Quality of Care Measures for the Management of Unhealthy Alcohol Use

Kimberly A. Hepner, PhD[a,*], **Katherine E. Watkins, MD, MSHS**[a], **Carrie M. Farmer, PhD**[b], **Lisa Rubenstein, MD, MSHS**[a,c,e], **Eric R. Pedersen, PhD**[a], and **Harold Alan Pincus, MD**[a,d]

[a]RAND Corporation, 1776 Main Street, Santa Monica, CA, 90407, USA

[b]RAND Corporation, 4570 Fifth Avenue, Pittsburgh, PA, 15213, USA

[c]VA Greater Los Angeles at Sepulveda, 16111 Plummer St. (152), North Hills, CA, 91343, USA

[d]Department of Psychiatry, Columbia University and New York-Presbyterian Hospital, 1051 Riverside Drive, Unit 09, New York, NY 10032, USA

[e]Department of Medicine and School of Public Health, UCLA, 10833 Le Conte Avenue, Los Angeles, California 90095, USA

Abstract

There is a paucity of quality measures to assess the care for the range of unhealthy alcohol use, ranging from risky drinking to alcohol use disorders. Using a two-phase expert panel review process, we sought to develop an expanded set of quality of care measures for unhealthy alcohol use, focusing on outpatient care delivered in both primary care and specialty care settings. This process generated 25 candidate measures. Eight measures address screening and assessment, 11 address aspects of treatment, and six address follow-up. These quality measures represent high priority targets for future development, including creating detailed technical specifications and pilot testing them to evaluate their utility in terms of feasibility, reliability, and validity.

Keywords

Unhealthy alcohol use; quality measures; expert panel

1. Introduction

Unhealthy alcohol use, which includes the range of elevated alcohol use from risky drinking to alcohol use disorders (U.S. Preventive Services Task Force, 2013), is prevalent in the United States. Estimates suggest approximately 20 percent of primary care patients drink at

[*]**Corresponding author:** Kimberly A. Hepner, RAND Corporation, 1776 Main Street, P.O. Box 2138, Santa Monica, CA, USA 90407-2138, hepner@rand.org; (310) 393-0411, x6381.

Conflicts of interest: none

unhealthy levels (Saitz, 2005; Vinson et al., 2010). Recent data suggests approximately 30% of Americans, an estimated 68 million, will have an alcohol use disorder (AUD) during their lifetime (Grant et al., 2015).

Unhealthy alcohol use is linked to medical concerns (e.g., hypertension, stroke), sleep disturbances, depression and suicidal ideation, problems at work, sexually transmitted infections, injury and increased risk of accidents (Brady, 2006; Caputo, Trevisani, & Bernardi, 2007; Cherpitel & Ye, 2008; Cook & Clark, 2005; Corrao, Bagnardi, Zambon, & La Vecchia, 2004; Harada et al., 2015; Saitz, 2003; Sanchez et al., 2015). AUDs are associated with substantial psychiatric and medical co-morbidities (Fergusson, Boden, & Horwood, 2009; Freiberg et al., 2010; Najt, Fusar-Poli, & Brambilla, 2011; Rehm et al., 2009; Roerecke & Rehm, 2014; Schneier et al., 2010; Schuckit, 2009), approximately 88,000 deaths annually (Stahre, Roeber, Kanny, Brewer, & Zhang, 2014), and an estimated $249 billion in economic costs in 2010 (Sacks, Gonzales, Bouchery, Tomedi, & Brewer, 2015), a figure that has been steadily rising (Bouchery, Harwood, Sacks, Simon, & Brewer, 2011).

1.1 Quality of Care for Unhealthy alcohol use

Providing appropriate care could reduce the consequences of alcohol use. Clinical practice guidelines describe recommended care across the range of severity, including screening for unhealthy alcohol use, providing a brief intervention and, if indicated, effective forms of psychotherapy, pharmacotherapy, and referral to self-help groups (Kleber et al., 2006; National Institute for Health and Clinical Excellence, 2011; U.S. Department of Veterans Affairs and U.S. Department of Defense, 2015). Yet despite the availability of effective interventions, studies suggest that the quality of care for unhealthy alcohol use is poor, with most individuals remaining undetected and untreated (Boyle & Davis, 2006; Castle, Yi, Hingson, & White, 2014; Hingson, Heeren, Edwards, & Saitz, 2012; IOM (Institute of Medicine), 2001; McCarty, 2007). For example, a recent meta-analysis highlighted that clinicians have considerable difficulty with the identification of problem drinking in clinical practice, identifying about half of those with AUD when relying on clinical judgement and correctly recording an AUD in the chart notes in only one in three cases (Mitchell, Meader, Bird, & Rizzo, 2012). Further, only 45 percent of patients with unhealthy alcohol use reported being asked about their drinking by a general medical practitioner, and less than half of these patients received any type of counseling regarding their drinking levels (D'Amico, Paddock, Burnam, & Kung, 2005). Recent data suggests that fewer than 5 percent of individuals with a past-year AUD received treatment for their unhealthy alcohol use from a health care practitioner (Grant et al., 2015), and most do not receive minimally adequate treatment (Wang et al., 2005).

1.2 Quality Care Measures for Unhealthy alcohol use

While studies suggest that most individuals with unhealthy alcohol use do not receive recommended care, efforts to improve care for unhealthy alcohol use have been hampered by the paucity of validated quality measures. Quality measures are frequently derived from clinical practice guidelines and assess the degree to which care recommended for a particular patient was received. Quality measures often focus on assessing process of care

(i.e., the actions of a provider with a particular patient) because these measures are typically more feasible, provide more actionable information to providers, and can be more responsive to change (McGlynn & Adams, 2014; Ryan & Doran, 2012). However, significantly more attention has been placed on quality measures for mental health than for substance use disorders (SUDs) (Waraich et al., 2010; Watkins, Farmer, De Vries, & Hepner, 2015). Few quality measures specifically assess care for either unhealthy alcohol use or SUDs more broadly. It was recently highlighted that very few National Quality Forum endorsed measures assess care for substance use disorders (Watkins et al., 2015). While there are now NQF-endorsed measures that assess alcohol screening and brief intervention, few measures assess care for alcohol use that does not meet the threshold for an AUD diagnosis, leaving out care for the large number of patients presenting with undiagnosed or lower, but still risky, levels of unhealthy alcohol use in primary care settings (Saitz, 2005; Solberg, Maciosek, & Edwards, 2008).

