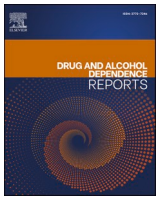
Drug and Alcohol Dependence Reports 8 (2023) 100183

Contents lists available at ScienceDirect

Drug and Alcohol Dependence Reports

journal homepage: www.elsevier.com/locate/dadr



Towards shortening the Brief Addiction Monitor-Revised (BAM-R)

Brittany E. Blanchard [a,*], Kevin G. Lynch [b], Carol A. Malte [c], Eric J. Hawkins [a,c], Dominick DePhilippis [d,e], David W. Oslin [d,f], James R. McKay [d,g], Andrew J. Saxon [a,c]

[a] School of Medicine, University of Washington, Seattle, WA, United States
[b] School of Medicine, University of Pennsylvania Perelman, Philadelphia, PA, United States
[c] Center of Excellence in Substance Addiction Treatment and Education (CESATE), Veterans Affairs Puget Sound Health Care System, Seattle, WA, United States
[d] University of Pennsylvania, Philadelphia, PA, United States
[e] VA Office of Mental Health and Suicide Prevention, Washington, D.C., United States
[f] VISN 4 Mental Illness Research, Education, and Clinical Center, Center of Excellence, Corporal Michael J. Crescenz Philadelphia VA Medical Center, Philadelphia, PA, United States
[g] Center of Excellence in Substance Addiction Treatment and Education (CEASATE), Corporal Michael J. Crescenz Philadelphia VA Medical Center, Philadelphia, PA, United States

HIGHLIGHTS

- The BAM-R is widely recommended for measurement-based care of substance use.
- We aimed to shorten the BAM-R to increase administration feasibility.
- No BAM-R items predicted 90-day SUD treatment retention or 12-month mortality.
- We identified 5 BAM-R items with sensitivity to change and clinical utility.

ARTICLE INFO

Keywords:
Brief Addiction Monitor-Revised (BAM-R)
Measurement-based care
Substance use disorders
Veterans

ABSTRACT

Introduction: The Brief Addiction Monitor-Revised (BAM-R) is a widely used, 17-item assessment of substance use, risk, and protective factors associated with recovery from substance use disorders. Despite wide adoption in the U.S. Department of Veterans Affairs (VA) and recommendations for use in measurement-based care (MBC), administration may not be feasible in many MBC settings due to time constraints. The purpose of this study was to derive a shortened version of the BAM-R for use in fast-paced healthcare settings.

Methods: BAM-R data from 32,002 Veterans were obtained through the VA's Corporate Data Warehouse. We used logistic regression models to identify items for removal based on prediction of two clinical outcomes (90-day substance use disorder (SUD) treatment retention and 12-month mortality) and item-level sensitivity to change during substance use treatment.

Results: Although no intake BAM-R items predicted SUD treatment retention or mortality, effect sizes for item-level sensitivity to change during substance use treatment varied from small to large. Seven items were judged as relevant for MBC of SUD. Among all BAM-R items, Heavy Alcohol Use, Self-Help, Drug Use, Craving, and Mood items demonstrated the greatest magnitude of sensitivity to change.

Conclusions: Although additional research is recommended before a shortened BAM-R can be implemented in non-specialty MBC settings, we identified 5 BAM-R items with perceived clinical utility and scores that demonstrated evidence of sensitivity to change. Shortening the BAM-R increases feasibility of use, though more work is needed to optimize measurement for SUD MBC.

1. Introduction

The Brief Addiction Monitor-Revised (BAM-R) is a well-known 17-item measure of substance use frequency, related risk, and recovery developed to support delivery of measurement-based care (MBC; Cacciola et al., 2013; McKay et al., 2009). MBC is operationalized as

* Corresponding author.
E-mail address: bblancha@uw.edu (B.E. Blanchard).

https://doi.org/10.1016/j.dadr.2023.100183
Received 30 March 2023; Received in revised form 3 August 2023; Accepted 4 August 2023
Available online 6 August 2023

"enhanced precision and consistency in disease assessment, tracking, and treatment to achieve optimal outcomes" (Harding et al., 2011, p. 1137), which includes repeated assessments of patient outcomes to guide shared decision-making between patients and clinicians (see Fortney et al., 2017; Goodman et al., 2013; Trivedi and Daly, 2007). The Addiction Medicine Practice-Based Research Network listed the BAM as a recommended assessment tool for implementation of MBC for opioid use disorders (OUD) and SUD (Clarke et al., 2021). The Kennedy Forum lists the BAM in its recommended core set of SUD outcome measures (Wrenn et al., 2017).

In 2016, the VA launched the national Measurement-Based Care in Mental Health Initiative to increase MBC implementation (Peterson et al., 2019). The VA has a full range of SUD specialty outpatient programming, including intensive outpatient (IOP), SUD outpatient clinics, and opioid treatment programs. As of December2022, the VA's healthcare system in the United States (U.S.) requires use of the BAM for MBC of Veterans being treated in VA SUD treatment programs (U.S. Department of Veterans Affairs, 2022).

