Are Improvements in Measured Performance Driven by Better Treatment or “Denominator Management”?

Alex H. S. Harris, PhD¹, Cheng Chen, MS¹, Anna D. Rubinsky, PhD¹, Katherine J. Hoggatt, PhD², Matthew Neuman, PhD¹, and Megan E. Vanneman, PhD^{1,3}

¹Center for Innovation to Implementation, VA Palo Alto Health Care System, Menlo Park, CA, USA; ²VA Greater Los Angeles Health Care System and University of California Los Angeles, Los Angeles, CA, USA; ³Center for Health Policy/Primary Care and Outcomes Research, Stanford University School of Medicine, Stanford, CA, USA.

BACKGROUND: Process measures of healthcare quality are usually formulated as the number of patients who receive evidence-based treatment (numerator) divided by the number of patients in the target population (denominator). When the systems being evaluated can influence which patients are included in the denominator, it is reasonable to wonder if improvements in measured quality are driven by expanding numerators or contracting denominators.

OBJECTIVE: In 2003, the US Department of Veteran Affairs (VA) based executive compensation in part on performance on a substance use disorder (SUD) continuity-of-care quality measure. The first goal of this study was to evaluate if implementing the measure in this way resulted in expected improvements in measured performance. The second goal was to examine if the proportion of patients with SUD who qualified for the denominator contracted after the quality measure was implemented, and to describe the facility-level variation in and correlates of denominator contraction or expansion.

DESIGN: Using 40 quarters of data straddling the implementation of the performance measure, an interrupted time series design was used to evaluate changes in two outcomes.

PARTICIPANTS: All veterans with an SUD diagnosis in all VA facilities from fiscal year 2000 to 2009.

MAIN MEASURES: The two outcomes were 1) measured performance—patients retained/patients qualified and 2) denominator prevalence—patients qualified/patients with SUD program contact.

KEY RESULTS: Measured performance improved over time ($P < 0.001$). Notably, the proportion of patients with SUD program contact who qualified for the denominator decreased more rapidly after the measure was implemented ($p = 0.02$). Facilities with higher pre-implementation denominator prevalence had steeper declines in denominator prevalence after implementation ($p < 0.001$).

CONCLUSIONS: These results should motivate the development of measures that are less vulnerable to denominator management, and also the exploration of “shadow measures” to monitor and reduce undesirable denominator management.

KEY WORDS: quality measurement; measurement reactivity; performance contracts.

J Gen Intern Med 31(Suppl 1):S21–7

DOI: 10.1007/s11606-015-3558-1

© Society of General Internal Medicine 2015

INTRODUCTION

Process measures of healthcare quality are usually formulated as the number of patients receiving guideline-congruent treatment (numerator) divided by the number of patients in the target population thought likely to benefit (denominator). The target population is commonly operationalized by particular diagnoses, positive screening tests, or utilization-based criteria. When the systems, programs, or clinicians who are being evaluated can influence which patients are included in the denominator, it is reasonable to wonder if improvements in measured quality are driven by expanding numerators, which is the goal, or contracting denominators, which may be an unintended consequence.

In their satirical commentary “6 EZ Steps to Improving Your Performance (or How to Make P4P Pay 4U!),” Hayward and Kent noted that “there is no better way to improve your quality than to devote your entire practice to the aggressive diagnosis and management of ultra-early or ultra-mild disease states.”¹ The suggestion is that providers can improve their measured performance by “denominator management”—flooding their quality measure denominators with patients who they are sure will meet the numerator criteria. A corollary of this strategy is to make sure that sicker, more complicated, or less compliant patients do not meet the denominator criteria in the first place.

Many unintended consequences of incentivized performance measures have been described and studied.^{2–8} However, the empirical study of denominator management, as well as methods for detecting it, has received far less attention.^{9–11} In one recent study, Roth and colleagues studied changes in measured performance and denominator qualification for the 2007 Healthcare Effectiveness Data and Information Set (HEDIS) measure on avoidance of antibiotics for the treatment of acute bronchitis. Using data from a primary care network, they found that measured performance significantly increased

Electronic supplementary material The online version of this article (doi:10.1007/s11606-015-3558-1) contains supplementary material, which is available to authorized users.

