



Published in final edited form as:

*Drug Alcohol Depend.* 2016 April 01; 161: 316–322. doi:10.1016/j.drugalcdep.2016.02.014.

## Validation of the alcohol use item banks from the Patient-Reported Outcomes Measurement Information System (PROMIS®)

Paul A. Pilkonis<sup>a,\*</sup>, Lan Yu<sup>a,b</sup>, Nathan E. Dodds<sup>a</sup>, Kelly L. Johnston<sup>a</sup>, Suzanne M. Lawrence<sup>a</sup>, and Dennis C. Daley<sup>a</sup>

<sup>a</sup>Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, PA 15213, United States

<sup>b</sup>Department of Medicine, University of Pittsburgh School of Medicine, Pittsburgh, PA 15213, United States

### Abstract

**Background**—The Patient-Reported Outcomes Measurement Information System (PROMIS®) includes five item banks for alcohol use. There are limited data, however, regarding their validity (e.g., convergent validity, responsiveness to change). To provide such data, we conducted a prospective study with 225 outpatients being treated for substance abuse.

**Methods**—Assessments were completed shortly after intake and at 1-month and 3-month follow-ups. The alcohol item banks were administered as computerized adaptive tests (CATs). Fourteen CATs and one six-item short form were also administered from eight other PROMIS domains to generate a comprehensive health status profile. After modeling treatment outcome for the sample as a whole, correlates of outcome from the PROMIS health status profile were examined.

**Results**—For convergent validity, the largest correlation emerged between the PROMIS alcohol use score and the Alcohol Use Disorders Identification Test ( $r = .79$  at intake). Regarding treatment outcome, there were modest changes across the target problem of alcohol use and other domains of the PROMIS health status profile. However, significant heterogeneity was found in initial severity of drinking and in rates of change for both abstinence and severity of drinking during follow-up. This heterogeneity was associated with demographic (e.g., gender) and health-profile (e.g., emotional support, social participation) variables.

**Conclusions**—The results demonstrated the validity of PROMIS CATs, which require only 4–6 items in each domain. This efficiency makes it feasible to use a comprehensive health status

---

\*Corresponding author at: Western Psychiatric Institute and Clinic, 3811 O'Hara Street, Pittsburgh, PA 15213, United States. Fax: +1 412 246 5840. pilkonispa@upmc.edu (P.A. Pilkonis).

#### Conflict of interest

There are no conflicts of interest for any authors.

#### Contributors

Paul A. Pilkonis, PhD, contributed to study conception, design, and implementation and took responsibility for drafting the manuscript. Lan Yu, PhD, provided data analysis and interpretation. Nathan Dodds, BS, Kelly L. Johnston, MPH, Suzanne Lawrence, MA, and Dennis C. Daley, PhD contributed to study implementation and data collection. All authors reviewed and approved the final manuscript.

profile within the substance use treatment setting, providing important prognostic information regarding abstinence and severity of drinking.

## Keywords

Alcohol use; Measurement; Item response theory; Patient-reported outcomes; PROMIS

---

## 1. Introduction

The Patient-Reported Outcomes Measurement Information System (PROMIS<sup>®</sup>) is an NIH Roadmap initiative devoted to developing better measurement tools for constructs relevant to the investigation and treatment of all diseases—constructs such as pain, fatigue, physical functioning, emotional distress, sleep, and alcohol and substance use (Baysse et al., 2010; Cella et al., 2007b, 2010; Fries et al., 2014; Pilkonis et al., 2011, 2013, 2015). PROMIS has created and refined a comprehensive methodology for developing item banks for these health-related constructs using both qualitative and quantitative techniques and modern psychometric methods, e.g., item response theory (IRT; Cella et al., 2007a, 2010; Hilton, 2011; Reeve et al., 2007). The use of IRT models to calibrate items results in greater precision at both the item and test levels and reduces respondent burden because of the small number of items (usually 4–6) needed to determine a respondent's status when items are administered as computerized adaptive tests (CATs). Additional discussion of the advantages of IRT and CAT is included as Supplementary material.

Following creation of item banks, our priority has been to validate them, most often in prospective observational studies. We report here on a study designed to validate the PROMIS alcohol use item banks with outpatients who had begun treatment for substance use disorder within the past 30 days. There were three specific goals for the protocol. The first was to demonstrate the utility (acceptability and ease of use) of the PROMIS measures, even when administered within an active clinical setting. The second goal was to demonstrate the convergent validity of the PROMIS alcohol item banks with existing alcohol instruments and to examine the responsiveness to change of the alcohol use item bank during treatment. All patients completed assessments at three points: shortly after intake (T1), one month following intake (T2), and three months following intake (T3). In addition to the alcohol use item banks, 14 other CATs and one 6-item short form from 8 other PROMIS domains were administered to generate a comprehensive health status profile for all participants. After modeling treatment outcome for alcohol use for the sample as a whole, the third goal was to identify correlates of outcome from the PROMIS health status profile.

To achieve the first goal (demonstration of utility), we conducted interviews with patients and their clinicians at T1 and T3 about their experience of completing the PROMIS measures. We provided feedback not only on alcohol use scores but also on the other domains of the health status profile, and we asked about the value of receiving such feedback for clinical care. A report of this work has documented the feasibility and perceived value of the PROMIS battery (Johnston et al., 2016).

