	11	2	1	0.64	X	X	X	X	X
	ADBI positive	37	24	28	8	3						
13. How often have any of your close friends had a problem with alcohol or drugs?	ADBI negative	54	29	12	3	1	0.61					
	ADBI positive	37	32	19	6	6						
	ADBI negative	61	25	10	2	2	0.61					

(continued)

Table 1 Continued

Item	Group	% never	% seldom	% sometimes	% often	% very often	AUC for ADBI	SOAPP -R-7	SOAPP -R-9	SOAPP -R-10	SOAPP -R-11	SOAPP -R-12
14. How often have others told you that you had a bad temper?	ADBI positive	41	32	17	4	5						
15. How often have you felt consumed by the need to get pain medication?	ADBI negative	62	25	9	4	0	0.62					
	ADBI positive	41	31	19	8	2						
16. How often have you run out of pain medication early?	ADBI negative	69	22	8	1	0	0.67	X	X	X	X	X
	ADBI positive	37	34	18	8	3						
17. How often have others kept you from getting what you deserve?	ADBI negative	68	23	6	2	1	0.60					
	ADBI positive	49	34	11	3	3						
18. How often, in your lifetime, have you had legal problems or been arrested?	ADBI negative	76	22	2	0	0	0.57			X	X	X
	ADBI positive	62	30	6	1	1						
19. How often have you attended an AA or NA meeting?	ADBI negative	88	8	2	1	1	0.58	X	X	X	X	X
	ADBI positive	72	10	6	8	4						
20. How often have you been in an argument that was so out of control that someone got hurt?	ADBI negative	88	8	3	0	0	0.57					
	ADBI positive	74	18	6	1	1						
21. How often have you been sexually abused?	ADBI negative	89	7	4	0	0	0.59	X	X	X	X	X
	ADBI positive	71	17	6	3	3						
22. How often have others suggested that you have a drug or alcohol problem?	ADBI negative	89	8	2	1	0	0.60					X
	ADBI positive	70	20	7	3	1						
23. How often have you had to borrow pain medications from your family or friends?	ADBI negative	93	5	2	1	0	0.60					
	ADBI positive	73	20	3	2	1						
24. How often have you been treated for an alcohol or drug problem?	ADBI negative	94	5	1	1	0	0.57		X	X	X	X
	ADBI positive	81	12	6	1	0						

value; a negative result (indicating lower risk) is obtained if the total score is less than the cut-off value. In its original validation study [17], the SOAPP-R was found to exhibit adequate sensitivity (0.81) and specificity (0.68).

The Aberrant Drug Behavior Index

In addition to completing the full-length SOAPP-R, all 428 subjects in the study had completed Aberrant Drug Behavior Index (ADBI) assessments. The ADBI is a combination measure of multiple sources of information. The first such source is the Prescription Drug Use Questionnaire (PDUQ), a 20-minute interview that was derived from the American Society of Addiction Medicine's definition of addiction in patients with chronic pain [22]. The PDUQ consists of 42 items; a cut-off of 11 or greater was used previously [17,18] and was likewise used herein. The second source of information included in the ADBI is the Prescription Opioid Therapy Questionnaire (POTQ), an 11-item instrument that is administered to the patient's physician [11]. A cut-off of 2 or greater was used based on previous work [17,18]. Finally, the third element of the ADBI is a toxicology screen. Each patient's urine sample had been tested at a Quest Diagnostics lab for 1) illicit substances and 2) additional medications that had not been prescribed. A positive ADBI result was obtained if the result of the PDUQ was positive or if the result of the PDUQ was negative, but both the POTQ and toxicology screen were positive.

Short Form Development

One candidate short form was developed for each possible test length under 24 items. Selection of items was conducted using least absolute shrinkage and selection operator (lasso) logistic regression, with the result of the ADBI defined as the outcome to be predicted by the model. Lasso regression differs from standard multiple regression models in that it uses a penalty on the regression parameter values to force some of the predictors to have coefficients equal to zero [23]. Such predictors may be interpreted as having little ability to predict the outcome, given the other predictors. The larger the inputted value of the penalty, the more coefficients are required to take on a zero value. To produce a short form with a prespecified test length, the penalty was increased until the correct number of items was included in the form. Results were obtained using the glmnet package in R [24].