Published online March 7, 2016

even when overall antibiotic prescribing did not decline, a situation mostly driven by the dramatic changes in diagnostic coding relevant to the denominator. Use of the diagnosis code 490 (Bronchitis, not otherwise specified), which is not included in the HEDIS measure specifications, increased from 1.5 % of total bronchitis visits in Year 1 to 84.6 % in Year 3.

Other studies have focused on selection or management strategies that impact inclusion of patients in the denominator of various public reporting metrics. Konetzka et al. found that nursing homes tend to re-hospitalize higher-risk, post-acute patients before the length-of-stay qualification criteria are met, thereby avoiding the inclusion of these patients in quality measures.¹¹ Other studies have found that implementation of healthcare report cards for surgeons performing coronary artery bypass graft (CABG) led to a decline in illness severity of patients receiving CABG, fewer intensive cardiac procedures among sicker patients eligible for CABG, less within-hospital variation in illness severity,¹² and increased racial disparities in CABG receipt.¹³ Thus, at least in the contexts examined so far, such as a primary care quality measure and CABG report cards, denominator management appears to occur when performance monitoring systems are implemented.

To further explore this aspect of quality measure validity and reactivity in a different context, we conducted a study of the implementation of another high-profile performance measure in a large integrated healthcare system, where the facilities and programs that were being evaluated had opportunities for denominator management. In order to improve the quality of specialty substance use disorder (SUD) treatment in its roughly 220 addiction treatment programs, which are nationally distributed over 131 VA facilities, the US Department of Veteran Affairs (VA) included an SUD Continuity of Care (CoC) measure in Network Directors’ performance contracts in fiscal year (FY) 2003. At the time, performance-based bonuses could amount to 10 % of VA Network Directors’ annual salaries.^{14–16} In addition, financial incentives were often provided to managers and clinicians at VA facilities, at the discretion of Network Directors, based upon the results of the performance contract.¹⁵

To qualify for the CoC measure denominator, patients needed to attend three visits at a VA specialty SUD treatment program within 30 days. Presumably (and anecdotally), it was easy for the programs to monitor which patients were at the threshold of qualifying, and either schedule the third qualifying visit within the qualification window or not, depending on whether the patient was expected to be retained and meet the numerator criteria. However, the extent to which this happened has not been systematically evaluated. For a variety of reasons,¹⁴ the CoC measure is no longer emphasized by VA. Thus, the purpose of this study was not to inform the implementation or interpretation of this specific measure, but rather to provide an empirical investigation of the measure reactivity and putative denominator management in a previously unexamined context.

We explored several specific hypotheses and questions. The first goal of this study was to evaluate if, as expected, inclusion

of the CoC measure in Network Directors’ performance contracts resulted in increases in measured performance. The second goal was to examine if implementation of the quality measure was associated with unintended effects; specifically, changes in the proportion of patients with SUD program contact who qualified for the denominator (“denominator prevalence”). Finally, we sought to describe the distribution of *changes* in denominator prevalence across VA facilities, and examine if these differences were correlated with other facility-level characteristics, including improvement of the CoC measure.

METHODS

Measures

The VA CoC Measure (“measured performance”) is defined as the percent of patients entering a specialty SUD treatment program who are retained in treatment for at least 90 days. Outpatients qualify for the denominator when they have three visits to an SUD specialty care clinic within a 30-day period, and did not have other SUD specialty care visits in the 90 days prior to the first of the three qualifying visits. Inpatients qualify for the denominator when discharged from an SUD inpatient stay lasting at least 4 days. The retention time period for outpatients starts the day after the third qualifying visit, and for inpatients, it starts at discharge or transfer to a non-SUD specialty unit. Veterans seen in multiple facilities are attributed to the facility where the last retention visit occurred in order to promote coordination between facilities. The vast majority of patients qualify under the outpatient criteria.

To be included in the numerator, patients must have at least two visits in an outpatient specialty SUD clinic every 30 days for three consecutive months. Appropriately documented telephone encounters are considered as retention visits in the second and third 30-day periods. SUD inpatient care less than 4 days in duration during the retention period is ignored. If the length of the SUD inpatient stay during the retention period is at least 4 days, then the patient is dropped from the previous qualifying track and re-qualifies for new monitoring at discharge from this stay. In each subsequent year, the 80th percentile of performance in the previous year is set as the new performance threshold facilities needed to be met in order to get the performance bonus.