With regard to the second goal, we examined convergent validity at the intake assessment by correlating scores on the PROMIS item banks with the Alcohol Use Disorders Identification

Test (AUDIT; Saunders et al., 1993); the CAGE Questionnaire (Ewing, 1984); the 6-item set of “recommended alcohol questions” developed by a National Institute on Alcohol Abuse and Alcoholism (NIAAA) task force (2003), which assess frequency and quantity of alcohol use and patterns of consumption; and the Comprehensive Alcohol Expectancy Questionnaire (CAEQ; Nicolai et al., 2010). In addition, we examined the PROMIS equivalents of clinically meaningful ranges and thresholds on the AUDIT, AUDIT-C, and CAGE, e.g., a score of 20 or higher on the AUDIT, a score of 2 or more on the CAGE.

The analysis of longitudinal outcome data regarding alcohol use presents special challenges. Before asking about aspects of drinking, the item banks for alcohol use and negative and positive consequences include a screening question: “Have you had a drink of alcohol in the past 30 days?” It is common for periods of abstinence to alternate with periods of use or abuse, especially in samples of patients for whom abstinence is a goal but in whom relapses may be common. Thus, the item banks provide information about two related but distinct phenomena: abstinence (versus exposure to any alcohol) and severity of use when exposed. Analyzing such data requires novel models, and the approach we chose was the two-part random-effects model for semi-continuous data described by Olsen and Schafer (2001). The model allows simultaneous analysis of binary (yes/no) and continuous (severity) data, yielding parameter estimates (intercepts and slopes) for functions that model both exposure and severity.

After modeling treatment outcome for the sample as a whole, the third goal was to identify correlates of outcome from the PROMIS health status profile. Comorbid conditions such as depression, anxiety, and sleep disturbance are common among individuals with substance use disorders (Stinson et al., 2005). However, comprehensive assessment of physical and mental health may not always be done during their care. The efficiency of the PROMIS measures makes it feasible to generate a health status profile at a small cost of time and effort. We predicted that substance use patients would show elevations on emotional distress, sleep disturbance, and cognitive concerns and that these variables would be related to treatment outcome. Other relevant demographic variables (gender, race, age, and educational attainment) were also included in the prognostic analyses.

## 2. Method

### 2.1. Inclusion criteria

Men and women 18 years and older who were able to read and understand English were enrolled in the protocol. Participants were required to have begun outpatient treatment for substance use within the past 30 days at the Center for Psychiatric and Chemical Dependency Services, a treatment program in Addiction Medicine Services at the University of Pittsburgh Medical Center. Only participants who indicated that they had had a drink of alcohol within the past 30 days were screened into the study because of the 30-day time frame for the alcohol use and consequences item banks. In addition, patients had to have attended at least one treatment session after their initial assessment to be eligible for the second and third assessments. Informed consent was obtained from all participants included in the study.

## 2.2. Exclusion criteria

Clinical exclusions were dementia, other major cognitive impairment, any major medical condition that may have a significant impact on the central nervous system (e.g., Parkinson's disease, stroke, aneurysm), or a history of any psychotic disorder that might compromise the validity of self-reports or interfere with questionnaire completion.

## 2.3. Sample

An initial sample of 225 patients completed the intake assessment (T1), 164 completed the second assessment (T2) one month later, and 158 completed the third assessment (T3) three months later. Attrition in a range from 26–60% has been documented in the first month of treatment in substance use samples (Graff et al., 2009; Loveland and Driscoll, 2014; SAMHSA, 2010; Stark, 1992). Our attrition of 27% at one month and 30% at three months is close to the lower bound of this range. The mean age of the sample was 38 (SD = 12). In terms of gender, ethnicity, and race, 44% of the sample was female, 6% identified as Hispanic, and 44% was non-white (with the large majority of these participants being African-American, 40% of the sample). In terms of educational attainment, 39% had a high school diploma or less. The majority of the sample (65%) had an annual household income of less than \$20,000 per year.

## 2.4. Test administration

Participants received the CAT versions of PROMIS measures, including alcohol use, emotional distress, sleep disturbance, fatigue, pain, physical functioning, cognitive concerns, and social participation and support. The fixed, six-item form for sexual interest was also administered. The PROMIS measures are scored on a T-score metric, with a mean of 50 (normed to the US general population) and a standard deviation of 10. With T-scores, a score of 60 (one SD above the mean) is often used as a clinically significant elevation, and there is evidence to support the validity of this threshold with PROMIS measures (Pilkonis et al., 2014).

There are five PROMIS alcohol use item banks: alcohol use, negative and positive consequences of alcohol use, and negative and positive expectancies regarding drinking (Pilkonis et al., 2013). The labels for the latter four item banks are self-explanatory, but it should be noted that the general “alcohol use” item bank focuses on problematic drinking rather than the frequency, pattern, and volume of consumption *per se*. Such items about consumption are included in the bank, but the most informative items that emerged from our IRT-calibration analyses were items reflecting struggles to control drinking and cravings for alcohol. Additional discussion of the development of the item banks and a more extended description of their rationale and content are included as Supplementary material. Tables providing the individual items in each bank are also in the Supplement.

The CAT algorithm required a minimum of 4 items and a maximum of 12, and it stopped the test when the standard error was less than .30. Across all tests at T1, T2, and T3, participants received an average of 5.8 items (SD = 1.8) for the alcohol use CAT, with a median of 5 items. Participants also completed “legacy” measures that assessed consumption, consequences, and expectancies regarding alcohol: The Alcohol Use Disorders Identification

Test (AUDIT; Saunders et al., 1993); the CAGE Questionnaire (Ewing, 1984), with a time frame of the past year; the 6-item set of NIAAA “recommended alcohol questions” (2003); and the Comprehensive Alcohol Expectancies Questionnaire (CAEQ; Nicolai et al., 2010). Participants also answered questions about their demographic characteristics and use of all substances both in their lifetimes and in the past 30 days. Items were displayed one at a time using Assessment Center, the PROMIS electronic testing platform (see <http://www.assessmentcenter.net>).