Once the items comprising a given short form were determined, a cut-off value associated with the form was defined in order to separate positive individuals from negative individuals. A classic method to select the cut-off is to use the Youden J index, in which the cut-off is chosen so that the quantity sensitivity + specificity – 1 is maximized. However, because the Youden J index gives equal weight to sensitivity and specificity, it does not account for the possibility that one of the two statistics (sensitivity or specificity) might be more critical than the other. Within the context of predicting aberrant drug-related behavior in patients with chronic pain, previous literature has clearly indicated the relative importance of sensitivity [14,17,18]. To reflect this emphasis on sensitivity, the particular cut-off value selected for usage in the current study was the one maximizing the Youden J index, among only the subset of cut-off values satisfying sensitivity $\geq$ specificity.

Short Form Evaluation

For each short form, as well as the full-length SOAPP-R, the receiver operating characteristic (ROC) curve was constructed using the ADBI as the outcome of interest, and the area under the curve (AUC) statistic was computed. Sensitivity and specificity were also calculated based on the cut-off value that was chosen via the procedure described in the previous section. Results of the different short forms were compared with one another, as well as with the full-length SOAPP-R. In addition, the coefficient alpha statistic was calculated for each form; a coefficient alpha of 0.70 or greater is considered to indicate sufficient internal consistency according to a well-known benchmark [25]. Each form's correlation with the short form of the Marlowe-Crowne Social Desirability Scale (MCSDS) was also computed to examine potential sensitivity to social desirability. A higher score on the MCSDS indicates a greater tendency to respond in a way that will be perceived as socially desirable. A lower score on a SOAPP-R form indicates lower risk (which would typically be perceived as socially desirable); hence, a correlation between a SOAPP-R form and the MCSDS that is closer to -1 suggests that the given SOAPP-R form is more sensitive to respondents' desire to appear socially desirable. On the other hand, a correlation between a SOAPP-R form and the MCSDS that is closer to zero suggests that the given SOAPP-R form is less sensitive to respondents' desire to appear socially desirable. Finally, the correlation between each short form and the full-length SOAPP-R was calculated using two methods: 1) the standard Pearson correlation between short form scores and full-length SOAPP-R scores and 2) the correlation obtained when applying Levy's [26] correction to account for the overlap of measurement errors between the short form and full-length form on common data. Bohlmeijer et al. [27] used a cut-off of 0.90 for the standard Pearson correlation and a cut-off of 0.80 for the corrected correlation as indicating substantial overlap between a short form and a full-length form.

The initial analysis involved selecting the items and the cut-off value for each short form based on the complete data set, then evaluating the performance of the short forms based on the same data set. One disadvantage of such an analysis is that when both model training (here, the selection of items and cut-off values) and evaluation are conducted on the same data, spuriously strong performance by the short forms may be observed [23]. To address the possibility of such inflated performance, two additional analyses were conducted

to obtain AUC, sensitivity, specificity, and coefficient alpha values with model training separated from evaluation. First, a "training-test" analysis was undertaken in which model training was performed for each short form using only the data from the original validation data set, and then the resulting short forms were evaluated using only the data from the cross-validation data set. Second, a "10-fold cross-validation" was undertaken in which the complete data set was randomly divided into 10 parts ("folds") of as equal size as possible. Model training was performed using a combined data set of nine of the 10 folds, after which evaluation of the resulting short forms was done on the remaining fold. Ten replications of this procedure were conducted so that each fold would serve as the evaluation data set exactly once. Results of the 10 replications were then aggregated.

Forms were also scrutinized in terms of their content. In particular, the content of each form that exhibited potential based on statistical characteristics was subjected to evaluation by an external set of 12 pain practitioners from a pain care center in Massachusetts. The practitioners included at least one of each of the following: pain physicians, pain psychologists, pain nurse practitioners, registered nurses, clinical social workers specializing in pain, and chiropractors. Feedback about the content of the forms was solicited in an open-ended format; additionally, participants were asked which form (including the full-length SOAPP-R) they would be most likely to use with their own patients. The option "none of the above" was presented as a possible answer choice for the latter question.

Item Evaluation

Item distributions were calculated separately for ADBI-negative individuals and ADBI-positive individuals. Each item's range of responses was examined based on these distributions. The AUC statistic was also computed for each item using the ADBI as the criterion to be predicted.