1.1. Psychometric evaluations of the BAM-R

Using exploratory factory analysis (EFA), initial psychometric testing during the BAM-R's development indicated a 3-factor structure with 11 items loading onto three factors: Recovery Protection, Physical & Psychological Problems, and Substance Use & Risk (Cacciola et al., 2013). These subscale scores demonstrated evidence of sensitivity to change over three months. Baseline Recovery Protection and Substance Use & Risk subscale scores demonstrated evidence of predictive validity (i.e., completion of an intensive substance use outpatient program; Cacciola et al., 2013).

Although the BAM-R contains items with both categorical and continuous response options, other psychometric evaluations have focused on the version of the BAM with categorical response options only (also called the BAM-D; see Nelson et al., 2014). Gaddy et al. (2018) examined the BAM-D factor structure using an EFA with a sample of 4955 veterans from the VA's Corporate Data Warehouse (CDW), a national dataset which stores veteran medical record data. Gaddy et al. found evidence supporting a 4-factor structure: Alcohol Use, Stressors, Risk, and Stability, comprised of 13 of the 17 BAM items, with subscale internal consistency varying from poor (Stability) to excellent (Alcohol use).

We are aware of only two other psychometric evaluations of the BAM-R. Hallinan et al. (2021) used a sample of 22,453 veterans from the VA's CDW. Veterans' BAM-R scores. Confirmatory factor analyses indicated the 4-factor structure fit the data better than the 3-factor structure. Further, more strict longitudinal measurement invariance assumptions were met for the 4-factor structure (scalar invariance), but not the 3-factor structure. In the only psychometric evaluation of the BAM-R with a non-Veteran sample (to our knowledge), Schumm et al. (2022) also found the 4-factor model was superior to the 3-factor model and varied internal consistency estimates across subscales. Because these represent the most stringent examinations of the BAM-R factor structure with the largest sample sizes to date, the 4-factor structure may be superior to the 3-factor structure.

1.2. Current study

MBC for SUD care is recommended, but in practice adoption has been slow (Resnick and Hoff, 2020). One reason may be due to the limited time in healthcare settings (Beehler and Wray, 2012). One way to reduce the time burden is to reduce the administration time of patient reported outcomes. Well-implemented disorder-specific MBC measures are brief (e.g., Patient Health Questionnaire –2 or 9). Because SUD is highly comorbid with mental health conditions, and patients with comorbidities are more likely to receive SUD treatment (Harris et al., 2019), multiple MBC measures may need to be administered in a single visit. The BAM-R takes about 5 min to administer, but recommendations by Boswell et al. (2015) state MBC measurement should take no more than 5 min. Thus, we aimed to create a shortened BAM-R that could be administered with additional MBC measures within 5 min total.

A brief yet comprehensive set of items to assess substance use and correlates relevant to treatment has the potential to quickly inform clinical decision-making. Further, as delivery of treatment of alcohol and SUDs becomes more common in non-addiction specialty settings, a brief MBC assessment tool that can be quickly and feasibly administered and interpreted is needed. To accomplish this, we focused on identifying BAM-R items with evidence of predictive validity and sensitivity to change during treatment. Of the available outcomes in the CDW dataset, we selected SUD treatment retention, a clinically relevant outcome that was examined in the development of the BAM-R (Cacciola et al., 2013). Given over 105,000 drug-related poisoning deaths occurred in U.S. in past year (Ahmad et al., 2023), we also selected mortality.

2. Material and methods

2.1. Participants

Data for the current study were retrieved through the VA's Corporate Data Warehouse (CDW). The CDW houses data from Veterans' electronic health records (EHR) across the U.S.. This study was approved by the Institutional Review Boards at the Philadelphia Veterans Affairs Medical Center and the VA Puget Sound Health Care System. We extracted data from records with BAM-R scores ($n = 40,679$). We then selected the first BAM-R administration from unique patients which met our inclusion criteria ($n = 33,899$). The inclusion criteria were 1) a new treatment episode between fiscal year (FY) 2014 and FY2019, defined as a SUD-specialty care visit (intake), with no SUD treatment visits in the prior 60 days (1 phone visit allowed), 2) a second SUD treatment visit in the next two weeks to ensure engagement in care (qualifying visit), and 3) completed BAM-R with no missing items in the week before or after an intake assessment (initial BAM; $n = 32,007$). Although the number of BAM-R assessments per Veteran ranged from 1 to 11, nearly 80% of Veterans had only one BAM-R. We used the initial BAM-R for Veterans with more than one BAM-R administration in their records except in the analyses examining changes in BAM-R scores over time. For sensitivity to change analyses, we used BAM-R scores from Veterans with a second BAM-R administration between 30 and 60 days after intake and no missing items on either the intake or follow-up BAM-R ($n = 7523$).

2.2. Measures

2.2.1. Demographics

Veteran age, race (Black, Native American, Asian/Pacific Islander, White, Multiple, Unknown), ethnicity (Hispanic/Latino, non-Hispanic/Latino, Unknown), and sex (i.e., female, male) were extracted from the medical record.

