Journal of Substance Abuse Treatment 58 (2015) 62–66

Contents lists available at ScienceDirect

Journal of Substance Abuse Treatment

Possible Alternatives to Diagnosis-Based Denominators for Addiction Treatment Quality Measures

Alex H.S. Harris, Ph.D. [a,*], Anna D. Rubinsky, Ph.D. [a], Katherine J. Hoggatt, Ph.D. [b]

[a] Center for Innovation to Implementation and Substance Use Disorder Quality Enhancement Research Initiative, VA Palo Alto Health Care System, Palo Alto, CA, United States
[b] VA Greater Los Angeles Health Care System and University of California Los Angeles, Los Angeles, CA, USA

ARTICLE INFO

Article history:
Received 15 August 2014
Received in revised form 3 June 2015
Accepted 4 June 2015

Keywords:
Quality measurement
Quality comparisons
Validity
Methodology

ABSTRACT

Consumers of healthcare quality measures are often unaware of how measured performance may be driven by diagnosing practices rather than the provision of high quality care. Reliance on quality metrics that depend on documented diagnoses can therefore subvert comparisons between clinicians, facilities or systems. In this study, three versions of an alcohol use disorder (AUD) treatment quality measure were calculated: method 1—the usual denominator including all diagnosed patients; method 2—a "population-based" denominator including the entire facility census; and method 3—an epidemiologically-derived denominator comprising the expected prevalence of AUD based on case-mix characteristics and geographic region. Performance rankings under the three specifications were calculated. Changes in percentile rank of up to 30–45% were observed between methods. Therefore, much of the observed between-facility differences on diagnosis-based quality measures may reflect variation in the propensity to diagnose rather than real differences in performance. Stakeholders must decide which of the validity threats produced by these different methods is least worrisome.

Published by Elsevier Inc.

1. Introduction

1.1. The challenge of diagnosis-based denominators

Efforts to improve the quality of substance use disorder (SUD) treatment depend on having valid definitions and measures of quality. Health care systems use quality measures to identify high and low performing facilities, to design and target quality improvement programs, to monitor the effects of system-wide initiatives, and to incentivize particular patient care and other practices (Donabedian, 2003). Measures of SUD treatment quality are generally expressed as proportions, with the number of individuals who receive guideline-congruent or evidence-based treatment in the numerator and the number of individuals who need treatment in the denominator. Accurately measuring this denominator, although often difficult, is nonetheless critically important. The validity of measured performance, and the conclusions drawn about quality of treatment, depends on whether the denominator adequately captures the target population.

For quality measures tailored to patients in specialty SUD treatment settings (e.g., outpatient follow-up after SUD residential treatment or detoxification), operationalizing the denominator can be relatively easy, for example by using a combination of qualifying diagnoses, treatment locations, and procedure codes (Harris et al., 2015). However, for quality measures that focus on entire systems of care, including primary care, behavioral health, and other specialty medical settings, operationalizing the denominator is more challenging. When the proportion of patients who have SUD accurately diagnosed and documented in the medical record varies substantially between facilities or systems, independent of actual SUD prevalence (Song et al., 2010), differences in measured quality may fail to reflect actual differences in quality of care.

Some health care systems may tend to only diagnose those patients who are interested or involved in treatment. Under-diagnosis may also occur because population-based assessments are infeasible due to providers' limited time and multiple competing demands, due to insurance or regulatory concerns (Manuel, Newville, Larios, & Sorensen, 2013), or in response to implementation of quality measures that introduce accountability to provide certain services once a diagnosis is made (Roth et al., 2012). Other systems may have active programs of assessment that identify broader groups of patients with particular disorders, only some of whom are interested in treatment. Such differences in case finding and propensity to diagnose may dramatically affect measured performance based on diagnosis-denominated quality measures.

The ideal denominator for many process quality measures would include all patients who would benefit from the evidence-based treatment, independent of which clinical diagnoses they happened to receive. As noted above, this does not pose as much of a problem for measures tailored to patients in specialty SUD settings, or in cases such as childhood immunizations where the target population is the entire population with very narrow exceptions. However, when no method exists to operationalize which patients in an integrated health

* Corresponding author at: VA Palo Alto Health Care System, Center for Innovation to Implementation and Substance Use Disorder, Quality Enhancement Research Initiative, 795 Willow Road, Menlo Park, CA 94025, United States. Tel.: +1 408 2121 320.
E-mail address: Alexander.Harris2@va.gov (A.H.S. Harris).

http://dx.doi.org/10.1016/j.jsat.2015.06.004
0740-5472/Published by Elsevier Inc.

care system might benefit from treatment, we often depend on clinical diagnoses as a proxy for potential to benefit. In this paper, we discuss how cross-system variation in the propensity to diagnose and document SUD may undermine the validity of cross-system comparisons using traditional diagnosis-based quality measures. We then consider alternate approaches to measuring quality and discuss their relative pros and cons.

1.2. An example

Allowing the clinicians and facilities whose performance is being assessed to influence the composition of the population they are responsible to treat can cause unintended consequences and seriously impact the subsequent validity of cross-system comparisons (Bradley et al., 2013; Doran, Fullwood, Reeves, Gravelle, & Roland, 2008; Doran et al., 2006; Hayward & Kent, 2008; Roth et al., 2012; Wachter, Flanders, Fee, & Pronovost, 2008). To illustrate this problem, assume that health care facilities A and B have the same underlying 6% prevalence of alcohol use disorder (AUD). Facility A only diagnoses 1/6th of those patients with AUD and facility B diagnoses 5/6th of those patients with AUD (1 vs. 5% of total patients respectively). Evidence-based treatment is provided to 50% of the diagnosed patients in facility A and to 25% of the diagnosed patients in facility B. Who is doing better in terms of measured performance and who is doing better in terms of real performance?