Results

Descriptive and demographic information was reported previously for this group of subjects [21]. The mean score on the full-length SOAPP-R was 20.4, with a standard deviation of 11.3. The ADBI was positive for 145 of the 428 subjects (34%). Four hundred twenty-six subjects had information on gender; among these, 243 were female (57%). Four hundred twenty-five subjects had information on age; among these, the mean age was 51.4 years, with a standard deviation of 13.0.

Table 1 displays the results of the item analysis. The entire range of response options was used for all items except item 24 ("been treated for an alcohol or drug problem"). Regarding the ability of items to predict ADBI status, the items with the highest AUC values were item 9 ("taken more pain medication than you were supposed to," AUC = 0.68), item 16 ("run out of pain medication early," AUC = 0.67), item 5 ("tension in the home," AUC = 0.67), and item 4 ("felt that things are just so overwhelming that you can't handle them," AUC = 0.66).

Table 2 presents results for the short forms and the full-length form, based on the analysis in which both model training and evaluation were conducted using the complete data set. The second column of Table 2 identifies the set of items that comprise each short form using their item numbers; see the first column of Table 1 for a list of the items along with their numbers. Table 2 shows that the aforementioned four items with the highest AUC values (items 9, 16, 5, and 4) were the first four items selected for the short forms by the lasso procedure; the item with the second-highest AUC value (item 16) was actually selected prior to the item with the highest AUC value (item 9). Turning to specific short forms, all forms between items 7 and 12 (inclusive) had AUC values higher than that of the full-length SOAPP-R (0.76) while exhibiting coefficient alpha values near or above the benchmark of 0.70. The eight-item form exhibited low specificity relative to the other forms, leading to the exclusion of this form from further consideration. All short forms of 13 items or more were also excluded from further consideration at this stage, given that they had lower AUC values than the nine-item, 10-item, and 12-item forms while being longer.

Also shown in Table 2 are several correlations of interest. As displayed in the eighth column of the table, the Pearson correlation between the full-length SOAPP-R and the MCSDS was –0.47. Correlations between the short forms and the MCSDS ranged between –0.47 and –0.18, suggesting reasonable neutrality with respect to social desirability. Additionally, all short forms of seven items or more exhibited a Pearson correlation over 0.90 with the full-length SOAPP-R. These same forms exhibited a correlation of 0.80 or higher with the full-length SOAPP-R after applying Levy's correction.

Based on the above analysis of the complete data set, five short forms were given further consideration via the training-test analysis and 10-fold cross-validation. These were the seven-item form, nine-item form, 10-item form, 11-item form, and 12-item form (hereafter the SOAPP-R-7, SOAPP-R-9, SOAPP-R-10, SOAPP-R-11, and SOAPP-R-12, respectively). The five short forms listed above will be referred to as the "short forms of interest;" the items comprising these specific forms are indicated in Table 1.

Tables 3 and 4 display the results of the training-test analysis and the 10-fold cross-validation, respectively. In the training-test analysis, the results regarding predictive validity (AUC, sensitivity, and specificity) generally exhibited shrinkage, both for the short forms and the full-length SOAPP-R. All short forms of interest had AUC values within 0.03 of that of the full-length SOAPP-R (0.74); the SOAPP-R-10 and SOAPP-R-12's values

Table 2 Results of the short forms and full-length SOAPP-R: complete data set analysis (N = 428)