2.2.2. Brief Addiction Monitor-Revised (BAM-R)

The BAM-R consists of 17 items which comprise three subscales. The current Use subscale consists of three items. Two items assess alcohol frequency, Alcohol Use (4) and Heavy Alcohol Use (5), and one item assesses drug use frequency – Drug Use (6). If use of "illegal/street drugs or abuse of prescription medication" is reported on Drug Use (6), a second set of items, Specific Drug Use (7a-7 g) is administered. Items 7a-7 g assesses frequency of drug use within drug classes (i.e., marijuana, sedatives/ tranquilizers, cocaine/crack, other stimulants, opiates, inhalants, and other drugs, respectively). The current Risk subscale items include Physical Health (1), Sleep (2), Mood (3), Craving (8), Risky Situations (11), and Social Problems (15). The Protective subscale items include Confidence (9), Self-help (10), Spirituality (12), Work (13), Income (14), and Social Support (16). All items, except Confidence (which queries the respondent about the subsequent 30-day period), are

assessed within a past 30-day timeframe. Item 1 is measured on a 5-point ordinal scale with response options ranging from 0 (*excellent*) to 15 (*good*) to 30 (*poor*). Items 8, 9, 12, 15, and 17 are measured on a 5-point ordinal scale with response options ranging from 0 (*not at all*) to 15 (*moderately*) to 30 (*extremely*). Item 14 is measured on a binary scale with 30 (yes) and 0 (no) response options. Likert-type and binary items are scored on a 0–30 scale for consistency with items 2–7 g, 10, 11, 13 and 16, which use a continuous 0 to 30-day scale. This also allows for each item to have equal weighting when creating sum scores. Items for each subscale (Use, Risk, Protective) are summed, with higher scores reflecting higher use, risk, and protective factors. For item abbreviations, BAM-R item number, and item content, refer to supplementary Figure 1.

2.2.3. Clinical outcomes

Clinical outcomes available in the EHR to examine predictive validity of the BAM-R subscales and items included SUD treatment retention and 12-month all-cause mortality. Ninety-day SUD treatment retention was defined as attending a second SUD-specialty treatment visit within 14 days of intake, followed by at least three SUD treatment visits over the next 90 days (i.e., at least one visit between 1 and 30 days, at least one visit between 31 and 60 days, and at least 1 visit between 61 and 90 days following specialty treatment admission). All clinical outcomes were modeled as binary.

2.3. Analytic plan

The overall goal of these analyses was to reduce the number of BAM-R items while maintaining content from all subscales to yield a measure as short as possible to assess substance use, risk factors, and protective factors feasibly in healthcare systems.

2.3.1. Item-Level predictive validity

We used a series of logistic regression models to determine whether the intake BAM-R item scores could predict clinical outcomes using a within-sample approach. To identify potential items for elimination, we planned a series of regularized models – least absolute shrinkage and selection operator (lasso) logistic regressions – with an out-of-sample prediction approach. Unlike standard regression models, a lasso regression applies a penalty for model complexity, (i.e., models containing larger numbers of items), resulting in a simpler model that is not overfitted. For cross-validation, we split the training and testing data sets using an 80%/20% split for models predicting SUD treatment retention. We analyzed the data using a 70%/30% split for 12-month mortality due to low base rate to ensure an adequate number of deaths in the testing data set. The best models identified in the training data were evaluated in the testing data.

Based on BAM-R scoring guidelines and previous psychometric work (Gaddy et al., 2018; Hallinan et al., 2021), we removed items 7a-7g (Specific Drug Use), 13 (Work), 14 (Income), and 17 (Progress) from item-level analyses. Specifically, items 7a-7g and 17 are not included in the subscale scores, and items 13–14 were removed by Gaddy et al. (2018) due to low factor loadings. In the first set of item-level analyses, age, race, and sex were added to the model first. Next, we added each item individually to these covariates, followed by a model with covariates and all BAM items. In the regularized regression models, the 13 BAM items were tested simultaneously with no covariates. To evaluate these models, we examined area under the curve (AUC) using the recommendation of AUC of 0.70 to indicate reasonable predictive value (Hosmer et al., 2013). Analyses were conducted using SAS/STAT software version 9.2 of the SAS System © for Windows,[1] as well as R version 4.1 (R Core Team, 2021) & RStudio (RStudio Team, 2020) with packages [caret (Kuhn, 2021), foreign (R Core Team, 2020), glmnet (Friedman et al., 2010), ISLR (James et al., 2021), leaps (Lumley, 2020), pROC (Robin et al., al.,2011), and tidyverse (Wickham et al., 2019)].

2.3.2. Sensitivity to change

To identify which item scores demonstrated evidence of sensitivity to change (and therefore have more utility for MBC), we examined the amount of change between scores on two BAM-R administrations. We compared BAM-R scores between intake and the next follow-up BAM-R administration within 30–60 days following initial BAM-R administration (n = 7523). We examined item-level and original subscale-level changes and estimated Cohen's d effect sizes of mean differences.