## 2.5. Statistical methods

**2.5.1. Psychometric analyses**—To investigate convergent validity, we computed bivariate Pearson correlations between the PROMIS alcohol CATs and the legacy measures (AUDIT, CAGE, NIAAA questions, and CAEQ), examining each correlation separately without controlling for covariates. We also examined the PROMIS equivalents of conventional thresholds and ranges of severity on the relevant legacy measures. For example, on the AUDIT, scores of 0–7 are below the threshold of concern, scores of 8–15 are indications for providing advice on the reduction of drinking, scores of 16–19 prompt brief counseling and continued monitoring, and scores of 20 or greater trigger further diagnostic evaluation.

**2.5.2. Outcome analyses**—We used the Olsen and Schafer (2001) two-part random effects model for semi-continuous longitudinal data to analyze scores from the PROMIS alcohol use CAT. The two-part model is a linear growth model for both the binary outcome (yes/no: any drink in the past 30 days) and the continuous outcome (when drinking, the severity score from the PROMIS alcohol use CAT). We examined the relationships of the intercepts and slopes of both the binary and continuous alcohol use variables to other relevant demographic and clinical variables to investigate which of these measures were associated with initial status (intercept) or rate of change (slope) on alcohol use. For this purpose, the scores on PROMIS domains other than alcohol use were of primary interest. The aim was to examine the potential value of the PROMIS health status profile as a prognostic tool in the treatment of substance use disorder. Additional discussion of the two-part random effects model is included as Supplementary material.

## 3. Results

### 3.1. Pattern of missing data

The pattern of missing data is summarized in Table 1, with 64% of the sample providing complete data. All participants provided data at T1. At T2 and T3, 27% and 30% of participants, respectively, failed to provide data, most often because they were no longer in treatment. Among respondents who did provide data at the two follow-up assessments, 35% were abstinent from alcohol at T2 and 37% at T3. Also, note that at T1 (which occurred after some time had elapsed from the initial screening), 26 of the 225 participants (12%) had passed the 30-day threshold for abstinence and opted out of the CAT for PROMIS alcohol use by denying any drinking in the past 30 days, i.e., they provided data only for the binary yes/no drinking question.

## 3.2. Descriptive statistics

**3.2.1. Alcohol, drug, and tobacco use**—Table 2 summarizes the results (means and standard deviations) from the PROMIS alcohol use, consequences, and expectancies CATs; the AUDIT; the AUDIT-C (the first three items from the AUDIT); the CAGE (used with a time frame of the past 12 months), and the six NIAAA questions. At baseline, there were significant elevations on all the alcohol use measures. The mean PROMIS alcohol use T-score was 60.8 (SD = 8.0), and in similar fashion, the means on the legacy measures were all above the usual thresholds for clinical concern: AUDIT = 20.7 (11.0), AUDIT-C = 7.4 (3.6), and CAGE = 2.8 (1.3).

In addition to alcohol, 97% of the sample had used other drugs during their lifetime, with a mean of 4.2 other substances (SD = 2.8). See Table 3 for a summary of lifetime substance use and use in the past 30 days. The most frequently used lifetime substances were cannabis (92%), cocaine (67%), and pain medications (46%). Twenty percent of the sample reported a history of injecting a drug. In the past 30 days, 76% of the sample reported using other drugs in addition to alcohol ( $M = 1.6$  substances,  $SD = 1.5$ ). On a lifetime basis, 82% of the sample reported smoking tobacco, with 65% as current smokers.

**3.2.2. PROMIS health status profile**—Table 2 also summarizes the means and standard deviations of the T-scores at baseline from the additional CATs included in the PROMIS health status profile. On the PROMIS measures, the most pronounced symptoms appeared in the areas of emotional distress and sleep disturbance (with means in a range from 58.0 to 63.4). Respondents also complained about fatigue, pain, and difficulties in social support and participation but at lower levels of severity. The only domains where symptoms were not as apparent were physical functioning, cognitive concerns, and sexual interest.

## 3.3. Convergent validity

The correlations at baseline between the PROMIS alcohol item banks and the alcohol legacy measures provided good evidence of the convergent validity of the different instruments. The largest correlation occurred between the PROMIS alcohol use score and the full AUDIT ( $r = .79$ ), with a slightly smaller correlation between PROMIS alcohol use and the AUDIT-C ( $r = .70$ ), the first three items from the AUDIT often used as a screening measure. The correlation between PROMIS alcohol use and the CAGE was .56. The correlations between PROMIS alcohol use and the six individual NIAAA questions ranged from .44 to .57, and the correlation between PROMIS alcohol use and the total score summed from the six questions was .72.

These correlations may be conservative estimates because time frames vary across the instruments, leading to potential discrepancies between them. The PROMIS item banks use a 30-day time frame. The first three items of the AUDIT (the AUDIT-C) have no time frame specified, whereas the last seven items refer to either the past year or lifetime. When used for screening, the CAGE has an unspecified time frame, although we used a time frame of the past 12 months in the current study. The first five NIAAA questions refer to the past 12 months, with the sixth question asking about lifetime occurrence (largest number of drinks in a 24-h period).