No. of items	Item(s) selected	Cut-off	Sensitivity for ADBI	Specificity for ADBI	AUC for ADBI	Coefficient alpha	Pearson correlation with MCSDS	Pearson correlation with full-length SOAPP-R	Corrected correlation with full-length SOAPP-R
1	16	≥1	0.63	0.69	0.67	–	−0.18	0.60	–
2	16, 9	≥1	0.82	0.44	0.70	0.69	−0.21	0.70	0.66
3	16, 9, 5	≥3	0.72	0.61	0.73	0.62	−0.27	0.79	0.72
4	16, 9, 5, 4	≥4	0.79	0.54	0.74	0.69	−0.32	0.85	0.77
5	16, 9, 5, 4, 21	≥5	0.71	0.64	0.76	0.65	−0.33	0.87	0.77
6	16, 9, 5, 4, 21, 19	≥5	0.79	0.61	0.78	0.63	−0.35	0.89	0.78
7	16, 9, 5, 4, 21, 19, 12	≥6	0.72	0.66	0.78	0.69	−0.36	0.91	0.80
8	16, 9, 5, 4, 21, 19, 12, 3	≥6	0.85	0.53	0.78	0.71	−0.38	0.92	0.80
9	16, 9, 5, 4, 21, 19, 12, 3, 24	≥7	0.79	0.60	0.79	0.71	−0.39	0.92	0.80
10	16, 9, 5, 4, 21, 19, 12, 3, 24, 18	≥8	0.74	0.70	0.79	0.71	−0.40	0.92	0.80
11	16, 9, 5, 4, 21, 19, 12, 3, 24, 18, 2	≥10	0.73	0.70	0.78	0.74	−0.41	0.93	0.81
12	16, 9, 5, 4, 21, 19, 12, 3, 24, 18, 2, 22	≥10	0.74	0.69	0.79	0.76	−0.42	0.94	0.82
13	16, 9, 5, 4, 21, 19, 12, 3, 24, 18, 2, 22, 11	≥10	0.77	0.64	0.78	0.77	−0.42	0.95	0.83
14	16, 9, 5, 4, 21, 19, 12, 3, 24, 18, 2, 22, 11, 1	≥12	0.78	0.64	0.78	0.79	−0.43	0.96	0.83
15	16, 9, 5, 4, 21, 19, 12, 3, 24, 18, 2, 22, 11, 1, 7	≥14	0.71	0.69	0.78	0.80	−0.44	0.96	0.83
16	16, 9, 5, 4, 21, 19, 12, 3, 24, 18, 2, 22, 11, 1, 7, 10	≥14	0.75	0.65	0.77	0.81	−0.45	0.97	0.84
17	16, 9, 5, 4, 21, 19, 12, 3, 24, 18, 2, 22, 11, 1, 7, 10, 15	≥14	0.77	0.62	0.77	0.83	−0.45	0.98	0.85
18	16, 9, 5, 4, 21, 19, 12, 3, 24, 18, 2, 22, 11, 1, 7, 10, 15, 23	≥12	0.88	0.50	0.77	0.83	−0.45	0.98	0.85
19	16, 9, 5, 4, 21, 19, 12, 3, 24, 18, 2, 22, 11, 1, 7, 10, 15, 23, 17	≥14	0.81	0.58	0.77	0.84	−0.46	0.98	0.86
20	16, 9, 5, 4, 21, 19, 12, 3, 24, 18, 2, 22, 11, 1, 7, 10, 15, 23, 17, 6	≥16	0.74	0.66	0.77	0.85	−0.45	0.99	0.86
21	16, 9, 5, 4, 21, 19, 12, 3, 24, 18, 2, 22, 11, 1, 7, 10, 15, 23, 17, 6, 14	≥17	0.71	0.68	0.77	0.86	−0.47	0.99	0.86
22	16, 9, 5, 4, 21, 19, 12, 3, 24, 18, 2, 22, 11, 1, 7, 10, 15, 23, 17, 6, 14, 20	≥17	0.74	0.67	0.77	0.86	−0.47	0.99	0.87
23	16, 9, 5, 4, 21, 19, 12, 3, 24, 18, 2, 22, 11, 1, 7, 10, 15, 23, 17, 6, 14, 20, 13	≥17	0.77	0.63	0.77	0.87	−0.47	1.00	0.87
24	16, 9, 5, 4, 21, 19, 12, 3, 24, 18, 2, 22, 11, 1, 7, 10, 15, 23, 17, 6, 14, 20, 13, 8	≥19	0.74	0.67	0.76	0.87	−0.47	1	0.87

Table 3 Results of the short forms and full-length SOAPP-R: training-test analysis (N = 207 for training, N = 221 for evaluation)

No. of items	Sensitivity for ADBI	Specificity for ADBI	AUC for ADBI	Coefficient alpha
1	0.70	0.50	0.63	–
2	0.84	0.37	0.68	0.63
3	0.93	0.28	0.72	0.57
4	0.84	0.45	0.71	0.65
5	0.89	0.41	0.72	0.61
6	0.85	0.41	0.71	0.65
7	0.84	0.49	0.71	0.71
8	0.74	0.55	0.71	0.72
9	0.75	0.54	0.72	0.71
10	0.89	0.39	0.73	0.70
11	0.75	0.54	0.72	0.73
12	0.70	0.59	0.73	0.75
13	0.78	0.49	0.72	0.77
14	0.84	0.50	0.73	0.79
15	0.86	0.47	0.74	0.81
16	0.86	0.47	0.74	0.82
17	0.85	0.49	0.74	0.83
18	0.85	0.47	0.73	0.84
19	0.82	0.51	0.73	0.84
20	0.82	0.51	0.73	0.84
21	0.82	0.50	0.73	0.85
22	0.82	0.50	0.74	0.85
23	0.79	0.52	0.73	0.86
24	0.74	0.60	0.74	0.86