In addition to the paucity of measures, existing measures suffer from several important limitations. The two most widely studied measures are based on utilization data and assess initiation and engagement with treatment for SUDs. There is some evidence that these measures are associated with modest improvements in patient outcomes (Dunigan et al., 2014; Garnick et al., 2014; Garnick et al., 2007; Harris, Humphreys, Bowe, Tiet, & Finney, 2010). Yet they rely on service utilization (e.g., number and timing of visits) and do not capture information about the process of care, such as whether the treatment delivered was evidence-based. This is an important omission and may explain why these measures have been only modestly associated with improved outcomes. Measures that assess whether evidence-based treatment was delivered, and the quality of the patient-provider interaction, may have stronger associations with outcomes and may better support quality improvement efforts.

1.3 The Present Study

To address the identified gaps in available measures, we sought to develop quality of care measures for unhealthy alcohol use, focusing on outpatient care delivered in both primary care and specialty care settings. We describe the process of developing these quality measures through a two-phase expert panel review process. The process included development of a preliminary list of measures based on literature review, one expert panel meeting with discussion of key themes, pre and post-meeting ratings of measures, and final selection of candidate measures for empirical validation in primary care and specialty care settings. Resulting quality measures focused on measure concepts (sometimes referred to as measure statements), rather than development of the full technical specifications for implementing the measures.

2. Materials and methods

We used a modified RAND/UCLA Appropriateness Method (Brook, 1995; McGory, Shekelle, & Ko, 2006; Shekelle, Maclean, Morton, & Wenger, 2001). This method entails providing experts with a synthesis of the best evidence on a particular topic and asking them to use their individual and collective expertise to generate judgments on a topic for which

there is little or no published evidence. This approach has been applied to numerous health conditions and interventions (Avery et al., 2011; Coulter, Adams, & Shekelle, 1995; Ostovar et al., 2010), as well to guideline development (Bernstein, Hofer, Meijler, & Rigter, 1997) and to rating treatment outcomes (Normand, Frank, & McGuire, 2002). In this study, we first identified candidate quality measures and then elicited two rounds of expert panel ratings with one face-to-face panel meeting between rounds. All procedures were approved by the RAND Human Subjects Protection Committee.

2.1 Identification of candidate quality measures

Twenty-five candidate measures were identified through a review of peer-reviewed literature on development and evaluation of quality measures related to care for alcohol and drug use (Horovitz-Lennon et al., 2009), clinical practice guidelines (Kleber et al., 2006; National Institute for Health and Clinical Excellence, 2011; U.S. Department of Veterans Affairs and U.S. Department of Defense, 2009), and measure databases (e.g., National Quality Forum). These original 25 measures are available from the first author upon request.

Like the majority of National Quality Forum endorsed measures (National Quality Forum, 2013), all 25 quality measures focused on process of care. Measures were defined using "IF…THEN" statements, where IF described the clinical presentation of targeted patients to whom the process applies (i.e., the denominator) and THEN described the process of care that should be applied under these circumstances (i.e., the numerator). For example, the measure for screening for co-occurring depression stated "IF patient has a new Alcohol Use Disorders Identification Test - Consumption (AUDIT-C) score ≥ 5 and no documented diagnosis of depression, THEN patient should be screened for Major Depressive Disorder within 30 days of the index visit." The performance rate for a quality measure is computed by dividing the number of patients who received the recommended process (i.e., the numerator) by the number of patients for whom the care process was recommended (i.e., the denominator).

2.2 Panel members

The panel was comprised of nine experts in prevention and management of unhealthy alcohol use. Panel members were selected to maximize diversity across a variety of characteristics including geographical location within the U.S., professional role (practitioner, researcher, and/or administrator), degree (M.D., Ph.D.), training background (internal/family medicine, psychiatry, psychology), institution type (academic medical center, VA, public sector, private sector academic), and treatment setting (primary care, mental health specialty care, and substance abuse specialty care). Panelists also had expertise in performance measurement, managed behavioral health care, and quality of care. The nine person panel consisted of six medical doctors and three clinical doctorates (PhDs). All were connected with medical centers with either administrative or research duties. Four were affiliated with the Veterans Health Administration (VHA) and all had more than 15 years of experience. Five were located in the northeast United States, three were located along the west coast, and one was located in the southeast. Panelists received an honorarium and travel expenses.

2.3 Round 1 Elicitation: Review and ratings of candidate measures

Prior to the face-to-face meeting, panelists received brief conceptual definitions of each candidate measure (i.e., IF-THEN statements), a measure rating form, and a summary of the supporting evidence for each candidate measure. Panelists were able to add comments regarding their ratings, suggestions for modifications (e.g., different time frames) or additional relevant literature, or to propose new candidate quality measures. Panel rating materials are available from the first author upon request.

