The American Journal of Drug and Alcohol Abuse, 35:123–127, 2009
ISSN: 0095-2990 print/ 1097-9891 online
DOI: 10.1080/00952990802707042

informa
healthcare

Performance Monitoring of Substance Use Disorder Interventions in the Veterans Health Administration

Keith Humphreys, Ph.D.
Center for Health Care Evaluation, Veterans Affairs, and Stanford University Medical Centers, Palo Alto, California, USA

Alex H. S. Harris, Ph.D.
Center for Health Care Evaluation, Veterans Affairs, and Stanford University Medical Centers, Palo Alto, California, USA

Daniel R. Kivlahan, Ph.D.
Veterans Affairs Substance Use Disorder Quality Enhancement Research Initiative and Veterans Affairs Puget Sound Health Care System and University of Washington, Seattle, Washington, USA

***Background*: Measuring and improving the quality of treatment for patients with substance use disorders are enduring challenges. *Methods*: This article describes how the Veterans Affairs health care system is using incentivized performance measures to promote more effective delivery of interventions for nicotine, illegal drug, and alcohol disorders. *Results*: The monitoring and incentive system has increased the delivery of evidence-based services, including screening for alcohol use disorders. *Conclusions*: Further work remains to be done to strengthen the connection between process-based measures and longer-term patient outcomes.**

Keywords Continuous quality improvement, substance use disorders, performance monitoring, veterans

INTRODUCTION

The Veterans Health Administration (VHA) directly provides more services targeting alcohol, illegal drug, and tobacco use than any other health care organization in the United States. Like all other areas of VHA health care services, substance use-related interventions are continuously evaluated in order to enhance their accessibility, efficiency, safety, and effectiveness (1). This article describes one of the most potent evaluation methods—incentivized performance monitoring—and explores how it can produce improvements in care quality both within VHA and in the broader U.S. health care system.

Address correspondence to Keith Humphreys, Ph.D., Center for Health Care Evaluation, VAPAHCS (152), 795 Willow Road, Menlo Park, CA 94025, USA. E-mail: KNH@stanford.edu

We begin with essential background on veterans and VHA. The veteran population currently stands at around 24 million. The number of surviving World War II and Korean War veterans is rapidly dwindling, whereas new veterans are being created daily by the current wars in Iraq and Afghanistan. The net effect of these two forces is to increase the proportion of young people and women in the veteran population. In 1996, Congress expanded eligibility for VHA services to all veterans. This was subsequently somewhat restricted in 2003, but the vast majority of veterans are still eligible for VHA services, and about 5 million access the system each year (2).

National survey data gathered from 2000–2003 (i.e., before there were many veterans from the current wars) showed high rates of past month substance use among veterans, including binge consumption of five or more alcoholic drinks on the same occasion (22.6%), having at least five such binges per month (7.5%), using marijuana (3.5%), using other illegal drugs (1.7%), and smoking cigarettes every day (18.8%). The survey study also showed that use of most addictive substances was somewhat higher among veterans than among nonveterans of similar age, sex, and geographic region (3). More recent data from veterans of the current wars in Iraq and Afghanistan reveals that substance use disorders are among the most prevalent mental health diagnoses among VHA users (4). Ensuring high quality services for substance use disorder is thus a higher priority than ever.

To address this substantial need, VHA provides a wide array of services for substance use disorders. In fiscal year 2006, the system comprised 226 specialty alcohol and drug treatment programs, which treated over 120,000 patients (5, 6). These specialty programs are complemented by substance use–focused services provided within mental health programs (e.g.,

case management for veterans with co-occurring disorders), behavioral medicine clinics (e.g., smoking cessation counseling and nicotine replacement therapy) and primary care settings (e.g., screening for alcohol misuse and brief interventions for problem drinking).

THE NATURE AND STRUCTURE OF THE VA PERFORMANCE MONITORING SYSTEM

Among the key strategies in the successful transformation of care in VHA was the investment in a system of performance measurement across several domains including quality of care, access to services, and patient satisfaction (7, 8). Both financial and nonfinancial incentives promoted improved performance on these measures. Regional directors, each of whom oversees the facilities in a particular geographic area, have a performance contract with salary bonuses of roughly 10% contingent on meeting specified quality standards (1).Regional directors establish corresponding financial performance incentives on these standards for managers and clinicians at facilities in their region. Facility-level results are posted on a VHA intranet site and widely disseminated. This data availability facilitate ongoing monitoring and generate nonfinancial incentives in the form of competition within and between networks (2). The centralized infrastructure for the system is lead by a national Office of Quality and Performance that works through quality managers at the regional and facility level. In a cooperative effort between clinicians and health services researchers, disease-specific quality enhancement research initiatives, including one for substance use disorders, focus on identifying and addressing barriers to improved quality (9).