were within 0.01. All short forms of interest had coefficient alpha values at or above 0.70. In the 10-fold cross-validation, all short forms of interest, with the exception of the SOAPP-R-7, exhibited mean AUC values 0.01 lower than that of the full-length SOAPP-R (0.76). The mean coefficient alpha values of the SOAPP-R-7 and SOAPP-R-9 were below the benchmark of 0.70, whereas the SOAPP-R-10, SOAPP-R-11, and SOAPP-R-12's mean values were at or above this benchmark.

The evaluation of content by the external set of 12 pain practitioners included all five short forms of interest, as well as the full-length SOAPP-R. When asked which form they would be most likely to use with their own patients, one participant (8%) selected the SOAPP-R-11, nine participants (75%) selected the SOAPP-R-12, and two participants (17%) selected the full-length SOAPP-R (no participant selected the SOAPP-R-7, SOAPP-R-9, SOAPP-R-10, or "none of the above"). Among the nine participants who selected the SOAPP-R-12, seven directly referred to test length and/or respondent burden when explaining their choice. One additional participant who selected the SOAPP-R-12 did not directly allude to test length or respondent burden, but stated that the SOAPP-R-12 covers most of the basic concerns that he/she has in evaluating a patient for risk of aberrant behavior. The final participant who selected the SOAPP-R-12 did not explain his/her choice. The participant who selected the SOAPP-R-11 stated that this screener was the best combination of questions with minimal overlap. The two participants who selected the full-length SOAPP-R alluded to its being comprehensive/extensive when explaining their choice.

Discussion

When chronic opioid therapy is considered, formal risk stratification is increasingly becoming the standard of care. As multiple sources of data must be considered in assessing risk, time and clinician burden become critical factors in undertaking patient care. Each component of assessment should meet sufficient standards before considering its use within an evaluation protocol, particularly where the reliability and predictive validity of assessment instruments can be challenged in medico-legal settings [28].

Each of the candidate short forms considered in this study underwent multiple modes of evaluation. The SOAPP-R-12 exhibited strong performance in both the statistical analysis and the evaluation by the external set of pain practitioners. This short form exhibited statistical characteristics comparable to those of previously developed computer-based versions of the SOAPP-R with

Table 4 Results of the short forms and full-length SOAPP-R: 10-fold cross-validation (N = 428)

No. of items	Sensitivity for ADBI	Specificity for ADBI	Mean AUC for ADBI	Mean coefficient alpha
1	0.63	0.66	0.66	–
2	0.79	0.42	0.67	0.56
3	0.72	0.59	0.72	0.54
4	0.75	0.55	0.71	0.59
5	0.75	0.56	0.74	0.58
6	0.76	0.57	0.74	0.61
7	0.76	0.56	0.74	0.62
8	0.73	0.58	0.74	0.66
9	0.73	0.61	0.75	0.69
10	0.71	0.65	0.75	0.70
11	0.70	0.68	0.75	0.71
12	0.74	0.64	0.75	0.73
13	0.72	0.64	0.75	0.74
14	0.76	0.60	0.75	0.77
15	0.73	0.65	0.75	0.79
16	0.74	0.63	0.75	0.80
17	0.77	0.60	0.75	0.81
18	0.74	0.65	0.76	0.82
19	0.74	0.64	0.76	0.83
20	0.79	0.58	0.76	0.84
21	0.73	0.65	0.76	0.84
22	0.75	0.64	0.76	0.85
23	0.77	0.64	0.76	0.86
24	0.74	0.67	0.76	0.86