The Comprehensive Alcohol Expectancy Questionnaire (CAEQ) has five subscales: social assertiveness and positive affect, tension reduction, cognitive impairment and physical discomfort, aggression, and sexual enhancement. Large correlations emerged between the PROMIS positive consequences and positive expectancies of alcohol use CATs and two of the positive CAEQ subscales: social assertiveness and positive affect ( $r$ 's = .71 and .62 for positive consequences and positive expectancies, respectively) and tension reduction ( $r$ 's = .53 and .59 for positive consequences and positive expectancies, respectively). In a similar way, substantial correlations emerged between the PROMIS negative consequences of alcohol use CAT and two of the negative CAEQ subscales: cognitive impairment and physical discomfort ( $r$  = .51) and aggression ( $r$  = .51). A comprehensive table, containing intercorrelations among all measures used in the protocol at intake, is included as Supplementary material.

Table 4 summarizes the PROMIS equivalents (mean scores) of the ranges on the AUDIT, AUDIT-C, and CAGE used for classifying different levels of risk. Consistent with usual clinical practice and past PROMIS results (Pilkonis et al., 2014), a PROMIS T-score of about 60 (one SD above the mean) is a marker of risk equivalent to the conventional thresholds for clinical concern from these legacy measures.

### 3.4. Outcome analyses

The means and variances of the intercepts and slopes of the binary (yes/no drinking) and continuous (PROMIS alcohol use score) variables were unconstrained initially for estimation. The mean of the intercept for the continuous variable and the means of the slopes for both the binary and continuous variables were all significant, indicating an intercept different from zero for the continuous variable and slopes different from zero for both the binary and continuous variables. The variances for the intercept of the continuous alcohol use variable and for the slope of the binary yes/no drinking variable were also significant, indicating heterogeneity in the baseline scores for PROMIS alcohol use and in the individual slopes for the yes/no drinking variable (abstinence). Conversely, there was no significant variation in the intercept for the binary variable (this result occurred by design—all participants were screened for presence of drinking in the recent past) or in the individual slopes for the continuous alcohol use variable.

On the T-score metric, the means and SDs of the observed scores for PROMIS alcohol use were 60.8 (8.0) for baseline, 58.3 (8.7) for 1-month follow-up, and 57.7 (8.1) for 3-month follow-up. The means and SDs of expected alcohol use scores estimated from the two-part model were 60.2 (5.5) at baseline, 58.3 (5.6) for one-month follow-up, and 54.4 (6.9) for 3-month follow-up. For the binary yes/no drinking variable, the observed percentages of participants having had a drink in the past 30 days were 88% at baseline, 65% at 1-month follow-up, and 63% at 3-month follow-up. The expected percentages estimated from the two-part model were 86% at baseline, 81% at 1-month follow-up, and 69% at 3-month follow-up. In general, agreement between observed and expected scores was good, but given the pattern of missing data, the two-part model estimated higher rates of any drinking at follow-up than we were able to observe among compliant respondents.

**3.4.1. Correlates of the initial severity of drinking**—Because participants were selected on the basis of their history of recent drinking, the intercept (baseline value) of the yes/no drinking variable was uninformative. Therefore, we constrained it to be zero in the versions of the two-part model that examined relationships between demographic and clinical variables and the other parameters of the model (the intercept of the continuous severity of use variable, the slope of the binary yes/no drinking variable, and the slope of the continuous severity of use variable). Demographic variables included age, gender, race, and education. Clinical variables included emotional distress (depression, anxiety, anger), pain (pain interference, pain behavior), fatigue, physical functioning, cognitive concerns, sleep (sleep disturbance, sleep-related impairment), social relatedness (emotional support, social isolation), social participation (satisfaction with social roles, satisfaction with leisure activities), sexual interest, and the presence versus absence of an interview at baseline providing feedback on these domains of the PROMIS health status profile. These variables were examined one at a time in repeated iterations of the two-part model.

Among the demographic variables examined, age (.015,  $p < .001$ ) and gender ( $-.010$ ,  $p < .001$ ) were associated significantly with initial severity of drinking (the intercept of the continuous score for alcohol use). Among the clinical variables investigated, emotional support ( $-.183$ ,  $p < .01$ ), cognitive concerns (.160,  $p < .01$ ), and sleep disturbance (.137,  $p < .01$ ) were associated significantly with initial severity of drinking. The coefficients reported here are unstandardized beta coefficients from the two-part model. The positive coefficients indicated that increasing age, cognitive concerns, and sleep disturbance were associated with more severe alcohol use at baseline. The negative coefficients indicated that participants with more emotional support and women (coded 1, with a mean T-score for PROMIS alcohol use of 59.1, versus men coded 0, with a mean T-score of 62.1) had less severe alcohol use at baseline.

**3.4.2. Correlates of change in the severity of drinking**—Given that there was significant variability in severity of use at baseline (intercept), we controlled for initial severity in all models examining correlates of the slope of the continuous PROMIS alcohol use score. Higher initial scores may allow for steeper declines, and we wanted to control for this potential effect as we examined all other correlates. The total AUDIT score at baseline was used as the measure of initial severity.