In Round 1, panel members were asked to rate the validity, feasibility of national implementation, and importance of each candidate measure on a 1 to 9 point scale, where 1 = definitely not valid/definitely not feasible/not at all important, 5 = uncertain or equivocal validity/uncertain or variable feasibility/moderately important, and 9 = definitely valid/ definitely feasible/very important. Validity, feasibility, and importance were selected as target domains due to their use in prior expert panel processes focused on quality measure development that incorporate the RAND/UCLA Appropriateness Method (Brook, 1995; McGory et al., 2006; Shekelle et al., 2001). These domains also map closely to the evaluation criteria for NQF measure endorsement (National Quality Forum, 2015). Panelists were provided the following definition of validity: "We define a measure to be valid if adequate scientific evidence or professional consensus exists to support a link between the performance of care specified by the measure and the accrual of health benefits to patients with unhealthy alcohol use (e.g. physical, mental, social); a practitioner or health plan with significantly higher rates of adherence to a measure would be considered a higher quality provider; and a majority of factors that determine adherence to a measure are under the influence of the practitioner or health plan (or are subject to influence, such as smoking cessation)." Feasibility of national implementation referred to the availability of data to generate the measure reliably. Importance referred to the existence of an important quality gap, a high prevalence of patients to whom the process applied or to the expectation of harmful consequences to patients if the measure was not adhered to.

2.4 Round 2 Re-Elicitation: Face-to-face panel meetings and ratings of final measures

Panelists then participated in one two-day in-person meeting. During the meeting, panelists were provided their individual Round 1 ratings and a summary of panelist ratings (i.e., mean and standard deviation for each candidate measure). At the meeting, project leaders led discussions prompted by cross-cutting topics emerging from the Round 1 ratings. The discussion started with overarching issues, followed by a discussion of each of the 25 candidate quality measures in turn. To facilitate discussion, measures were grouped by their Round 1 validity ratings as being rated highly (6.5 validity ratings and above; five measures), controversial (varying validity ratings from 3.9 to 6.3; seven measures), medium (ratings of 5.0 to 6.3; eight measures), or low (ratings of 4.1 to 4.4; five measures). Other studies utilizing the RAND/UCLA Appropriateness Method have focused solely on validity ratings (e.g., (McGory et al., 2006; Wenger, Roth, Shekelle, & Acove Investigators, 2007), so this was a useful domain to organize discussion.

During the discussion of individual measures, most were reworded or clarified based on group discussion. New measures and alternative specifications were proposed by panelists

and new evidence was presented in some cases for the appropriateness of these new measures. Panelist ratings and discussion resulted in some Round 1 measures being dropped (e.g., screen for Posttraumatic Stress Disorder, physical exam, family involvement, offer of employment needs), and resulted in some new measures being developed (e.g., screen for suicidal ideation, repeat brief intervention). During discussion, an additional 24 measures were drafted based on refinements of the candidate measures and proposal of new measures by panelists. This set of 49 measures included some that represented alternative approaches to measuring a particular process of care. These 49 measures were rated for validity using the 1 to 9 rating measure at the end of the meeting. Due to time constraints, these final ratings focused only on validity and excluded feasibility and importance ratings.

2.5 Analysis of Panelist Ratings

Following the Round 2 ratings, we selected measures with a final mean validity rating of 6.0 or greater. This value, although slightly lower than the rating of 7.0 utilized in other studies using the RAND/UCLA Appropriateness Method (e.g., (McGory et al., 2006; Shekelle et al., 2001), was selected to ensure that all measures with a higher likelihood of validity would be candidates for testing in future work. Occasionally more than one measure statement for the same process was rated as 6.0 or higher and we selected the highest rated measure statement for that particular process. Measures were reviewed and modified slightly for clarity, consistency, and applicability in both primary care and specialty care settings by the study team.

3. Results

For Round 1, mean validity ratings for individual quality measures ranged from 3.9 to 7.4, feasibility ranged from 4.2 to 7.7, and importance ranged from 4.2 to 7.6. Detailed results of the first round ratings are not reported in detail, as these ratings are used primarily to facilitate discussion and consensus. Of the 49 measures rated in Round 2, the project team selected 25 quality measures for the final set based on the criteria outlined above (i.e., validity rating above 6.0 and ensuring measures did not assess the same process of care). Table 1 includes each final measure statement, the phase of care the process occurs (screening and assessment, treatment, follow-up), and the mean and standard deviation for the validity rating. Eight measures address screening and assessment (mean validity ratings ranged from 6.1 to 7.6), 11 address aspects of treatment (mean validity ratings ranged from 6.2 to 8.2), and six address follow-up (mean ratings ranged from 6.8 to 7.7). Time frames for each measure varied, with most screening, assessment, and treatment measures recommended within the first 30 days of the index visit (i.e., date of positive unhealthy alcohol use screen), with shorter time frames for more severe patients with immediate concerns (e.g., follow-up treatment offered within seven days of detoxification). Follow-up measures were generally endorsed for the next routine visit or within three to six months.

4. Discussion

Using a modified RAND/UCLA Appropriateness Method (Brook, 1995; Shekelle et al., 2001), we used an expert panel review process and generated 25 quality of care measures that assess care for unhealthy alcohol use across primary care and specialty care outpatient

settings. These measures assess multiple phases of care including screening and assessment, treatment, and follow-up, and reflect panelist expertise regarding how care is delivered in both primary care and specialty mental health care settings. Unlike some existing quality measures, these measures address care across the broad spectrum of unhealthy alcohol use severity found across both primary care and specialty settings, rather than focusing on care for patients diagnosed with an alcohol use disorder. The measures also differ from some existing measures in that they are specific to care for unhealthy alcohol use; existing predominantly measures focus on substance use disorders more broadly (e.g., initiation of treatment, engagement of treatment; (Garnick, Lee, Horgan, Acevedo, & Washington Circle Public Sector, 2009)). The greater breadth in terms of addressing the needs of the full spectrum of unhealthy alcohol use severity combined with a specific focus on alcohol assessment, treatment and follow-up could strengthen the links between the quality of measured care and outcomes (Garnick, Horgan, Acevedo, McCorry, & Weisner, 2012). These measures could also better support quality improvement by identifying care deficits more specifically.