similar average test lengths [21]. Moreover, compared with the full-length SOAPP-R, the SOAPP-R-12 reduces the number of items by 50%, a figure that has previously been identified in other fields as constituting a significant decrement in test length [29,30]. Therefore, the results recommend the SOAPP-R-12 for the prediction of aberrant medication-related behaviors among chronic pain patients. Nevertheless, we emphasize that the current research represents only the first step in validating this form. Inferences from the current study are limited by its retrospective nature, particularly as the items comprising the SOAPP-R-12 (and other short forms) were administered as part of the full-length screener, rather than being administered as an intact set. As such, the results may be influenced by context effects and are not necessarily representative of the screening properties that would be obtained for an intact set. Future validation work should administer both the SOAPP-R-12 and the full-length form to the same participants, as the ultimate test of a screener is its performance when given prospectively in its operational format [31]. Testing should be performed in diverse settings, including both paper-based and computer-based administration when possible, in order to evaluate the generalizability of results. The cut-off value determined herein for the SOAPP-R-12 should also be validated, considering that sensitivities and specificities with respect to individual cut-off values demonstrated greater instability than the overall summary measure of discriminatory power, the AUC statistic. Hence, although the current study suggests the benefits of using the SOAPP-R-12, this form should not be considered fully validated unless the above steps are undertaken and the form's positive results are confirmed.

The validation cohort for the study included subjects who had previously been prescribed opioids for their pain. The study did not test whether the SOAPP-R forms are valid for predicting whether patients not on opioids would exhibit aberrant medication-related behaviors if opioids were prescribed to them. Therefore, the results cannot be generalized to opioid-naïve subjects. Nevertheless, the use of the obtained sample reflects real-world clinical practice in which a prescriber would want to consider the patient's risk for long-term opioid therapy.

Another important consideration that applies in a general screening context is that certain statistical properties of a screener are dependent on the prevalence of the outcome being predicted. In particular, an unduly low prevalence may contribute to a low positive predictive value, indicating that an unacceptably high percentage of positive findings are in fact false positives. Conversely, an unduly high prevalence may contribute to a low negative predictive value, indicating that an

unacceptably high percentage of negative findings are in fact false negatives. This phenomenon can occur even for screeners with acceptable sensitivity and specificity values [32]. In the current study, 34% of subjects had a positive result on the ADBI, suggesting that the prevalence of aberrancy was neither unduly high nor unduly low in the given population. Nevertheless, as the prevalence of aberrancy differs from population to population, the effect of prevalence on the screening characteristics of the SOAPP-R-12 (as well as the full-length SOAPP-R and other screeners) should be considered.

Finally, it should be emphasized that while the SOAPP-R was designed to include items that are subtle in terms of their scoring, and thus to be less prone to overt deception than the original SOAPP, the possibility of such deception is present in any screener regardless of length. Therefore, the SOAPP-R-12 should be considered as providing supplementary data that should be interpreted in the context of other information. Nevertheless, it may be useful in assisting providers with their assessments. Indeed, should the SOAPP-R-12 be validated in additional research, it may enhance the efficiency of screening and increase utilization.

Conclusions

The SOAPP-R-12 exhibits potential as a short screener to assess risk of aberrant medication-related behaviors among chronic pain patients. This form should be tested prospectively in future studies.

Acknowledgments

The authors are indebted to Jennifer Brownstein for assisting with the study.

References

1 Institute of Medicine (US) Committee on Advancing Pain Research, Care, and Education. Relieving pain in America: A blueprint for transforming prevention, care, education, and research. Washington, DC: National Academies Press; 2011.

2 Cahana A, Dansie EJ, Theodore BR, Wilson HD, Turk DC. Redesigning delivery of opioids to optimize pain management, improve outcomes, and contain costs. Pain Med 2013;14:36–42.

3 Gilson AM, Ryan KM, Joranson DE, Dahl JL. A reassessment of trends in the medical use and abuse of opioid analgesics and implications for diversion control: 1997–2002. J Pain Symptom Manage 2004;28:176–88.

4 Joranson DE, Ryan KM, Gilson AM, Dahl JL. Trends in medical use and abuse of opioid analgesics. JAMA 2000;283:1710–4.

5 SUPPORT Study Principal Investigators. A controlled trial to improve care for seriously ill hospitalized patients: The study to understand prognoses and preferences for outcomes and risks of treatments (SUPPORT). JAMA 1995;274:1591–8.

6 Chou R, Fanciullo GJ, Fine PG, et al. Clinical guidelines for the use of chronic opioid therapy in chronic noncancer pain. J Pain 2009;10:113–30.

7 Fine P, Webster L, Argoff C. American academy of pain medicine response to prop petition to the FDA that seeks to limit pain medications for legitimate noncancer pain sufferers. Pain Med 2012;13: 1259–64.