Among the demographic variables examined, gender (.003,  $p < .01$ ; control variable AUDIT, .005, *ns*) was associated significantly with change in severity of drinking (the slope of the continuous score for alcohol use). Among the clinical variables examined, the presence of an interview providing feedback on the PROMIS health status profile at baseline ( $-.324$ ,  $p < .001$ ; control variable AUDIT, .009,  $p < .05$ ) was also associated significantly with change in severity of drinking. The positive coefficient for gender indicated that women (coded 1 versus men coded 0) had positive (or increasing) slopes, i.e., worse outcomes. Initial severity was not significant in this model. The negative coefficient for the presence of an interview indicated that receiving an interview incorporating feedback about the PROMIS health status profile was associated with negative (or decreasing) slopes, i.e., better outcomes. Initial severity was also significant in this model, but the positive coefficient for



the AUDIT indicated that higher initial severity was associated with positive (or increasing) slopes, i.e., worse outcomes.

### 3.4.3. Correlates of change in the presence versus absence of any drinking—

Among the demographic variables examined, gender (.021,  $p < .05$ ) was associated significantly with change in abstinence (the slope of the binary yes/no drinking variable). Among the clinical variables examined, anger (.399,  $p < .01$ ), satisfaction with performance of social roles (-.416,  $p < .05$ ), emotional support (-.414,  $p < .05$ ), and satisfaction with discretionary social activities (-.407,  $p < .05$ ) were associated significantly with change in abstinence. The positive coefficients for gender and anger indicated that women (coded 1 versus men coded 0) and participants with higher levels of anger at baseline had positive (or increasing) slopes, i.e., higher probabilities of any drinking during follow-up (worse outcomes). The negative coefficients for emotional support and for satisfaction with social participation (both in conventional social roles and in discretionary social activities) indicated that higher levels of these domains at baseline were associated with negative (or decreasing) slopes, i.e., lower probabilities of any drinking during follow-up (better outcomes).

## 4. Discussion

In the current sample, there was good evidence of convergent validity between the PROMIS alcohol use item banks and other traditional measures of alcohol use, especially the AUDIT. Over the three-month observational period, the PROMIS alcohol use CAT was also responsive to changes in drinking despite the relatively modest improvements in mean levels of alcohol use. The mean PROMIS alcohol use score changed from 60.8 at baseline to 57.7 at 3-month follow-up, reflecting an effect size of .39. Nonetheless, the slopes for both the binary variable of presence versus absence of drinking and the continuous variable of severity of alcohol use derived from the PROMIS CAT were significantly different from zero in the two-part random effects model for semi-continuous longitudinal data that we used to analyze treatment outcome. This result of modest change was also true of the other domains of the PROMIS health status profile. For example, the mean score for PROMIS anxiety (the highest score on the profile at baseline) changed only from 63.4 to 60.3 over the same 3-month period (effect size again = .39). We take this as evidence of the difficulty in treating substance use disorder (and related problems) in the short term, given that it is a chronic, relapsing condition often requiring several years to sustain long-term recovery (Dennis and Scott, 2007).

In this general context, however, we were able to identify some correlates of (a) initial severity of drinking and (b) rates of change for both severity of drinking and any exposure to alcohol (abstinence). The correlates of initial severity of drinking were age, gender, emotional support, cognitive concerns, and sleep disturbance. The correlates of rate of change in severity were fewer: gender, and participation in an interview providing feedback on the PROMIS health status profile at intake. The correlates of rate of change in abstinence reflected both similarities and differences with those related to severity of drinking. Overlap occurred with gender and emotional support, but relationships also emerged with satisfaction with participation in social roles, satisfaction with participation in discretionary social

activities, and anger. Summarizing across these results, major themes included worse outcomes for women (on both our continuous and binary outcome variables) and the relationship of the social context to abstinence. In the latter case, emotional support, satisfaction with social roles, and satisfaction with discretionary social activities were all associated with the absence of any drinking. This result is consistent with prior evidence that social support predicted long-term abstinence rates across various addictions (Marlatt and Witkiewitz, 2005). Active involvement in mutual support programs is also associated with long-term abstinence, improved functioning, and greater self-efficacy (Donovan et al., 2013; Laudet et al., 2006).

Examining the impact of a feedback interview (about many different aspects of health, not just alcohol use) was a goal of the project, consistent with increasing use of patient-reported outcomes in clinical research and increasing emphasis on patient-centered approaches to clinical care and research. Therefore, it is noteworthy that this innovation was associated with greater declines in severity of alcohol use over the course of treatment. Generating a comprehensive health status profile and conveying this information to both patients and their clinicians can now be done with a minimum of participant burden using PROMIS CATs. Collecting and sharing such information is a clinically meaningful intervention, and its positive impact on severity of drinking in the current sample provides support for continuing to investigate its use in clinical practice.

The two-part random effects model for semi-continuous longitudinal data allowed us to examine separately the influences on any exposure to alcohol (abstinence) and severity of use when drinking, and the capacity to distinguish between these different but related constructs contributes to its utility. Our findings support the position that the two constructs can be related to different correlates and that, as a result, they deserve to be examined separately. It should be noted that other discussions of analyzing data in which information about severity is contingent on the presence of some behavior or event are available (e.g., Liu and Verkuilen, 2013; Reardon and Raudenbush, 2006).

Given that validation is an evolving process (rather than a single outcome), it is important to note some of the limitations of the current work. A key limitation was the constraints imposed by the legacy measures used. In general, the AUDIT and CAGE are used most commonly for screening rather than assessment of outcome. The NIAAA questions are difficult to interpret as a “scale” and are intended more to establish patterns of drinking, often in an epidemiological context. As expected, the AUDIT had the strongest relationship with the PROMIS alcohol use CAT because of the greatest similarity in content, i.e., the AUDIT asks about usual consumption of alcohol but has an even larger number of questions regarding problematic aspects of drinking and negative consequences, consistent with the most informative items from the PROMIS alcohol use and negative consequences banks. The relationship with the AUDIT-C was less strong (although still large), given that the AUDIT-C focuses on consumption alone. The CAEQ served in predicted ways to establish convergent validity with the PROMIS negative and positive expectancy item banks, but we had no measure of positive consequences in the current assessment battery. Future research should consider the use of additional scales to establish convergent and discriminant validity.