Our work addresses the multiple calls for additional, valid measures of care for AUDs and unhealthy alcohol use (Garnick, Horgan, & Chalk, 2006; Herbstman & Pincus, 2009; McCorry, Garnick, Bartlett, Cotter, & Chalk, 2000; Watkins et al., 2015). Further, some of the measures identified focus on the quality of psychosocial interventions (e.g., brief intervention, psychotherapy), a high priority area for measure development highlighted in a recent Institute of Medicine report (2015).

4.1 Implementation Issues

Our expert panel process generated 25 quality measures to assess care for unhealthy alcohol use. While this is an important step, and begins to address a need for these measures, there are several additional steps that are required before the highest priority measures can be identified and these measures can be implemented in a manner that conforms to the criteria for endorsement by the National Quality Forum. In this section, we highlight core considerations related to developing, evaluating, and implementing quality measures for unhealthy alcohol use.

4.1.1 Developing technical specifications—As described earlier, the quality measures resulting from this expert panel process are in the form of measure statements (i.e., IF… THEN statements), representing the measure concepts rather than the detailed technical specifications required to implement the measures. Technical specifications include the specific diagnosis codes, procedure codes, and time frames to reliably identify the eligible population and whether the recommended process of care occurred. For example, these specifications would indicate whether patients with co-occurring drug use disorders or other psychiatric diagnoses should be included. Some measures in our set are adaptations of existing measures (e.g., initiation and engagement) for which detailed specifications could be adapted, whereas new measures would need more development. Specifications have been developed for a subset of these measures as part of a multiyear study to evaluate their predictive validity (Mattox et al., 2016).

4.1.2 Testing measures to evaluate reliability, feasibility, and predictive validity

—Our expert panel process generated several measures that met a minimum a priori threshold for validity (i.e., 6 or higher), based on panel ratings. While this suggests these measures have face validity, it will still be important to evaluate these measures in terms of their reliability, feasibility, and predictive validity. For example, measures will need to be tested to ensure that they can be reliably generated across different health care systems. Relatedly, measures must be assessed for their feasibility for implementation. Availability of necessary data elements is a primary driver of feasibility. For example, this panel process generated measures that will likely draw on multiple data sources, including administrative data, electronic health records, and medical record review. While it has been suggested that only utilizing administrative data for quality measures may be insufficient and multiple sources of data collection may yield a more comprehensive picture of care quality (Kilbourne, Keyser, & Pincus, 2010), administrative data-based measures remain more feasible than quality measures requiring other types of data. Measures that incorporate electronic health record data or medical record review, while less feasible, capture clinical detail and complexity that administrative data typically cannot. Examples include discussion of treatment options and quality of psychotherapy. Incorporating electronic health record or medical record review data elements will be a particular challenge for specialty behavioral health settings, which are less likely than general medical settings to have implemented electronic medical records.

Finally, quality measures need to be evaluated in terms of their relationship to important clinical and patient outcomes. While better process of care should improve patient outcomes (Donabedian & Bashur, 2002; Ryan & Doran, 2012), it is important to validate existing and new quality measures in terms of their ability to predict patient outcomes before using them to incentivize changes in care (McGlynn & Adams, 2014; McLellan, Chalk, & Bartlett, 2007). Demonstrating this predictive validity helps to increase the likelihood that improvements in the quality of care delivered actually result in improved outcomes. Additional data on reliability, feasibility, and validity will provide essential information to guide identification of high priority measures from among the 25 candidate measures developed from our expert panel process.

4.1.3 Consider routine clinical practice rather than ideal practice—There is a risk of developing measures that cannot be met by most providers with typical resources. Indeed, during the panel process, some panel members were concerned that the processes of care assessed by some of the measures may be difficult to achieve in some practice settings due to lack of resources. Thus, the 25 candidate measures resulting from our panel process took under consideration the resources of a typical practice, aiming not to set unreasonably high measure standards. This involved not always requiring the process to be completed on the same day, or allowing another provider to complete the process. Time frames for completion of the process were often extended to ensure adequate time for completion. As the measures are further developed and evaluated, it will be important to ensure the resulting measures are based in an understanding of the demands of real world clinical practice.

4.1.4 Considering treatment setting—As measures are developed, evaluated, and implemented, the setting of care for which the measure applies must be considered. Regardless of whether care is received in the primary or specialty care setting, patients should receive appropriate care. Yet, historically, quality measures have often been targeted to particular care settings, with different expectations for appropriate care based on the capabilities of that setting. For example, for a newly identified patient with unhealthy alcohol use, requiring only a screen for co-occurring depression may be an appropriate standard for care delivered in primary care settings, but a diagnostic assessment for depression would be a more appropriate standard for specialty mental health and substance use specialty settings. Higher standards for appropriate care can be set for patients who are seen in specialty care settings because of the qualifications and resources of those settings compared to primary care practices. Yet an alternative perspective is that the health care system is responsible for delivering high quality care to patients with unhealthy alcohol use, regardless of whether they are seen in specialty care or not. This is an important issue for further consideration as quality measures for unhealthy alcohol use are further developed and evaluated.

4.1.5 Accountability measures and unintended consequences—Quality measures that are publicly reported or tied to incentives (financial or otherwise), are often referred to as accountability or performance measures. Before implementing quality measures, it is important to consider how the measures could be used for accountability (e.g., financial or other incentives), whether they are used to monitor the performance of individual providers or of a facility, and what unintended consequences might follow. Once incentivized, measures can be "gamed" to improve performance rates. For example, if a quality measure assesses whether a provider delivered a brief intervention following identification of unhealthy alcohol use, performance could appear better by simply not identifying patients with unhealthy alcohol use (Bradley et al., 2013; Harris, Rubinsky, & Hoggatt, 2015). Ensuring that there are quality measures that track identification of patients with unhealthy alcohol use is an essential strategy that is needed to accurately assess the quality of downstream care.