2.3.3. Subset-Level predictive validity

Next, we examined the predictive validity of the 5 items we selected for a shortened BAM-R (see Section 4.4) and compared this to the original BAM-R subscale scores in predicting 90-day SUD treatment retention in specific subsamples. Substance use-specific subsamples were defined as follows: a) No use (n = 15,799; defined as Item 5 = 0 and Item 6 = 0 which indicates no heavy alcohol or drug use on initial BAM); b) Alcohol only (n = 9023 defined as Item 5 > 0, and Item 6 = 0); c) Opioid-primary (n = 1853; defined as Item 7e > than item 4,7a,7b,7c,7d,7f,7 g; d) Stimulant-primary (n = 2502; Items 7c or 7d > items 4,7a,7b,7e,7f,7 g); and e) Cannabis-primary (n = 2825; Item 7a > items 4,7b,7c,7d,7e,7f,7 g). These analyses use the same out-of-sample approach with an 80%/20% split for testing and training datasets. We then compared performance of the retained 5 items to each of the original 3 subscales separately and simultaneously (i.e., the full BAM-R currently recommended for MBC). The following models were examined: Model 1: 5 selected BAM items as main effects; Model 2: Risk (sum of items 1,2,3,8,11,15); Model 3: Protective (sum of item 9,10,12,13,14,16); Model 4: Use (sum of item 4,5,6); and Model 5: Risk, Protective, and Use sum scores as main effects.

3. Results

3.1. Sample description

The sample mean age was 49.6 (SD = 13.4). Based on demographics from the EHR, approximately 30% of Veterans identified as Black, 7% as Hispanic/Latino, and 61% as White, with 7% of the sample identified as female Veterans. See Table 1 for descriptive statistics. For clinical outcomes, 15,694 (49%) of Veteran records indicated 90-day SUD treatment retention, and 649 (2%) indicated 12-month mortality.

Table 1
Descriptive statistics.

Patient Demographics	Initial BAM-R (N = 32,002)		Follow-up BAM-R (N = 7523)	
Age (*M, SD*)	49.6	(13.4)	50.1	(13.1)
Race *N* (%)				
Asian/Pacific Islander	515	(1.6)	91	(1.2)
Black	9722	(30.4)	2388	(31.7)
Native American	332	(1.0)	92	(1.2)
White	19,478	(60.9)	4556	(60.6)
Multiple	405	(1.3)	95	(1.3)
Unknown	1550	(4.8)	301	(4.0)
Ethnicity *N* (%)				
Hispanic/Latino	2292	(7.2)	429	(5.7)
Non-Hispanic/Latino	28,767	(89.9)	6899	(91.7)
Unknown	943	(3.0)	195	(2.6)
Sex (%)				
Female	2362	(7.4)	529	(7.0)
Male	29,640	(92.6)	6994	(93.0)

[1] Copyright © 2008 SAS Institute Inc. SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc., Cary, NC, USA.

3.2. Item-level predictive validity

Item-level predictive validity analyses from standard logistic regression models indicated no individual items, nor the full set of intake BAM-R total scores, demonstrated evidence of predicting the available clinical outcomes. AUC for 90-day SUD treatment retention models ranged from 0.51–0.56. For 12-month mortality models, AUC ranged from 0.63–0.66. The larger AUCs for 12-month mortality compared to SUD treatment retention were primarily driven by inclusion of age in the model (AUC = 0.62). Table 2 shows the AUC values for logistic regression models. The regularized lasso regression eliminated 6 items when predicting 90-day SUD treatment retention and all but one item when predicting 12-month mortality. Because the AUC for both models was below 0.70 (0.57 and 0.66, respectively), the lasso regression findings were not used to eliminate items for the shortened BAM.

3.3. Sensitivity to change

Items which demonstrated the most change between intake and follow-up (i.e., moderate-to-large effect sizes, in descending order) were Heavy Alcohol Use (5), Self-Help (10), Drug Use (6), Craving (4), Mood (3), Progress (17), Risky Situations (11), Physical Health (1), Social Problems (15), and Sleep (2). Table 3 provides item means, standard deviations, and change score effect sizes (Cohen's *d*) for all items and subscales.

3.4. Item selection

We sought to retain as few items as possible to enhance utility of the BAM-R. Because our predictive validity analyses did not inform item removal, we selected items based on 1) perceived clinical utility as judged by expert clinicians (co-authors), and 2) largest magnitude of change over time. Items selected for perceived clinical utility in non-addiction specialty MBC settings included the following: Health (1), Sleep (2), Mood (3), Heavy Alcohol Use (5), Drug Use (6), Craving (8), and Self-Help (10). We chose to retain the items with the largest magnitude of sensitivity to change [i.e., Heavy Alcohol Use (5), Self-Help (10), Drug Use (6), Craving (4), and Mood (3), respectively]. Each subscale derived from previous work (i.e., 3- and 4-factor structures; Cacciola et al., 2013; Hallinan et al., 2021) is represented by 1–2 of the selected items.

3.5. BAM subset predictive validity

Consistent with our item-level predictive validity analyses, the selected 5-item subset and the three original BAM-R subscale intake scores failed to exhibit adequate prediction of 90-day SUD treatment retention across the full sample and substance use-specific subsamples (AUC range = 0.50 - 0.58; see Table 4).

Table 2
Areas under the curve for item-level univariate and full BAM models.