In summary, the results supported the utility and validity of the PROMIS alcohol use item banks and, within limits, their comparability to other legacy measures of alcohol use. The findings also documented the modest but statistically significant impact of treatment for substance use disorder over a three-month period and the value of examining the presence of any drinking versus severity of drinking as distinct but related outcomes. The results identified both demographic correlates (gender, age) and clinical correlates (social context, cognitive concerns, sleep, anger) of these variables in terms of both initial status and outcome over a three-month period. Such findings support the value of including broader indicators of health and functioning in the assessment and treatment of patients with substance use disorder. In addition, providing such information to both patients and clinicians had a positive impact on severity of drinking, suggesting that it can and should be integrated into clinical care. The brevity and efficiency of the PROMIS item banks (especially when implemented as CATs) now make such efforts feasible.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

### Role of funding source

This project was funded by the Patient-Centered Outcomes Research Institute (1P2PI000189; PI: Paul A. Pilkonis, PhD).

We acknowledge the work of our colleagues within the Center for Psychiatric and Chemical Dependency Services at the University of Pittsburgh Medical Center who collaborated in this work: Trey Ghee, LSW, Dorothy Sandstrom, MS, and Janis McDonald.

## References

- Buysse DJ, Yu L, Moul DE, Germain A, Stover A, Dodds NE, Johnston KL, Shablesky-Cade MA, Pilkonis PA. Development and validation of patient-reported outcome measure for sleep disturbance and sleep-related impairments. *Sleep*. 2010; 33:781–792. [PubMed: 20550019]
- Cella D, Gershon R, Lai JS, Choi S. The future of outcomes measurement: item banking tailored short forms, and computerized adaptive assessment. *Qual Life Res*. 2007a; 16:133–144. [PubMed: 17401637]
- Cella D, Yount S, Rothrock N, Gershon R, Cook K, Reeve B, Ader D, Fries JF, Bruce B, Rose M. The Patient-Reported Outcomes Measurement Information System (PROMIS<sup>®</sup>): progress of an NIH Roadmap cooperative group during its first two years. *Med Care*. 2007b; 45:S3–S11.
- Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, Amtmann D, Bode R, Buysse D, Choi S, Cook K, Devellis R, DeWalt D, Fries JF, Gershon R, Hahn EA, Lai JS, Pilkonis P, Revicki D, Rose M, Weinfurt K, Hays R. The Patient-Reported Outcomes Measurement Information System (PROMIS<sup>®</sup>) developed and tested its first wave of adult self-reported outcome item banks: 2005–2008. *J Clin Epidemiol*. 2010; 63:1179–1194. [PubMed: 20685078]
- Dennis ML, Scott CK. Managing addiction as a chronic condition. *Addict Sci Clin Pract*. 2007; 4:45–55. [PubMed: 18292710]
- Donovan DM, Ingalsbe MH, Benbow J, Daley DC. 12-step interventions and mutual support programs for substance use disorders: an overview. *Soc Work Public Health*. 2013; 28:313–332. [PubMed: 23731422]
- Ewing JA. Detecting alcoholism: the CAGE questionnaire. *JAMA*. 1984; 252:1905–1907. [PubMed: 6471323]