4.2 Limitations

We acknowledge some limitations. While an expert panel approach is generally an acceptable method for generating expert opinions (Powell, 2003; Rowe & Wright, 1999), it is possible that the views of the nine experts who participated may differ from other experts in this area. For example, some types of clinicians who delivered care for unhealthy alcohol use were not included (e.g., social workers, addiction counselors). Further, panelists were asked to rate and discuss measure statements (i.e., IF…THEN statements), which represent the concept of the measure, but did not review full detailed specifications for each measure. We used this approach to generate a large group of quality measures for unhealthy alcohol use that could be promising candidates for future work. As highlighted earlier, the selected measures will need further development and evaluation, and most would yet not be suitable currently for widespread implementation because they still need detailed specifications to support implementation across health care systems. Further, we present a larger list of candidate measures than will likely be ultimately implemented. The validity ratings

presented Table 1 provide some indication of the relative value of these candidate quality measures at the conceptual level, but the priority of measures could evolve as they are developed and evaluated. Due to time constraints, we did not elicit final ratings on feasibility and importance. Finally, the 2009 VA/DoD clinical practice guidelines for substance use disorders were reviewed in preparation for this expert panel process (U.S. Department of Veterans Affairs and U.S. Department of Defense, 2009) and updated guidelines were recently released (U.S. Department of Veterans Affairs and U.S. Department of Defense, 2015). While the resulting measures are consistent with the new guidelines, they were not available prior to the panel process.

5. Conclusions

We identified 25 quality measures that represent an initial step towards increasing the number of measures to assess care for unhealthy alcohol use. These measures are high priority targets for further development, evaluation, and implementation, and address the need for more mechanisms to monitor and improve quality of care for unhealthy alcohol use.

Acknowledgments

References

Avery AJ, Dex GM, Mulvaney C, Serumaga B, Spencer R, Lester HE, Campbell SM. Development of prescribing-safety indicators for GPs using the RAND appropriateness method. British Journal of General Practice. 2011; 61(589):e526–e536. [PubMed: 21801572]

Bernstein SJ, Hofer TP, Meijler AP, Rigter H. Setting standards for effectiveness: A comparison of expert panels and decision analysis. International Journal for Quality in Health Care. 1997; 9:255–263. [PubMed: 9304424]

Bouchery EE, Harwood HJ, Sacks JJ, Simon CJ, Brewer RD. Economic costs of excessive alcohol consumption in the U.S., 2006. American Journal of Preventive Medicine. 2011; 41(5):516–524. [PubMed: 22011424]

Boyle AR, Davis H. Early screening and assessment of alcohol and substance abuse in the elderly: Clinical implications. Journal of Addictions Nursing. 2006; 17(2):95–103.

Bradley KA, Chavez LJ, Lapham GT, Williams EC, Achtmeyer C, Rubinsky AD, Kivlahan DR. When quality indicators undermine quality: Bias in a quality indicator of follow-up for alcohol misuse. Psychiatric Services. 2013; 64(10):1018–1025. [PubMed: 23852137]

Brady J. The association between alcohol misuse and suicidal behaviour. Alcohol and Alcoholism. 2006; 41(5):473–478. [PubMed: 16891335]

Brook, RH. The RAND/UCLA appropriateness method. Santa Monica, CA: RAND Corporation; 1995.

Caputo F, Trevisani F, Bernardi M. Alcohol misuse and traffic accidents. Lancet. 2007; 369(9560): 463–464.

Castle IP, Yi H, Hingson RW, White AM. State variation in underreporting of alcohol involvement on death certificates: Motor vehicle traffic crash fatalities as an example. Journal of Studies on Alcohol and Drugs. 2014; 75:299–312. [PubMed: 24650824]

Cherpitel CJ, Ye Y. Alcohol-attributable fraction for injury in the U.S. general population: Data from the 2005 national alcohol survey. Journal of Studies on Alcohol and Drugs. 2008; 69:535–538. [PubMed: 18612569]

Cook RL, Clark DB. Is there an association between alcohol consumption and sexually transmitted diseases? A systematic review. Sexually Transmitted Diseases. 2005; 32(3):156–164. [PubMed: 15729152]

Corrao G, Bagnardi V, Zambon A, La Vecchia C. A meta-analysis of alcohol consumption and the risk of 15 diseases. Preventive Medicine. 2004; 38:613–619. [PubMed: 15066364]

Coulter I, Adams A, Shekelle P. Impact of varying panel membership on ratings of appropriateness in consensus panels: A comparison of a multi- and single disciplinary panel. Health Services Research. 1995; 30:577–591. [PubMed: 7591782]

D'Amico EJ, Paddock SM, Burnam A, Kung FY. Identification of and guidance for problem drinking by general medical providers: Results from a national survey. Medical Care. 2005; 43(3):229–236. [PubMed: 15725979]

Donabedian, A., Bashur, R. An introduction to quality assurance in health care. Oxford: Oxford University Press; 2002.

Dunigan R, Acevedo A, Campbell K, Garnick DW, Horgan CM, Huber A, Ritter GA. Engagement in outpatient substance abuse treatment and employment outcomes. Journal of Behavioral Health Services and Research. 2014; 41(1)doi: 10.1007/s11414-11013-19334-11412

Fergusson DM, Boden JM, Horwood LJ. Tests of causal links between alcohol abuse or dependence and major depression. Archives of General Psychiatry. 2009; 66(3):260–266. [PubMed: 19255375]

Freiberg MS, McGinnis KA, Kraemer K, Samet JH, Conigliaro J, Ellison RC, Justice AC. The association between alcohol consumption and prevalent cardiovascular diseases among HIV infected and uninfected men. Journal of Acquired Immune Deficiency Syndromes. 2010; 53(2):247–253. [PubMed: 20009766]

Garnick DW, Horgan CM, Acevedo A, Lee MT, Panas L, Ritter GA, Wright D. Criminal justice outcomes after engagement in outpatient substance abuse treatment. Journal of Substance Abuse Treatment. 2014; 46(3):295–305. [PubMed: 24238717]