Model/Item(#)	AUC for 90-day Retention	AUC for 12-month Mortality
Age, Race, Sex	.511	.628
Physical Health (1)	.523	.649
Sleep (2)	.518	.635
Mood (3)	.524	.632
Any Alcohol Use (4)	.553	.629
Heavy Alcohol Use (5)	.547	.629
Drug Use (6)	.528	.630
Craving (8)	.523	.632
Confidence (9)	.530	.629
Self-help (10)	.529	.629
Risky Sit (11)	.530	.636
Spirituality (12)	.520	.629
Social Problems (15)	.521	.630
Social Support (16)	.523	.628
All BAM items	.567	.661

Note. AUC = area under the curve. See supplementary Figure 1 for item abbreviations and content.

Table 3
Item-level means, standard deviations, and change effect sizes.

Item(#)/Scale	Intake		Follow Up		Effect Size
	Mean	SD	Mean	SD	Cohen's *d*
Self-Help (10)	2.96	6.67	8.30	9.96	0.64
Progress (17)	3.39	1.30	3.97	1.03	0.50
Protective Subscale	79.24	31.12	94.30	31.94	0.48
Confidence (9)	3.72	1.40	4.02	1.21	0.22
Social Support (16)	14.92	12.43	17.56	12.14	0.21
Spirituality (12)	3.08	1.55	3.38	1.48	0.20
Work (13)	7.17	10.21	8.55	10.97	0.13
Income (14)	0.61	0.49	0.65	0.48	0.08
Any Alcohol (4)	1.14	2.86	0.62	2.55	−0.19
Sleep (2)	16.20	11.36	12.24	10.98	−0.35
Soc Prob (15)	2.39	1.34	1.93	1.10	−0.38
Physical Health (1)	3.38	1.05	2.91	1.03	−0.45
Risk (11)	8.20	10.79	3.84	7.73	−0.47
Mood (3)	15.46	11.29	10.06	10.16	−0.50
Craving (8)	2.91	1.39	2.18	1.11	−0.58
Drug Use (6)	6.51	10.17	1.79	5.81	−0.59
Risk Subscale	82.52	39.47	56.62	34.06	−0.70
Heavy Alcohol (5)	7.37	10.10	1.49	4.97	−0.78
Use Subscale	23.05	23.39	5.57	13.33	−0.95

Note. Please refer to supplementary Figure 1 for item and subscale content. Follow-up is defined as a BAM administration between 30 and 60 days after intake.

4. Discussion

The purpose of this study was to derive a shorter BAM-R for non-SUD specialty healthcare settings. For predictive validity analyses, we used logistic regression models to inform item selection. No combination of intake BAM-R item scores predicted 90-day SUD treatment retention or 12-month mortality. We selected 5 items based on sensitivity to change estimates and expert opinion of perceived item-level clinical utility, which also represent each of the previously identified factors of the BAM-R. Further predictive validity analyses indicated this 5-item subset was equivalent to the model with all three BAM-R subscales included, though neither predicted 90-day treatment retention or 12-month mortality.

4.1. Predictive validity

Unlike Cacciola et al. (2013), we did not find that intake BAM-R subscale scores successfully predicted treatment retention. One potential explanation for this discrepancy is the difference in retention length examined (90- versus 30-day SUD treatment retention). Alternatively, our sample was larger and more inclusive in terms of treatment type (all

Table 4
Areas under the curve for item sets.

Model	90-day SUD Treatment Retention					
	Full sample	No Use	Alcohol-Only	Opioid-Primary	Stimulant-Primary*	Cannabis-Primary
1	.57	.52	.54	.57	.56	.52
2	.54	.53	.54	.52	.56	.57
3	.56	.50	.55	.55	.58	.51
4	.56	.53	.55	.52	.58	.55
5	.53	.51	.52	.50	.55	.52

Note. Stimulant-primary subsample consists of stimulant and cocaine/crack items.
Model 1 = 5-item BAM; Model 2 = Risk (sum of items 1,2,3,8,11,15); Model 3 = Protective (sum of item 9,10,12,13,14,16); Model 4 = Use (sum of item 4,5,6); Model 5 = Risk, Protective, and Use.

SUD outpatient treatment v. intensive outpatient only), race (more gender and racially diverse sample v. primarily Black men), and site (national, multisite v. one clinic). Consistent with our findings, earlier work found that intake Addiction Severity Index's substance use status, drug and alcohol composite scores, medical, legal, family/social, and psychiatric domains did not predict SUD treatment retention (McCaul et al., 2001).

Our findings that no intake BAM-R subscale or combination of items predicted outcomes call into question the predictive validity of the BAM-R regarding SUD treatment retention and mortality. Because no predictive validity analyses have been conducted using subscales derived from the 4-factor solution or at the item level, our findings are difficult to contextualize. Although we sought to replicate and extend Cacciola et al. (2013) original work, in hindsight, examining criterion-related validity may be more appropriate to psychometrically evaluate MBC measures, which are meant to assess current state and change over time.