Denominator Prevalence was defined as the number of patients qualified for the measure denominator divided by the number of patients receiving at least one visit in any SUD specialty treatment program. This definition was chosen because we thought that denominator management would most likely operate within SUD specialty care by managing patient visits to avoid meeting the denominator specifications. However, denominator management might have also occurred by restricting access to SUD specialty care. Therefore, an alternative definition of denominator prevalence was examined—the number of patients qualified for the measure

denominator divided by the number of patients with an SUD diagnosis (primary or secondary) recorded in any clinical encounter for each facility and observation period. Results pertaining to the alternative definition are presented in [Appendix](#) (available online).

Pre-implementation Facility-level Characteristics. Pre-implementation measured performance and denominator prevalence were calculated with data from FY2002. Pre-implementation SUD prevalence was defined as the number of patients with an SUD diagnosis recorded in any clinical encounter divided by the number of patients with at least one healthcare encounter during FY2002.

Design and Data Analysis

The official VA CoC performance measure algorithm was applied to VA National Patient Care Database data for each of 40 quarters from FY2000 to FY2009. We used an interrupted time series design, analyzed with mixed-effects segmented linear regression, to assess 1) changes in the level of measured performance and the rate of change in measured performance, and 2) changes in the level of and rate of change in denominator prevalence.^{17,18} Four parameters were estimated for each outcome: 1) the baseline (Quarter 1 (Q1) level of the outcome), 2) the rate of change (slope) in the period before the measure was implemented (i.e., slope from Q1-Q12), 3) the immediate change (gap) in the outcome when the measure was implemented, and 4) the change in the slope for the post-implementation period compared to the pre-implementation period. The CoC measure was implemented in Q12. All models were adjusted for strong seasonal effects using a decomposition method and centered moving averages to estimate seasonal indexes.¹⁹

Using the regression models above examining denominator prevalence, we then estimated parameters 3 and 4 (“denominator reactivity”) for each of the VA facilities, described the distributions and facility-level 95 % confidence intervals (CIs), and evaluated their associations with several facility-level characteristics including pre-implementation SUD prevalence, pre-implementation denominator prevalence, pre-

implementation performance on the CoC measure, and pre-to-post implementation improvement on the CoC measure.

RESULTS

Implementation of the CoC measure was not associated with an immediate change in performance, but the rate of improvement in the post-implementation period increased significantly relative to the rate of improvement in the pre-implementation period (see Table 1 and Fig. 1). At the start of the observation period (12 quarters before implementation of the measure), the average performance nationwide was 20.8 %, which slowly and not significantly increased 0.2 % per quarter on average throughout the pre-implementation period. Just before the measure was added to performance contracts, average performance was 23.1 % and there was not an immediate change when the measure was implemented. However, the rate of increase more than tripled to 0.7 % per quarter ($p < 0.001$) after the measure was implemented, meaning that the estimated performance was 48.3 % by the end of the observation period.

At the start of the observation period (12 quarters before implementation of the measure), the proportion of patients with addiction treatment program contact who met the denominator criteria was 23.6 %, which decreased by 0.1 % ($p < 0.001$) per quarter throughout the pre-implementation period (see Table 1 and Fig. 2). Just before the measure was added to performance plans, denominator prevalence was 22.1 %. When the measure was implemented, there was not a significant drop in denominator prevalence; however, the post-implementation rate of change became significantly more negative compared to the pre-implementation rate ($p < 0.001$), and was only 15.9 % at the end of the observation period. Results pertaining to the alternative version of denominator prevalence are presented in [Appendix](#) (available online).

In Fig. 3, forest plots of the 95 % CIs for immediate post-implementation *changes* in denominator prevalence across VHA facilities are shown; associated data is illustrated in Table 2. Except for a few outlying facilities, variability in immediate post-implementation *changes* in denominator prevalence was limited and not associated with other facility-level

Table 1 Segmented Linear Regression Analyses of Interrupted Time Series of Measured Performance and Denominator Prevalence

Effect	Estimate (SE)	t-value	p value
Outcome: measured performance			
Baseline	0.2077 (0.0086)	24.23	< 0.001
Pre-implementation slope	0.0019 (0.0016)	1.63	0.1112
Gap following implementation	-0.0092 (0.0093)	-0.99	0.3293
Change in slope after implementation	0.0074 (0.0012)	6.15	<0.001
Outcome: Denominator Prevalence			
Baseline	0.2364 (0.0027)	84.63	<0.001
Pre-implementation slope	-0.0013 (0.0003)	-3.50	0.001
Gap following implementation	0.0031 (0.0030)	1.01	0.319
Change in slope after implementation	-0.0009 (0.0004)	-2.51	0.0167