Using the usual construction of quality measures, with a diagnosis-based denominator, facility A is performing twice as well as facility B in providing evidence-based treatment for AUD (50 vs. 25%). Therefore, facility A is superior in terms of measured performance. However, facility A is treating only 8.3% (1/6th × 50%) of their patients with AUD and facility B is treating 20.8% (5/6th × 25%) of their patients with AUD, meaning that facility B is doing over twice as well as facility A in real performance. The dramatic reversal of rank in measured vs. real performance is entirely driven by the differences in the gap between diagnosed and actual prevalence.

1.3. Alternative denominators

In Fig. 1, we describe alternative methods of calculating process quality measures; the usual diagnosis-denominated construction is labeled method 1. If we assume for a moment that every patient with a disorder is diagnosed (and none without the diagnosis), and that the underlying prevalence of the disorder in question is constant across facilities being evaluated, then changing the denominator from the number of patients who are diagnosed within the observation period, as in method 1, to the total number of patients served at each facility (method 2) will not affect the rank order of facilities. If rank order changes dramatically between methods 1 and 2, then either one or both of the assumptions above are problematic. The question then is which problem is more worrisome: (1) distortions caused by differences between real and diagnosed prevalence or (2) distortions caused by the probably erroneous assumption that real prevalence is constant across facilities. If we are more worried about the later, we should prefer method 1 (which allows for variation in prevalence across facilities). If we are more worried about the former, we should prefer method 2 (which avoids rewarding under-diagnosis). However, if we are worried about both assumptions, we need another approach.

Method 1 (M1)	# of diagnosed patients receiving evidence-based treatment for disorder / # of patients diagnosed with diagnosis
Method 2 (M2)	# of diagnosed patients receiving evidence-based treatment for disorder / Total number of patients served
Method 3 (M3)	# of diagnosed patients receiving evidence-based treatment for disorder / Age-, gender-, and regionally-adjusted estimated number of patients with disorder

Fig. 1. Three alternative methods for constructing process quality measures.

To avoid assuming that the diagnosed prevalence reflects the real prevalence, and as a hedge against the unrealistic assumption of equal prevalence across facilities or systems, we might instead compute epidemiologically-based estimates of the real prevalence, adjusted for demographic and geographic factors (labeled method 3 in Fig. 1). These estimates would not be perfect measures of real prevalence, but they might be less bad than the two alternatives already mentioned.

Although it is impossible to know which of these methods maps most accurately onto real process quality, the present study sought to examine two related questions: (1) Does it matter which method we choose? If rank order remains invariant regardless of method, then deciding which method has better properties is more academic than practical; (2) If rank order does change when different methods are used, is there any evidence that facilities' better performance under method 1 is due to under-diagnosis of the conditions that determine the denominator, relative to their expected prevalence?

2. Materials and methods

We examined these three alternative methods for calculating a measure of pharmacotherapy for alcohol use disorder (AUD) used by the US Veterans Health Administration (VHA) for system monitoring and quality improvement purposes (Trafton et al., 2013). The VHA measure is defined as the proportion of patients during a measurement year who received a clinical diagnosis of AUD (denominator) who also filled a prescription for an FDA-approved medication for the treatment of AUD, including oral and extended release naltrexone, acamprosate, and disulfiram (numerator). This measure is very similar to the measure of the same construct developed by the Washington Circle (Thomas et al., 2013).

Data to construct the measures were derived from the fiscal year 2010 (FY10) VHA National Patient Care Database and Decision Support System inpatient and outpatient pharmacy files. VHA is organized into 138 major facilities that are the unit of aggregation for accountability and reporting of most VHA quality measures. Using these data, we calculated alternative versions of the AUD pharmacotherapy quality measure using the methods described in Fig. 1. The numerator followed the measure specifications given above and was consistent across all methods. The denominator for method 1, following the usual measure specifications, was the number of patients in each facility with a recorded clinical diagnosis of AUD in FY10. The denominator for method 2 was a count of all patients in each facility with any clinical encounter during FY10, regardless of diagnosis. Note that another form of this alternative would be the number of patients treated per 1000, which might yield a measure on a scale that is easier to interpret. Patients who were seen at more than one facility were assigned to the facility in which they received the most outpatient clinical encounters.

The denominator for method 3 was an epidemiologically-derived expected count of patients who met diagnostic criteria for AUD given the facility's gender and age distribution and geographic region. To make these estimates, we identified a source of publically-available data on AUD prevalence, the National Surveys on Drug Use and Health (NSDUH). The NSDUH is an annual survey sponsored by the Substance Abuse and Mental Health Services Administration (SAMHSA) that collects data on substance use and mental health from randomly sampled civilians aged 12 or older residing in households or non-institutionalized group quarters or living on military bases. Confidential computer-assisted interviews are conducted in participants homes using Audio Computer-Assisted Self-Interview software, and a series of questions are asked to assess alcohol or other drug disorders in the past year based on *Diagnostic and Statistical Manual of Mental Disorders*, 4th edition (DSM-IV) specifications. Detailed information on the study

Download English Version:

https://daneshyari.com/en/article/328786

Download Persian Version:

https://daneshyari.com/article/328786

Daneshyari.com