4.2. Item selection

We selected 5 items which demonstrated the largest magnitude of sensitivity to change and perceived clinical utility. Each of the previously proposed 3- and 4-factor structures (Cacciola et al., 2013; Hallinan et al., 2021) are also represented by 1–2 of the selected items.: Frequency of use is a standard metric in substance use treatment, and we retained items assessing frequency of heavy alcohol use and drug use. As noted by Clarke et al. (2021), the Craving (8) item is actionable for MBC and could facilitate conversations to increase therapeutic alliance and motivation. Further, a recent meta-analysis of 237 studies found that craving predicts drug use reliably (Vafaie and Kober, 2022). The Mood (3) item assesses depression, anxiety, and anger, all of which can be addressed with evidence-based treatments that can be delivered in non-addiction specialty integrated care settings (Zhang et al., 2019). Finally, the Self-help (10) item may facilitate conversations between patients and providers to identify potential additional sources of community recovery support.

In addition to the five items we selected, other BAM items may be clinically useful in non-SUD specialty settings. Although we did not evaluate each drug use item (7a-7 g), which asks about frequency of use of specific drugs, this is recommended for SUD MBC (Marsden et al., 2019). Additionally, we believe Health (1) and Sleep (2) are relevant and clinically actionable in faster-paced healthcare settings, and both demonstrated adequate sensitivity to change.

To our knowledge, there are currently no measures with strong evidence of psychometric soundness brief enough to deliver SUD MBC in healthcare settings. Although the Drug Abuse Screening Test (DAST) is used in clinical and research settings, this measure only assesses past year drug use (yes/no) and drug-related consequences. The Alcohol Use Disorders Identification Test – Consumption (AUDIT-C) is commonly used in clinical settings but only screens for alcohol use disorder. The shortened BAM-R items provide more nuanced drug use information (i.e., past-month use heavy alcohol and drug use). Of all the measures for MBC recommended by the Addiction Medicine Practice-Based Research Network, the BAM was the only measure which assesses drug and alcohol use, as well as risk and protective factors. Alternatively, another approach for MBC for opioid use disorder proposed by Marsden et al. (2019) called for assessing frequency of use and the 11 *DSM-5* criteria via clinical interview and recommends a single patient-reported outcome item to measure perceptions in improvement of OUD symptoms following treatment if time is limited. However, this approach also lacks psychometric evidence.

Although more work is needed, the potential benefits of a shortened BAM-R include standardized collection of data that is clinically relevant and actionable in fast-paced, non-SUD specialty healthcare settings. Regular assessment of these items can also help create a shared language and facilitate dialog between patients and providers about substance use and treatment. Although we did not administer this shortened BAM-R in clinical settings, we suspect this will take about 2 min.

4.3. Strengths and limitations

Notable strengths of the current study include use of the largest, nationally representative sample to attempt to produce a shorter version of the BAM-R. Although use of the VA's CDW allowed for such a large and diverse sample to be analyzed, we were limited to the data available in the CDW, which restricted the clinical outcomes available. By only including data from fully completed BAM-Rs for most analyses (no missingness) and at least 2 BAM-Rs for sensitivity to change analyses, our findings may not be generalizable to the full population of Veterans with an SUD, including those who choose to discontinue treatment early. Because inpatient treatment was not specified as an exclusion criterion, our sample may include individuals who completed outpatient and inpatient SUD treatment. Another potential limitation may be our retention definition, which could have excluded data from Veterans who were retained in treatment but had visits outside the 30-day timeframe we imposed or transitioned to inpatient care. Requiring more than one session per 30-day period to indicate retention might have produced different results. It is also possible that follow-up BAM-R administration scores or change scores might be better predictors of clinical outcomes, rather than intake scores, as effective MBC would reduce the strength of the association between intake scores and treatment outcomes. Because these data were from VA's CDW, our findings may not be generalizable beyond VA healthcare settings.

4.4. Future directions

Future research to shorten the BAM-R or identify items for use in SUD MBC should examine item-level convergent and concurrent criterion-related validity (e.g., biomarkers of heavy alcohol use, drug use, craving). For example, the Mood item could be correlated with Patient Health Questionnaire – 9 (PHQ-9) scores. If future research chooses to explore predictive validity, efforts may want to examine intake BAM-R scores among untreated samples and/or use change scores for samples receiving SUD treatment. Future research could determine the effectiveness of the shortened BAM-R (or another measure) in guiding MBC by randomizing patients or clinics to: (1) MBC with the shortened BAM-R, (2) MBC with some other instrument, or (3) treatment as usual without MBC. Further, based on recommendations by Lewis et al. (2019), in addition to an independent analysis of the selected item scores' sensitivity to change, future work could focus on identifying items to track that are considered most important by patient, providers, and clinic administration.

Future BAM-R research may benefit from determining whether items are more appropriately modeled as formative, reflective, or with mixed indicators with a vanishing tetrad test (Bollen and Ting, 2000). If the BAM-R should be modeled as formative, future psychometric work should focus on evaluating predictive validity, known groups validity, test-retest reliability, and clinical validity in MBC contexts (Murray and Booth, 2018).

5. Conclusions

Considering the limited measures available for implementation of MBC in general healthcare settings for treating people with SUD, we attempted to shorten the BAM-R, the current recommended measure for SUD MBC. Although intake BAM-R item scores did not demonstrate evidence of predictive validity, we found evidence of sensitivity to change for multiple items which were identified by experts as clinically actionable and meaningful. Further evaluation of the clinical utility and psychometric properties (e.g., test-retest reliability, criterion validity) of the BAM-R items and other brief measures is needed to optimize MBC for people receiving SUD treatment in healthcare settings.