All models were adjusted for strong seasonal effects using a decomposition method and centered moving averages to estimate seasonal indexes

Figure 1. Effect of quality measure implementation on measured performance.

variables (Table 3). Changes in the rate of change in denominator prevalence were more varied across VHA facilities. Facilities with higher pre-implementation denominator prevalence had greater decreases in the slope of denominator prevalence over time ($p < 0.001$).

DISCUSSION

As expected, including an SUD CoC quality measure in Network Directors' performance contracts was associated with an increase in measured performance, from 23 % just before the measure was implemented to 48 % by the end of the observation period. Annually increasing the threshold of performance needed to receive the performance bonus, based on the 80th percentile of the previous year's distribution, may

have contributed to the sustained linear trend. Additionally, the overall proportion of patients with SUD program contact that qualified for the measure decreased more rapidly over time following implementation, and exhibited significant facility-level variation. An alternative definition of denominator prevalence focused on a broader group of patients with SUD, not only those in specialty programs, revealed a somewhat different pattern (Online Appendix), and highlights the importance of testing multiple specifications of denominator prevalence so that different effects might be revealed.

Given the data available, it is impossible to tell what proportion of the observed changes in denominator prevalence was due to improvements in coding or patient management versus undesirable denominator management solely intended to increase performance numbers (i.e., gaming). Measurement Reactivity Theory suggests that individuals or entities will

Figure 2. Effect of quality measure implementation on denominator prevalence (# qualified/ # receiving SUD specialty care).

Figure 3. Distribution of post-implementation “denominator reactivity” parameters across 133 VA facilities.

alter their behavior in response to the knowledge that they are being monitored, especially if they know and can influence the parameters being assessed.²⁰ Financial incentives and other contingencies such as public reporting will, in most cases, increase this reactivity. In fact, the logic of tying quality measures to incentives assumes that measurement reactivity will occur. However, the *intended* effect of measurement reactivity is on the process of care (numerator), and *not* to stimulate undesirable management of the denominator.

In the case of VA’s SUD CoC measure, facility leaders and program staff knew that they were being monitored, knew the specific parameters being assessed, had the ability to affect the processes of care being targeted by the measure, and had financial and reputational consequences tied to performance. As predicted, the intended and unintended reactivity in the measure occurred to varying degrees across the system. When changes in measured performance are driven at least in part by denominator contraction rather than numerator expansion, the validity of performance data may be threatened and negative clinical outcomes may occur.

Knowing the conditions under which unintended measurement reactivity can occur helps suggest ways that it might be minimized in the future. Clearly, it is important to design denominators that are harder to manage. In the case of the VA SUD CoC measure, which had utilization-based denominator criteria, it was entirely possible for SUD programs to pick and choose who qualified by making simple scheduling adjustments (e.g., scheduling SUD patients’ potentially qualifying third outpatient visits outside of the 30 day monitoring

window or not scheduling a third outpatient visit). Other than anecdotally, we do not know if this kind of selection occurred. Another alternative explanation might involve legitimate improvements in coding or management of patients that might have impacted denominator prevalence. Either or both of these mechanisms would explain the denominator contraction observed in this study.

Process quality measures are usually defined as the number of patients who received the treatment divided by the number of patients who were identified as needing the treatment; however, alternatives have been proposed that may be less sensitive to denominator management. Bradley et al. examined a “population-based” quality measure defined as patients both screened and treated divided by *all* patients.¹⁰ Such population-based measures might encourage both high-quality screening and treatment, and their denominators are virtually ungameable. However, this concept ignores real between-facility differences in the prevalence of screen-positive patients. Another option is to model the expected case-mix and regionally-adjusted denominator prevalence of qualifying patients using epidemiologic data.²¹ Due to complexity and probably opacity, the use of modeled denominators in quality measures would probably meet resistance among those being evaluated, but may have promise as “shadow measures” to detect potential problems with denominator prevalence.