8 Atluri S, Sudarshan G, Manchikanti L. Assessment of the trends in medical use and misuse of opioid analgesics from 2004 to 2011. Pain Physician 2014; 17:E119–28.

9 Butler SF, Budman SH, Fernandez KC, et al. Development and validation of the current opioid misuse measure. Pain 2007;130:144–56.

10 Cheatle MD. Psychological dependence and prescription opioid misuse and abuse. Pain Med 2014; 15:541–3.

11 Michna E, Ross EL, Hynes WL, et al. Predicting aberrant drug behavior in patients treated for chronic pain: Importance of abuse history. J Pain Symptom Manage 2004;28:250–8.

12 Peirce GL, Smith MJ, Abate MA, Halverson J. Doctor and pharmacy shopping for controlled substances. Med Care 2012;50:494–500.

13 Friedman R, Li V, Mehrotra D. Treating pain patients at risk: Evaluation of a screening tool in opioid-treated pain patients with and without addiction. Pain Med 2003;4:182–5.

14 Butler SF, Budman SH, Fernandez K, Jamison RN. Validation of a screener and opioid assessment measure for patients with chronic pain. Pain 2004; 112:65–75.

15 Jamison RN, Link CL, Marceau LD. Do pain patients at high risk for substance misuse experience more pain? A longitudinal outcomes study. Pain Med 2009;10:1084–94.

16 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;10:1426–33.

17 Butler SF, Fernandez K, Benoit C, Budman SH, Jamison RN. Validation of the revised screener and opioid assessment for patients with pain (SOAPP-R). J Pain 2008;9:360–72.

18 Butler SF, Budman SH, Fernandez KC, Fanciullo GJ, Jamison RN. Cross-validation of a screener to predict opioid misuse in chronic pain patients (SOAPP-R). J Addict Med 2009;3:66–73.

19 Dugdale DC, Epstein R, Pantilat SZ. Time and the patient-physician relationship. J Gen Intern Med 1999;14(suppl 1):S34–40.

20 Koyyalagunta D, Bruera E, Aigner C, et al. Risk stratification of opioid misuse among patients with cancer pain using the SOAPP-SF. Pain Med 2013; 14:667–75.

21 Finkelman MD, Kulich RJ, Zacharoff KL, et al. Shortening the screener and opioid assessment for patients with pain-revised (SOAPP-R): A proof-of-principle study for customized computer-based testing. Pain Med 2015;16:2344–56.

22 Compton P, Darakjian J, Miotto K. Screening for addiction in patients with chronic pain and "problematic" substance use: Evaluation of a pilot assessment tool. J Pain Symptom Manage 1998; 16:355–63.

23 Hastie T, Tibshirani R, Friedman JH. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd edition. New York: Springer; 2009.

24 R Core Team. R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria, 2013. Available at: http://www.R-project.org/ (accessed January 5, 2015).

25 Putnam SP, Rothbart MK. Development of short and very short forms of the children's behavior questionnaire. J Pers Assess 2006;87:103–13.

26 Levy P. The correction for spurious correlation in the evaluation of short-form tests. J Clin Psychol 1967; 23:84–6.

27 Bohlmeijer E, ten Klooster PM, Fledderus M, Veehof M, Baer R. Psychometric properties of the five facet mindfulness questionnaire in depressed adults and development of a short form. Assessment 2011; 18:308–20.

28 Kulich RJ, Driscoll J, Prescott JC, et al. The daubert standard, a primer for pain specialists. Pain Med 2003;4:75–80.

29 Leidy NK, Knebel A. In search of parsimony: Reliability and validity of the functional performance inventory-short form. Int J Chron Obstruct Pulmon Dis 2010;5:415–23.

30 Marsh HW, Ellis LA, Parada RH, Richards G, Heubeck BG. A short version of the self description questionnaire II: Operationalizing criteria for short-form evaluation with new applications of confirmatory factor analyses. Psychol Assess 2005;17:81–102.

31 Smith GT, McCarthy DM, Anderson KG. On the sins of short-form development. Psychol Assess 2000;12:102–11.

32 Derogatis LR, Culpepper WJ. Screening for psychiatric disorders. In: Maruish ME, ed. The Use of Psychological Testing for Treatment Planning and Outcomes Assessment. Mahwah, NJ: Lawrence Erlbaum Associates; 2004:68–69.