CRediT authorship contribution statement

Brittany E. Blanchard: Methodology, Formal analysis, Writing – original draft. **Kevin G. Lynch:** Conceptualization, Methodology, Investigation, Formal analysis, Supervision, Writing – review & editing. **Carol A. Malte:** Conceptualization, Investigation, Data curation, Project administration, Writing – review & editing. **Eric J. Hawkins:** Conceptualization, Investigation, Writing – review & editing. **Dominick DePhilippis:** Conceptualization, Investigation, Writing – review & editing. **David W. Oslin:** Conceptualization, Investigation, Writing – review & editing. **James R. McKay:** Conceptualization, Funding acquisition, Investigation, Writing – review & editing. **Andrew J. Saxon:** Conceptualization, Funding acquisition, Investigation, Writing – review & editing.

Declaration of Competing Interest

None of the authors have any conflicts of interest to report. The views expressed in this article are those of these authors only and do not reflect the position or policies of any of the affiliated agencies.

Role of Funding Source

This work was supported by the National Institutes of Health/National Institute of Drug Abuse (UG1 DA013714-18) NIDA Clinical Trials Network Pacific Northwest Node and the National Institutes of Health/National Center for Advancing Translational Sciences' UW Institute of Translational Sciences KL2 program (KLTR002317). Dr. Blanchard is also supported by the National Institute of Drug Abuse Loan Repayment Program (L30DA056956). Funders were not involved in any aspect of the research or manuscript preparation.

Acknowledgments

N/A

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.dadr.2023.100183.

References

Ahmad, F.B., Cisewski, J.A., Rossen, L.M., Sutton, P., 2023. Provisional Drug Overdose Death Counts. National Center for Health Statistics. https://www.cdc.gov/nchs/nvss/vsrr/drug-overdose-data.htm#dashboard.

Beehler, G.P., Wray, L.O., 2012. Behavioral health providers' perspectives of delivering behavioral health services in primary care: a qualitative analysis. BMC Health Serv. Res. 12 (337), 1–12. https://doi.org/10.1186/1472-6963-12-337.

Bollen, K.A., Ting, K.F., 2000. A tetrad test for causal indicators. Psychol. Methods 5 (1), 3–22. https://doi.org/10.1037/1082-989X.5.1.3.

Boswell, J.F., Kraus, D.R., Miller, S.D., Lambert, M.J., 2015. Implementing routineoutcome monitoring in clinical practice: benefits, challenges, and solutions. Psychother. Res. 25 (1), 6–19. https://doi.org/10.1080/10503307.2013.817696.

Cacciola, J.S., Alterman, A.I., DePhilippis, D., Drapkin, M.L., Valadez Jr, C., Fala, N.C., McKay, J.R., 2013. Development and initial evaluation of the Brief Addiction Monitor (BAM). J. Subst. Abuse Treat. 44 (3), 256–263. https://doi.org/10.1016/j.jsat.2012.07.013.

Clarke, D.E., Ibrahim, A., Doty, B., Patel, S., Gibson, D., Pagano, A., Schwartz, R.P., 2021. Addiction Medicine Practice-Based Research Network (AMNet): assessment tools and quality measures. Subst. Abuse Rehabil. 12, 27–39. https://doi.org/10.2147/SAR.S305972.

Fortney, J.C., Unützer, J., Wrenn, G., Pyne, J.M., Smith, G.R., Schoenbaum, M., Harbin, H.T., 2017. A tipping point for measurement-based care. Psychiatr. Serv. 68 (2), 179–188. https://doi.org/10.1176/appi.ps.201500439.

Friedman, J., Tibshirani, R., Hastie, T., 2010. Regularization paths for generalized linear models via coordinate descent. J. Stat. Softw. 33 (1), 1–22. https://doi.org/10.18637/jss.v033.i01.

Gaddy, M.A., Casner, H.G., Rosinski, J., 2018. Factor structure and measurement invariance of the Brief Addiction Monitor. J. Subst. Use Treat. 90, 29–37. https://doi.org/10.1016/j.jsat.2018.04.010.

Goodman, J.D., McKay, J.R., DePhilippis, D., 2013. Progress monitoring in mental health and addiction treatment: a means of improving care. Prof. Psychol. 44 (4), 231. https://doi.org/10.1037/a0032605.

Hallinan, S., Gaddy, M., Ghosh, A., Burgen, E., 2021. Factor structure and measurement invariance of the Revised Brief Addiction Monitor. Psychol. Assess. 33 (3), 273–278. https://doi.org/10.1037/pas0000973.

Harding, K.J.K., Rush, A.J., Arbuckle, M., Trivedi, M.H., Pincus, H.A., 2011. Measurement-based care in psychiatric practice: a policy framework for implementation. J. Clin. Psychiatry 72 (8), 1136–1143. https://doi.org/10.4088/JCP.10r06282whi.

Harris, M.G., Bharat, C., Glantz, M.D., Sampson, N.A., Al-Hamzawi, A., Alonso, J., World Health Organization's World Mental Health Surveys collaborators, 2019. Cross-national patterns of substance use disorder treatment and associations with mental disorder comorbidity in the WHO World Mental Health Surveys. Addiction 114 (8), 1446–1459.