Although almost certain to be unpopular among those being monitored or assessed, and perhaps infeasible in many contexts, another strategy is to not fully disclose the measure specifications. It might be possible to disclose only the details of a measure that pertain to its intended effects. Although unintended numerator management is also possible, not knowing exactly which patients were qualifying for measures might discourage denominator management.

Designing measures that are less easily managed is certainly important. However, given that many existing measures are vulnerable, it is equally important to design and implement “shadow measures” to help quality managers detect

Table 2. Distribution of Post-Implementation “Denominator Reactivity” Parameters Across 133 VA Facilities

Parameter	Facility-level mean (SD)	Minimum	Maximum
Gap in denominator prevalence	−0.0040 (0.0802)	−0.3873	0.5419
Change in slope denominator prevalence	−0.0054 (0.0107)	−0.0526	0.0264

Table 3. Facility-level Correlates of Gap and Change in Slope of Denominator Prevalence Following Implementation

Effect	Estimate (SE)	t-value	p value
Outcome: gap in denominator prevalence			
Intercept	0.0170 (0.0196)	0.86	0.3893
FY02-03 improvement in CoC performance	0.1832 (0.1050)	1.74	0.0835
Pre-implementation denominator prevalence	−0.2455 (0.1333)	−1.06	0.0678
SUD prevalence	0.1929 (0.3831)	0.50	0.6154
Outcome: change in slope of denominator prevalence			
Intercept	0.0027 (0.0026)	1.03	0.3070
FY02-FY03 improvement in CoC performance	−0.0045 (0.0141)	−0.32	0.7470
Pre-implementation denominator prevalence	−0.0796 (0.0179)	−4.44	< 0.001
SUD prevalence	0.0703 (0.0514)	1.36	0.1740

denominator management when it occurs. The purpose of shadow measures would be to provide warning signals to quality managers that data from the main quality measures are suspect. The simplest shadow measure could track denominator prevalence. Sizable contractions of denominator prevalence, like those revealed in Fig. 3, could trigger more in-depth investigation to determine if the changes in denominator were justified due to better coding or care management or perhaps evidence of less desirable denominator management. A more refined shadow measure would not only check for changes in and deviation from expected rates of denominator prevalence, but also the composition of the qualified patients. Specifically, if the case-mix of qualified patients changed after a measure was implemented, as has been found elsewhere,^{12,13} performance measure-motivated patient selection might be suspected. It may be possible to develop a method to penalize measured performance by the extent to which it is caused by undercounting the target population. Developing observed-expected measures of denominator prevalence might be feasible where high quality epidemiologic data exist.²¹

Several limitations of the current study are worth noting. First, the examined data are over a decade old, and the CoC measure is no longer emphasized by VA. However, the purpose of this study was to examine if changes in measured performance can be due to denominator restriction rather than the intended changes in the numerator. As such, these analyses are not meant to inform implementation and interpretation of the CoC measure, but rather to highlight the general problem of denominator management. It is also possible that providers' responses to incentives may have changed over time given the increased prevalence of incentives. Second, the reason that denominator prevalence was falling before the measure was implemented is unknown. Third, although interrupted time series is a strong quasi-experimental design,^{17,18} the lack of randomly selected or matched control series makes causal inference tenuous.

Developing quality measures that stimulate intended measurement reactivity while avoiding unintended reactivity is challenging. Previous research and the results from this study have shown that denominator management may be a real threat to quality measure validity. These findings should motivate the development of measures that are less vulnerable to

denominator management, and the exploration and evaluation of “shadow measures” to monitor and reduce undesirable denominator management.

Acknowledgements:

Contributors: There are no contributors other than the authors.

Prior presentations: Some of these results were presented at the Addiction Health Services Research Conference on October 15, 2015 and at the VA HSR&D/QUERI National Conference on 10 July 2015.

Corresponding Author: Alex H. S. Harris, PhD; Center for Innovation to Implementation VA Palo Alto Health Care System, 795 Willow Road, Menlo Park, CA 94025, USA (e-mail: Alexander.Harris2@va.gov).

Compliance with Ethical Standards: This study was funded by the VA Health Services Research and Development (HSR&D) Service (IIR 10-370-2), VA Substance Use Disorder Quality Enhancement Research Initiative (SUD QUERI; SUDQ-LIP-1208; SUDQ-LIP-1314), a VA HSR&D/QUERI Career Development Award to Dr. Hoggatt, a VA HSR&D Postdoctoral Fellowship to Dr. Vanneman, and a VA HSR&D Research Career Scientist Award (RCS-14-232) to Dr. Harris. The views expressed herein are not necessarily those of the Department of Veterans Affairs.