Garnick DW, Horgan CM, Acevedo A, McCorry F, Weisner C. Performance measures for substance use disorders–what research is needed? Addiction Science & Clinical Practice. 2012; 7:18. [PubMed: 23186374]

Garnick DW, Horgan CM, Chalk M. Performance measures for alcohol and other drug services. Alcohol Research & Health. 2006; 29(1):19–26. [PubMed: 16767849]

Garnick DW, Horgan CM, Lee MT, Panas L, Ritter GA, Davis S, Reynolds M. Are Washington Circle performance measures associated with decreased criminal activity following treatment? Journal of Substance Abuse Treatment. 2007; 33(4):341–352. [PubMed: 17524596]

Garnick DW, Lee MT, Horgan CM, Acevedo A, Washington Circle Public Sector, Workgroup. Adapting Washington Circle performance measures for public sector substance abuse treatment systems. Journal of Substance Abuse Treatment. 2009; 36(3):265–277. [PubMed: 18722075]

Grant BF, Goldstein RB, Saha TD, Chou SP, Jung J, Zhang H, Hasin DS. Epidemiology of DSM-5 alcohol use disorder: Results from the national epidemiologic survey on alcohol and related conditions III. JAMA Psychiatry. 2015; 72(8):757–766. [PubMed: 26039070]

Harada MY, Gangi A, Ko A, Liou DZ, Barmparas G, Li T, Ley EJ. Bicycle trauma and alcohol intoxication. International Journal of Surgery. 2015; 24:14–19. [PubMed: 26493212]

Harris AHS, Humphreys K, Bowe T, Tiet Q, Finney JW. Does meeting the HEDIS substance abuse treatment engagement criterion predict patient outcomes? Journal of Behavioral Health Services and Research. 2010; 37(1):25–39. [PubMed: 18770044]

Harris AHS, Rubinsky AD, Hoggatt KJ. Possible alternatives to diagnosis-based denominators for addiction treatment quality measures. Journal of Substance Abuse Treatment. 2015; 58:62–66. [PubMed: 26251046]

Herbstman BJ, Pincus HA. Measuring mental healthcare quality in the United States: A review of initiatives. Current Opinion in Psychiatry. 2009; 22(6):623–630. [PubMed: 19745741]

Hingson RW, Heeren T, Edwards EM, Saitz R. Young adults at risk for excess alcohol consumption are often not asked or counseled about drinking alcohol. Journal of General Internal Medicine. 2012; 27(2):179–184. [PubMed: 21935753]

Horovitz-Lennon, M., Watkins, KE., Pincus, HA., Shugarman, LR., Smith, B., Mattox, T., Mannle, TE, Jr. Veterans Health Administration mental health program evaluation technical manual. Santa Monica, CA: RAND Corporation; 2009.

IOM (Institute of Medicine). Crossing the quality chasm: A new health system for the 21st century. Washington, DC: National Academies Press; 2001.

IOM (Institute of Medicine). Psychosocial interventions for mental and substance use disorders: A framework for establishing evidence-based standards. Washington, DC: National Academies Press; 2015.

Kilbourne AM, Keyser D, Pincus HA. Challenges and opportunities in measuring the quality of mental health care. Canadian Journal of Psychiatry. Revue Canadienne de Psychiatrie. 2010; 55(9):549–557. [PubMed: 20840802]

Kleber, HD., Weiss, RD., Anton, RF., Jr, George, TP., Greenfield, SF., Kosten, TR., Connery, HS. Practice guideline for the treatment of patients with substance use disorders. Second. Washington, DC: American Psychiatric Press; 2006.

Mattox, T., Hepner, KA., Kivlahan, DR., Farmer, CM., Rosenbluth, S., Hoggatt, KJ., Watkins, KE. Candidate quality measures to assess care for alcohol misuse: Technical specifications. Santa Monica, CA: RAND Corporation; 2016.

McCarty D. Performance measurement for systems treating alcohol and drug use disorders. Journal of Substance Abuse Treatment. 2007; 33:353–354. [PubMed: 17596902]

McCorry F, Garnick DW, Bartlett J, Cotter F, Chalk M. Developing performance measures for alcohol and other drug services in managed care plans. Washington Circle Group. Joint Commission Journal on Quality Improvement. 2000; 26(11):633–643. [PubMed: 11098426]

McGlynn EA, Adams JL. What makes a good quality measure? JAMA. 2014; 312(15):1517–1518. [PubMed: 25321907]

McGory ML, Shekelle PG, Ko CY. Development of quality indicators for patients undergoing colorectal cancer surgery. Journal of the National Cancer Institute. 2006; 98(22):1623–1633. [PubMed: 17105985]

McLellan AT, Chalk M, Bartlett J. Outcomes, performance, and quality: What's the difference? Journal of Substance Abuse Treatment. 2007; 32(4):331–340. [PubMed: 17481456]

Mitchell AJ, Meader N, Bird V, Rizzo M. Clinical recognition and recording of alcohol disorders by clinicians in primary and secondary care: Meta-analysis. British Journal of Psychiatry. 2012; 201(2):93–100. [PubMed: 22859576]

Najt P, Fusar-Poli P, Brambilla P. Co-occurring mental and substance abuse disorders: A review on the potential predictors and clinical outcomes. Psychiatry Research. 2011; 186(2–3):159–164. [PubMed: 20728943]

National Institute for Health and Clinical Excellence. Alcohol dependence and harmful alcohol use (CG115). 2011. Retrieved April 19, 2016, from http://guidance.nice.org.uk/CG115

National Quality Forum. Quality positioning system. 2013. Retrieved January 2, 2015, from http://www.qualityforum.org/QPS/QPSTool.aspx

National Quality Forum. Measure evaluation criteria and guidance for evaluating measures for endorsement. 2015. Retrieved August 20, 2016, from http://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=79434