- Fries JF, Witter J, Rose M, Cella D, Khanna D, Morgan-Dewitt E. Item response theory, computerized adaptive testing, and PROMIS: assessment of physical function. *J Rheumatol*. 2014; 41:153–158. [PubMed: 24241485]
- Graff FS, Morgan TJ, Epstein EE, McCrady BS, Cook SM, Jensen NK, Kelly S. Engagement and retention in outpatient alcoholism treatment for women. *Am J Addict*. 2009; 18:277–288. [PubMed: 19444731]
- Hilton TF. The promise of PROMIS® for addiction. *Drug Alcohol Depend*. 2011; 119:229–234. [PubMed: 22238781]
- Johnston KL, Lawrence SM, Dodds NE, Yu L, Daley DC, Pilkonis PA. Evaluating PROMIS® instruments and methods for patient-centered outcomes research: patient and provider voices in a substance use treatment setting. *Qual Life Res*. 2016; 25:615–624. [PubMed: 26353907]
- Laudet AB, Morgen K, White WL. The role of social supports spirituality, religiousness, life meaning and affiliation with 12-step fellowships in quality of life satisfaction among individuals in recovery from alcohol and drug problems. *Alcohol Treat Q*. 2006; 24:33–73. [PubMed: 16892161]
- Liu Y, Verkuilen J. Item response modeling of presence-severity items: application to measurement of patient-reported outcomes. *Appl Psychol Meas*. 2013; 37:58–75.
- Loveland D, Driscoll H. Examining attrition rates at one specialty addiction treatment provider in the United States: a case study using a retrospective chart review. *Subst Abuse Treat Prev*. 2014; 9:41.
- Marlatt, GA., Witkiewitz, K. Relapse prevention for alcohol and drug problems. In: Marlatt, GA., Donovan, DD., editors. *Relapse Prevention*. Guilford; New York: 2005. p. 1-44.
- National Institute on Alcohol Abuse and Alcoholism (NIAAA). [accessed on 4.11.11] Task Force On Recommended Alcohol Questions: National Council On Alcohol Abuse And Alcoholism Recommended Set Of Alcohol Consumption Questions. 2003. Oct 15–16. <<http://www.niaaa.nih.gov/Resources/ResearchResources/TaskForce.htm>>
- Nicolai J, Demmel R, Moshagen M. The comprehensive alcohol expectancy questionnaire: confirmatory factor analysis, scale refinement, and further validation. *J Pers Assess*. 2010; 92:400–409. [PubMed: 20706926]
- Olsen MK, Schafer JL. A two-part random-effects model for semicontinuous longitudinal data. *J Am Stat Assoc*. 2001; 96:730–745.
- Pilkonis PA, Choi SW, Reise SP, Stover AM, Riley WT, Cella D. Item banks for measuring emotional distress from the Patient-Reported Outcomes Measurement Information System (PROMIS®): depression, anxiety, and anger. *Assessment*. 2011; 18:263–283. [PubMed: 21697139]
- Pilkonis PA, Yu L, Colditz J, Dodds NE, Johnston KL, Maihoefer C, Stover AM, Daley DC, McCarty D. Item banks for alcohol use from the Patient-Reported Outcomes Measurement Information System (PROMIS®): use consequences, and expectancies. *Drug Alcohol Depend*. 2013; 130:167–177. [PubMed: 23206377]
- Pilkonis PA, Yu L, Dodds NE, Johnston KL, Maihoefer C, Lawrence SM. Validation of the depression item bank from the Patient-Reported Outcomes Measurement Information System (PROMIS®) in a three-month observational study. *J Psychiatr Res*. 2014; 56:112–119. [PubMed: 24931848]
- Pilkonis P, Yu L, Dodds NE, Johnston KL, Lawrence SM, Hilton TF, Daley DC, Patkar AA, McCarty D. Item banks for substance use from the Patient-Reported Outcomes Measurement Information System (PROMIS®): severity of use and positive appeal of use. *Drug Alcohol Depend*. 2015; 156:184–192. [PubMed: 26423364]
- Reardon S, Raudenbush SW. A partial independence item response model for surveys in which responses to filter questions determine whether subsequent questions are asked. *Sociol Methodol*. 2006; 36:257–300.
- Reeve BB, Hays RD, Bjorner JB, Cook KF, Crane PK, Teresi JA, Thissen D, Revicki DA, Weiss DJ, Hambleton RK, Liu H, Gershon R, Reise SP, Cella D. PROMIS Cooperative Group. Psychometric evaluation and calibration of Health-Related Quality of Life item banks: plans for the Patient-Reported Outcome Measurement Information System (PROMIS®). *Med Care*. 2007; 45:S22–S31. [PubMed: 17443115]
- SAMHSA. Treatment Episode Data Set: Discharges from Substance Abuse Treatment Services. Substance Abuse and Mental Health Services Administration; Rockville, MD: 2010. BHSIS Series S-67, HHS Publication No. (SMA) 14-48172013

- Saunders JB, Aasland OG, Babor TF, Delafuente JR, Grant M. Development of the Alcohol-Use Disorders Identification Test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption. *Addiction*. 1993; 88:791–804. [PubMed: 8329970]
- Stark MJ. Dropping out of substance abuse treatment: a clinically oriented review. *Clin Psychol Rev*. 1992; 12:93–116.
- Stinson FS, Grant BF, Dawson DA, Ruan WJ, Huang B, Saha T. Comorbidity between DSM-IV alcohol and specific drug use disorders in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Drug Alcohol Depend*. 2005; 80:105–116. [PubMed: 16157233]

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.drugalcdep.2016.02.014>.

**Table 1**

Missing data patterns for the PROMIS Alcohol Use item bank.

Study visit			N	%
T1	T2 <sup>a</sup>	T3 <sup>b</sup>		
1	1	1	145	64
1	1	0	19	8
1	0	1	13	6
1	0	0	48	21

<sup>a</sup>Missing data at T2:  $N=61$  (27%).<sup>b</sup>Missing data at T3:  $N=67$  (30%).

Table 2

Means and standard deviations for all study measures.