Conflict of Interest: All authors are employees of the US Department of Veterans Affairs. No other conflict of interest exists for any of the authors.

REFERENCES

1. Hayward RA, Kent DM. 6 EZ steps to improving your performance: (or how to make P4P pay 4U!). *JAMA*. 2008;300:255–6.
2. Wachter RM, Flanders SA, Fee C, Pronovost PJ. Public reporting of antibiotic timing in patients with pneumonia: lessons from a flawed performance measure. *Ann Intern Med*. 2008;149:29–32.
3. Baker DW, Gaseem A, American College of Physicians' Performance Measurement Committee. Evidence-based performance measures: preventing unintended consequences of quality measurement. *Ann Intern Med*. 2011;155:638–40.
4. Bratzler DW, Loeb JM, Krusenowski D. Preventing unintended consequences of quality measurement for venous thromboembolism prophylaxis. *Ann Intern Med*. 2012;156:326–7.
5. Karve AM, Ou FS, Lytle BL, Peterson ED. Potential unintended financial consequences of pay-for-performance on the quality of care for minority patients. *Am Heart J*. 2008;155:571–6.
6. Werner RM, Asch DA. The unintended consequences of publicly reporting quality information. *JAMA*. 2005;293:1239–44.
7. Riskin L, Campagna JA. Quality assessment by external bodies: intended and unintended impact on healthcare delivery. *Curr Opin Anaesthesiol*. 2009;22:237–41.
8. Saini SD, Vijan S, Schoenfeld P, Powell AA, Moser S, Kerr EA. Role of quality measurement in inappropriate use of screening for colorectal cancer: retrospective cohort study. *BMJ*. 2014;348:g1247.

9. **Roth S, Gonzales R, Harding-Anderer T, et al.** Unintended consequences of a quality measure for acute bronchitis. *Am J Manag Care*. 2012;18:e217–24.
10. **Bradley KA, Chavez LJ, Lapham GT, et al.** When quality indicators undermine quality: bias in a quality indicator of follow-up for alcohol misuse. *Psychiatr Serv*. 2013;64:1018–25.
11. **Konetzka RT, Polsky D, Werner RM.** Shipping out instead of shaping up: rehospitalization from nursing homes as an unintended effect of public reporting. *J Health Econ*. 2013;32:341–52.
12. **Dranove D, Kessler D, McClellan M, Satterthwaite M.** The effects of report cards on health care providers: is more information better? *J Polit Econ*. 2003;111:555–88.
13. **Werner RM, Asch DA, Polsky D.** Racial profiling: the unintended consequences of coronary artery bypass graft report cards. *Circulation*. 2005;111:1257–63.
14. **Harris AH, Humphreys K, Bowe T, Kivlahan DR, Finney JW.** Measuring the quality of substance use disorder treatment: evaluating the validity of the Department of Veterans Affairs continuity of care performance measure. *J Subst Abuse Treat*. 2009;36:294–305.
15. **Humphreys K, Harris AH, Kivlahan DR.** Performance monitoring of substance use disorder interventions in the Veterans health administration. *Am J Drug Alcohol Abuse*. 2009;35:123–7.
16. **Perlin JB, Kolodner RM, Roswell RH.** The Veterans health administration: quality, value, accountability, and information as transforming strategies for patient-centered care. *Am J Manag Care*. 2004;10:828–36.
17. **Duncan T, Duncan S.** A latent growth curve modeling approach to pooled interrupted time series analysis. *J Psychopathol Behav Assess*. 2004;26:271–8.
18. **Shadish R, Cook TD, Campbell DT.** Experimental and quasi-experimental designs for generalized causal inference. Boston: Houghton Mifflin Company; 2001.
19. **Fuller W.** Introduction to statistical time series. 2nd ed. New York: Wiley; 2009.
20. **Heppner PP, Wampold BE, Kivlighan DM.** Research design in counseling. 3rd ed. Belmont; Thomson; 2008.
21. **Harris AHS, Rubinsky AD, Hoggatt KJ.** Possible alternatives to diagnosis-based denominators for addiction treatment quality measures. *J Subst Abuse Treat*. 2015;58:62–66.