Normand SL, Frank RG, McGuire TG. Using elicitation techniques to estimate the value of ambulatory treatments for major depression. Medical Decision Making. 2002; 22:245–261. [PubMed: 12058782]

Ostovar R, Rashidian A, Pourreza A, Rashidi BH, Hantooshzadeh S, Ardebili HE, Mahmoudi M. Developing criteria for cesarean section using the RAND appropriateness method. BMC Pregnancy and Childbirth. 2010; 10(1):1–8. [PubMed: 20059767]

Powell C. The Delphi technique: Myths and realities. Journal of Advanced Nursing. 2003; 41(4):376–382. [PubMed: 12581103]

Rehm J, Mathers C, Popova S, Thavorncharoensap M, Teerawattananon Y, Patra J. Global burden of disease and injury and economic cost attributable to alcohol use and alcohol-use disorders. Lancet. 2009; 373:2223–2233. [PubMed: 19560604]

Roerecke M, Rehm J. Cause-specific mortality risk in alcohol use disorder treatment patients: A systematic review and meta-analysis. International Journal of Epidemiology. 2014; 43(3):906–919. [PubMed: 24513684]

Rowe G, Wright G. The Delphi technique as a forecasting tool: Issues and analysis. International Journal of Forecasting. 1999; 15(4):353–375.

Ryan AM, Doran T. The effect of improving processes of care on patient outcomes: Evidence from the United Kingdom's quality and outcomes framework. Medical Care. 2012; 50(3):191–199. [PubMed: 22329994]

Sacks JJ, Gonzales KR, Bouchery EE, Tomedi LE, Brewer RD. 2010 national and state costs of excessive alcohol consumption. American Journal of Preventive Medicine. 2015; 49(5):e73–79. [PubMed: 26477807]

Saitz, R. Overview of medical and surgical complications. In: Graham, A.Schultz, T.Mayo-Smith, M.Ries, R., Wilford, B., editors. Principles of addiction medicine. 3rd. Chevy Chase, MD: American Society of Addiction Medicine; 2003. p. 1027-1052.

Saitz R. Clinical practice. Unhealthy alcohol use. New England Journal of Medicine. 2005; 352(6):596–607. [PubMed: 15703424]

Sanchez K, Walker R, Campbell AN, Greer TL, Hu MC, Grannemann BD, Trivedi MH. Depressive symptoms and associated clinical characteristics in outpatients seeking community-based treatment for alcohol and drug problems. Substance Abuse. 2015; 36(3):297–303. [PubMed: 25084694]

Schneier FR, Foose TE, Hasin DS, Heimberg RG, Liu SM, Grant BF, Blanco C. Social anxiety disorder and alcohol use disorder co-morbidity in the national epidemiologic survey on alcohol and related conditions. Psychological Medicine. 2010; 40(6):977–988. [PubMed: 20441690]

Schuckit MA. Alcohol-use disorders. Lancet. 2009; 373:492–501. [PubMed: 19168210]

Shekelle PG, Maclean C, Morton SC, Wenger N. Assessing care of vulnerable elders: Methods for developing quality indicators. Annals of Internal Medicine. 2001; 135(8)

Solberg LI, Maciosek MV, Edwards NM. Primary care intervention to reduce alcohol misuse ranking its health impact and cost effectiveness. American Journal of Preventive Medicine. 2008; 34(2):143–152. [PubMed: 18201645]

Stahre M, Roeber J, Kanny D, Brewer RD, Zhang X. Contribution of excessive alcohol consumption to deaths and years of potential life lost in the United States. Preventing Chronic Disease. 2014; 11:130293.

U.S. Department of Veterans Affairs and U.S. Department of Defense. VA/DoD clinical practice guideline for management of substance use disorders, version 2.0. 2009. Retrieved September 10, 2013, from www.healthquality.va.gov/guidelines/MH/sud/sud_full_601f.pdf

U.S. Department of Veterans Affairs and U.S. Department of Defense. VA/DoD clinical practice guideline for management of substance use disorders, version 3.0. 2015. Retrieved April 19, 2016, from http://www.healthquality.va.gov/guidelines/MH/sud/

U.S. Preventive Services Task Force. Screening and behavioral counseling interventions in primary care to reduce alcohol misuse: Recommendation statement. 2013. AHRQ publication no. 12-05171-EF-3Retrieved January 18, 2014, from http://www.uspreventiveservicestaskforce.org/uspstf12/alcmisuse/alcmisusefinalrs.htm

Vinson DC, Manning BK, Galliher JM, Dickinson LM, Pace WD, Turner BJ. Alcohol and sleep problems in primary care patients: A report from the AAFP national research network. Annals of Family Medicine. 2010; 8(6):484–492. [PubMed: 21060117]

Wang PS, Lane M, Olfson M, Pincus HA, Wells KB, Kessler RC. Twelvemonth use of mental health services in the United States: Results from the national comorbidity survey replication. Archives of General Psychiatry. 2005; 62(6):629–640. [PubMed: 15939840]

Waraich P, Saklikar RS, Aube D, Jones W, Haslam D, Hamill K. Quality measures for primary mental healthcare: A multistakeholder, multijurisdictional Canadian consensus. Quality & Safety in Health Care. 2010; 19(6):519–525. [PubMed: 20702436]

Watkins KE, Farmer CM, De Vries D, Hepner KA. The Affordable Care Act: An opportunity for improving care for substance use disorders? Psychiatric Services. 2015; 66(3):310–312. [PubMed: 25727120]

Wenger NS, Roth CP, Shekelle P, Acove Investigators. Introduction to the assessing care of vulnerable elders-3 quality indicator measurement set. Journal of the American Geriatrics Society. 2007; 55(Suppl 2):S247–252. [PubMed: 17910544]

Highlights

- A two-phase expert panel process generated 25 quality measures.

- Measures assess primary and specialty outpatient care for unhealthy alcohol use.