Hosmer Jr, D.W., Lemeshow, S., Sturdivant, R.X, 2013. Applied Logistic Regression, 398. John Wiley & Sons.

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2021). ISLR: data for an Introduction to Statistical Learning with Applications in R. https://CRAN.R-project.org/package=ISLR.

Kuhn, M. (2021). caret: classification and regression training. https://CRAN.R-project.org/package=caret.

Lewis, C.C., Boyd, M., Puspitasari, A., Navarro, E., Howard, J., Kassab, H., Kroenke, K., 2019. Implementing measurement-based care in behavioral health: a review. JAMA Psychiatry 76 (3), 324–335. https://doi.org/10.1001/jamapsychiatry.2018.3329.

Lumley, T. (2020). leaps: regression subset selection. https://CRAN.R-project.org/package=leaps.

Marsden, J., Tai, B., Ali, R., Hu, L., Rush, A.J., Volkow, N., 2019. Measurement-based care using DSM-5 for opioid use disorder: can we make opioid medication treatment more effective? Addiction 114 (8), 1346–1353. https://doi.org/10.1111/add.14546.

McCaul, M.E., Svikis, D.S., Moore, R.D., 2001. Predictors of outpatient treatment retention: patient versus substance use characteristics. Drug Alcohol Depend. 62 (1), 9–17. https://doi.org/10.1016/S0376-8716(00)00155-1.

McKay, J.R., Drapkin, M.L., Goodman, J., DePhilippis, D., 2009. Brief Addiction Monitor (BAM): A New Performance Measure in VA SUD Care. Department of Veterans Affairs, Washington, DC.

Murray, A.L., Booth, T., 2018. Causal indicators in psychometrics. Wiley Handb. Psychom. Test. 187–207.

Nelson, K.G., Young, K., Chapman, H., 2014. Examining the performance of the Brief Addiction Monitor. J. Subst. Abuse Treat. 46, 472–481. https://doi.org/10.1016/j.jsat.2013.07.002.

Peterson, K., Anderson, J., and Bourne, D. (2019). Evidence brief: use of patient reported outcome measures for measurement based care in mental health shared decision-making. https://europepmc.org/article/nbk/nbk536143.

R Core Team (2020). Foreign: Read Data Stored By 'Minitab', 'S', 'SAS', 'SPSS', 'Stata', 'Systat', 'Weka', 'dBase', ... https://CRAN.R-project.org/package=foreign.

R Core Team, 2021. R: A Language and Environment For Statistical Computing. R Foundation for Statistical Computing. https://www.R-project.org/.

Resnick, S.G., Hoff, R.A., 2020. Observations from the national implementation of measurement based care in mental health in the department of veterans affairs. Psychol. Serv. 17 (3), 238–246. https://doi.org/10.1037/ser0000351.

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.C., Müller, M., 2011. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinform. 12 (77), 1–8. http://www.biomedcentral.com/1471-2105/12/77/.

RStudio Team (2020). RStudio: Integrated Development For R. http://www.rstudio.com/.

Schumm, J., Wong, C., Okrant, E., Tharp, J.A., Embree, J., Lester, N., 2022. Factor structure of the Brief Addiction Monitor in a non-veteran substance use disorder outpatient treatment sample. Drug Alcohol Depend. Rep. 5, 100125.

Trivedi, M.H., Daly, E.J., 2007. Measurement-based care for refractory depression: a clinical decision support model for clinical research and practice. Drug Alcohol Depend. 88, S61–S71. https://doi.org/10.1016/j.drugalcdep.2007.01.007.

U.S. Department of Veterans Affairs, 2022. VHA Programs For Veterans with Substance Use Disorders. December 8. https://www.va.gov/vhapublications/publications.cfm?Pub=1.

Vafaie, N., Kober, H., 2022. Association of drug cues and craving with drug use and relapse: a systematic review and meta-analysis. JAMA Psychiatry 79 (1), 641–650. https://doi.org/10.1001/jamapsychiatry.2022.1240.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T.L., Miller, E., Bache, S.M., Müller, K., Ooms, J., Robinson, D., Seidel, D.P., Spinu, V., Yutani, H., 2019. Welcome to the tidyverse. J. Open Source Softw. 4 (43), 1686. https://doi.org/10.21105/joss.01686.

Wrenn, G., Fortney, J., Kennedy, P., Harbin, H., Carneal, G., Daviss, S., Heiman, H.J., Simon, K., Sladek, R., Unützer, J., Vinson, S., 2017. Issue Brief: A Core Set of Outcome Measures For Behavioral Health Across Service Settings. The Kennedy Forum. MBC_supplement.pdf (thekennedyforum.org.

Zhang, A., Franklin, C., Jing, S., Bornheimer, L.A., Hai, A.H., Himle, J.A., Ji, Q., 2019. The effectiveness of four empirically supported psychotherapies for primary care depression and anxiety: a systematic review and meta-analysis. J. Affect. Disord. 245, 1168–1186. https://doi.org/10.1016/j.jad.2018.12.008.