Measure	Baseline (T1)			1-Month (T2)			3-Month (T3)		
	M	SD	n	M	SD	n	M	SD	n
PROMIS alcohol use measures									
Alcohol use	60.8	8.0	199	58.3	8.7	107	57.7	8.1	98
Negative consequences	55.0	6.6	194	53.9	6.9	104	51.6	8.5	93
Negative expectancies	53.7	8.2	225	52.3	7.9	165	53.3	10.1	158
Positive consequences	56.2	7.9	196	53.8	8.1	100	54.0	7.9	94
Positive expectancies	57.7	8.3	225	54.5	8.6	163	52.7	9.1	158
AUDIT	20.7	11.0	224	15.7	10.2	165	13.5	9.9	159
AUDIT-C	7.4	3.6	224	4.7	3.5	164	4.2	3.5	159
CAGE	2.8	1.3	224	2.9	1.3	164	2.9	1.4	157
NIAAA alcohol questions (6-item version) <sup>d</sup>									
Q1: frequency of drinking	6.5	2.1	223	6.3	2.3	162	6.1	2.3	154
Q2: average number of drinks/day	5.0	2.1	224	4.9	2.1	161	4.7	2.1	157
Q3: maximum number of drinks/24 h period	6.9	1.9	223	6.7	2.1	161	6.4	2.1	156
Q4: frequency of maximum number of drinks	4.1	2.4	223	3.9	2.4	163	4.2	2.4	156
Q5: frequency of binge drinking	4.9	2.9	223	4.7	2.9	163	4.5	2.9	156
Q6: maximum number of drinks/24 h period—lifetime	7.1	2.0	222	6.9	2.2	164	6.8	2.1	157
CAEQ—factors 1–5									
F1: social assertiveness & positive affect	55.4	12.3	224	51.6	12.0	162	48.9	12.5	156
F2: tension reduction	25.9	6.2	223	24.4	5.5	165	22.8	5.9	157
F3: cognitive impairment & physical discomfort	30.7	8.0	224	30.8	7.0	164	31.1	7.9	157
F4: aggression	11.0	4.4	225	10.9	3.9	163	11.0	4.2	157
F5: sexual enhancement	16.4	5.1	225	15.8	4.7	165	15.8	5.0	157
Other PROMIS Measures									
Higher scores indicate greater impairment									
Anxiety	63.4	8.0	223	61.5	9.5	165	60.3	10.1	158
Depression	62.5	8.5	223	60.7	9.6	165	59.6	10.0	158
Sleep-related impairment	59.0	8.8	223	57.5	8.4	164	56.6	10.9	157

Measure	Baseline (T1)			1-Month (T2)			3-Month (T3)		
	M	SD	n	M	SD	n	M	SD	n
Anger	58.6	9.9	223	56.7	10.2	165	56.9	10.7	157
Sleep disturbance	58.0	9.2	222	55.7	8.7	165	54.5	9.2	158
Social isolation	57.4	8.4	224	56.6	8.2	165	55.1	8.9	158
Fatigue	56.6	8.9	223	55.3	7.9	164	54.7	10.4	158
Pain interference	55.0	10.7	222	54.3	10.1	165	54.3	10.5	157
Pain behavior	54.5	10.2	223	53.6	10.7	165	53.2	10.4	158
Cognitive concerns	39.6	8.5	223	38.8	8.0	165	37.6	9.2	158
Lower scores indicate greater impairment									
Interest in sexual activity	52.9	10.3	225	52.0	10.4	164	52.3	10.1	157
Physical function	49.2	9.1	223	48.3	8.6	165	49.0	9.3	158
Emotional support	46.5	8.6	223	47.4	8.3	165	47.8	8.4	158
Satisfaction w/discretionary social activities	44.5	7.9	224	45.8	7.4	163	46.5	8.5	158
Satisfaction w/social roles	44.4	8.6	223	44.9	8.1	165	45.8	7.9	158

Note: the PROMIS item bank scores are T-scores with a population mean of 50 (SD = 10).

<sup>4</sup>The specific wording of the NIAAA questions and response options is included as Supplementary material.



**Table 3**

Self-reported substance use.

<b>Substance</b>	<b>Ever used (%)</b>	<b>Used in past 30 days (%)</b>
Cannabis	91.9	55.1
Cocaine	67.1	30.7
Pain medication	45.8	23.6
Hallucinogens	38.2	4.0
Sedatives	37.3	17.8
Amphetamines	34.7	6.2
Heroin	25.8	13.3
Inhalants	24.4	2.7
Other	24.4	4.9

**Table 4**

PROMIS Alcohol Use equivalents for clinical tiers of the AUDIT, AUDIT-C, and CAGE.

	T1						T2						T3					
	AUDIT		PROMIS Alcohol Use		AUDIT		PROMIS Alcohol Use		AUDIT		PROMIS Alcohol Use		AUDIT		PROMIS Alcohol Use			
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD		
AUDIT score																		
0-7	3.9	2.0	49.0	6.2	3.3	2.2	46.7	7.7	3.2	2.6	49.5	5.7						
8-15	11.4	2.4	55.6	4.7	10.6	2.4	56.0	5.9	10.5	2.5	53.0	6.3						
16-19	17.2	1.3	58.1	4.0	17.3	1.1	59.9	3.7	17.4	1.0	59.4	5.1						
20+	29.1	5.1	65.1	5.7	26.5	4.8	63.9	5.4	25.4	4.8	63.0	6.1						
	AUDIT-C		PROMIS Alcohol Use		AUDIT-C		PROMIS Alcohol Use		AUDIT-C		PROMIS Alcohol Use							
AUDIT-C score	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD						
Men																		
0-3	1.3	1.5	52.7	8.1	0.9	1.1	51.5	10.2	1.2	1.3	51.2	7.7						
4+	9.1	2.6	63.0	7.3	7.0	2.2	60.7	7.8	7.3	2.2	59.7	6.4						
Women																		
0-2	0.9	0.9	41.6	5.1	0.4	0.8	51.3	10.3	0.6	0.9	48.9	9.2						
3+	7.4	2.7	60.2	6.6	6.4	2.7	59.4	6.6	6.3	2.3	60.8	7.1						
	CAGE		PROMIS Alcohol Use		CAGE		PROMIS Alcohol Use		CAGE		PROMIS Alcohol Use							
CAGE score	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD						
0-1	0.4	0.5	51.1	7.0	0.4	0.5	49.2	7.3	0.2	0.4	49.8	5.6						
2+	3.4	0.7	62.6	6.8	3.4	0.7	60.5	7.6	3.3	0.8	58.8	7.9						

Note: The PROMIS item bank scores are T-scores with a population mean of 50 (SD = 10).