- Quality measures identified represent high priority targets for future development.

Table 1

Final Consensus-Based Quality Measures for Alcohol Misuse

Measure	Measure Statement	Treatment Phase	*Validity Mean (SD)*
Adequate Identification of Unhealthy Alcohol Use	Proportion of all patients screened for unhealthy alcohol use	Screening and Assessment	7.0 (1.7)
Assess for AUD	IF patient has a new AUDIT-C score 5, THEN patient should be assessed for an alcohol use disorder (AUD) within 30 days.	Screening and Assessment	6.4 (1.0)
Assess for Major Depressive Disorder in Specialty Care	IF patient has a new AUDIT-C >= 8 or new diagnosis of AUD, THEN patient should be assessed for major depressive disorder within 30 days before or after index visit (specialty care).	Screening and Assessment	7.1 (2.4)
Screen for Major Depressive Disorder in Primary Care	IF patient has a new AUDIT-C >= 8 or new diagnosis of AUD, THEN patient should be screened for major depressive disorder within 30 days before or after index visit (primary care).	Screening and Assessment	6.4 (2.2)
Screen for Suicidal Ideation	IF patient has AUDIT-C >=8 or an AUD diagnosis, THEN patient should be screened for suicide within 30 days before or after index visit.	Screening and Assessment	6.1 (1.9)
Liver Function Test	IF patient has AUDIT-C >=8 or an AUD diagnosis, THEN patient should be screened for liver disease within 30 days before or after index visit.	Screening and Assessment	6.3 (2.3)
Screening for Other Substance Use in Any Care Setting	IF patient has a new AUDIT-C score 5, THEN patient should be screened for other substance use (including tobacco) within 30 days of the index visit.	Screening and Assessment	6.9 (1.9)
Assessment of Drug Use in Specialty Care	IF patient has a new AUDIT-C score >=8 and is seeing a mental health specialist, THEN patient should be assessed for other substance use (including type, frequency, and recency) within 30 days of the index visit.	Screening and Assessment	7.6 (1.4)
Brief Intervention	IF patient has a new AUDIT-C score 5, THEN patient should receive a brief intervention within 30 days following the index visit.	Treatment	8.2 (1.1)
Discuss Treatment Options	IF patient has a new AUDIT-C score 8 or an AUD, THEN patient should receive counseling regarding treatment options within 30 days following the index visit.	Treatment	7.1 (1.1)
Psychotherapy Offer	IF patient has AUDIT-C >= 8 or an AUD, THEN patient should be offered psychotherapy.	Treatment	6.2 (1.9)
Psychotherapy Dose	IF patient has AUDIT-C >= 8 or an AUD and one or more psychotherapy visits, THEN patient should receive at least 4 visits within the first 12 weeks.	Treatment	7.2 (2.0)
Psychotherapy Quality	IF patient has AUDIT-C >= 8 or an AUD and one or more psychotherapy visits, THEN the visit should include elements of an evidence-based psychotherapy.	Treatment	6.3 (2.4)
Pharmacotherapy for Alcohol Dependence	IF patient has a newly identified diagnosis of alcohol dependence, THEN patient should receive pharmacotherapy for alcohol dependence within 90 days following identification.	Treatment	7.1 (1.8)
Referral to Recovery Support in the Community	IF patient has a new diagnosis of alcohol dependence, THEN patient should be referred to recovery support in the community (e.g. Alcoholics Anonymous) within six months following identification.	Treatment	6.3 (1.9)

Measure	Measure Statement	Treatment Phase	*Validity Mean (SD)*
Offer of Housing Services	IF patient has a new AUDIT-C score 8 or an AUD and a documented housing need, THEN patient should be offered housing services within 30 days following identification of the need.	Treatment	6.1 (1.5)
Integrated Co-Occurring Disorder Treatment	IF patient has a current mental health diagnosis and a new AUDIT-C >= 8 or an AUD, THEN there should be evidence that both conditions are addressed as evidenced by treatment goals for both conditions, an integrated treatment plan, OR continuous engagement for both conditions.	Treatment	6.9 (0.9)
Treatment Initiation for Alcohol Dependence	IF patient has a newly identified diagnosis of an AUD, THEN patient should have either an inpatient AUD admission or both an initial AUD-related outpatient visit and an additional AUD-related visit within 30 days of the index visit.	Treatment	6.2 (1.2)
Treatment Engagement for Alcohol Dependence	IF patient has a newly identified AUD diagnosis and has initiated treatment, THEN patient should receive two additional alcohol-related visits within 30 days following treatment initiation.	Treatment	6.9 (1.4)
Reassess Alcohol Use	IF patient has a new AUDIT-C score 5, THEN patient should have their quantity and frequency of drinking reassessed within 30 days using a structured metric.	Follow-up	6.8 (1.9)
Repeat Brief Intervention	IF patient has a new AUDIT-C >= 5, THEN patient should receive TWO brief interventions within two months of the index visit.	Follow-up	7.6 (1.7)
Pharmacotherapy Evaluation and Management	IF patient started on new medication for alcohol dependence, THEN patient should have at least one alcohol-related follow-up encounter within 30 days of the index visit.	Follow-up	7.1 (1.1)
Any Alcohol-Related Follow-Up	IF patient has a new AUDIT-C score 5, THEN alcohol should be addressed at next routine visit.	Follow-up	6.9 (0.9)
Detox Follow-up	IF patient receives medication assisted detox, THEN patient should receive alcohol-related outpatient follow-up within 7 days.	Follow-up	7.3 (1.1)
Follow-up Consistent with Chronic Care Management	IF patient receives alcohol treatment, THEN patient should be re-evaluated quarterly and treatment adjusted if necessary.	Follow-up	7.7 (0.9)

Note: Panelists rated statements that referred to "abuse/dependence" and this language has been updated to indicate "alcohol use disorder." Index visit refers to the visit in which alcohol misuse was detected.