Selection of performance measures has been largely driven by the recommendations included in evidence-based clinical practice guidelines developed by VHA in collaboration with the Department of Defense (available on line at www.oqp.med.va.gov). Based on these guidelines and a systematic evaluation process, a work group of managers, clinicians, researchers, and policy makers operationalizes potential performance measures and recommends them to the Undersecretary for Health for national adoption (10). Those selected are determined to have the most compelling evidence base, highest potential to reduce disease burden in the VHA population, feasibility of reliable measurement, and a gap between current and optimal performance. VHA also takes account of what important bodies in the performance management field, such as the National Quality Forum and the Washington Circle, are recommending as this often yields opportunities to benchmark VHA care against that in other healthcare systems.

Typically, at least a year of pilot data is required before a measure is included in the incentivized performance system. Other measures are removed from the incentivized performance system when there is no further room for improvement, though they are still monitored and can be restored for a specific region if performance deteriorates.

Current performance measures relevant to substance use disorders (see Table 1) monitor population-based screening for alcohol misuse with the first 3 items of the Alcohol Use Disorders Identification Test (AUDIT-C) (11), appropriate clinical follow-up for those who screen positive for alcohol misuse, evidence-based approaches to smoking cessation (12), and treatment retention over 90 days for patients beginning new episodes of substance use disorder specialty care (13).

THE LOGIC BEHIND THE VHA PERFORMANCE MONITORING SYSTEM

Why does VHA, or for that matter any health care system, need a performance monitoring system at all? Consider two related questions: (1) Why are there no bad restaurants in San Francisco? and (2) Why do prison cafeterias serve lousy food? The answer to the first question is that San Francisco has a large number of people who eat out frequently and have excellent

TABLE 1
Substance use disorder (SUD) related performance measures in VHA

	FY08 Targets
Alcohol Misuse Screening and Brief Intervention	
Percent of all outpatients screened annually for alcohol misuse using the AUDIT-C with documented item responses and total score	93%
Percent of patients who screen positive on the AUDIT-C (total score >= 5) who have documented brief alcohol counseling	N/A
Tobacco Use Cessation	
Percent of tobacco users with documentation of tobacco use cessation counseling in the past 12 months	84%
Percent of tobacco users offered medication to support smoking cessation within the past 12 months	73%
Percent of tobacco users offered referral to a tobacco use cessation program within the past 12 months	93%
Continuity of Specialty care for SUD	
Percentage of patients beginning new episodes of SUD specialty care with at least two visits each 30 days for 90 days	47%

judgment about the quality of food. Further, when they sign the bill, like all restaurant goers, San Francisco diners can assess if the quality they received justifies the cost. Through normal market forces, this results in poor restaurants (i.e., those who do not deliver quality appropriate to the price they charge) going out of business and good restaurants staying in existence. In short, because consumers can detect quality and cost, it makes sense for restaurants to invest in producing the best possible food at the most attractive price.

If serving good food is rewarded, why do prisons, even prisons near San Francisco, serve lousy food? Even prisoners of the most refined taste must eat whatever the cafeteria serves. Lack of investments in quality food thus cannot be punished in the market place by prisoners because they don't have the choice to eat somewhere else.

The above examples are relevant to substance use disorder treatment in different respects. First, unlike in the San Francisco restaurant example, consumers have little idea how to tell whether the substance use disorder treatment they are receiving is good or bad and are usually shielded from knowing the costs of care. Second, as in the prison example, Substance Use Disorder (SUD) treatment is usually sought in the public sector, in which consumers have little choice. SUD treatment thus typically lacks the required elements for a market to improve quality, namely, the ability of consumers to detect quality and costs and then to reward it by choosing higher value services over lower value services.

In short, without a performance measurement system that detects and rewards quality, public sector substance use disorder services have no internal incentive to improve quality. Indeed, they actually have a disincentive, in that investments in quality consume resources without creating any competitive advantage in the marketplace. By creating a system that recognizes quality and rewards it financially, the VHA performance monitoring system overcomes these structural problems.

HAS PERFORMANCE MONITORING INFLUENCED THE RANGE AND QUALITY OF ADDICTION-RELATED SERVICES?

The effort to improve quality of care through performance measurement has been an iterative process as shown by the successful VHA experience with alcohol screening. Given the high prevalence of alcohol use disorders within VHA and evidence that it often went unrecognized or unaddressed by primary care providers, one of the earliest performance measures required annual alcohol screening with any validated questionnaire. Baseline performance was only 2% in 1996. By the end of 1997, when the performance measure was first required, 40% of patients were screened, and this increased to 85% by 2000; however, most sites were screening only with the CAGE questionnaire (11). Although this screening resulted in some increased identification of patients with more severe problems, only a minority of patients with the full spectrum of alcohol misuse (i.e., including drinking beyond recommended levels) reported receiving any alcohol-related advice. The VHA Office of Quality and Performance became especially concerned by survey data in the late 1990s indicating that, among patients who reported a need for help with their drinking, fewer than 1 in 6 heavy drinkers or very heavy drinkers reported receiving it.

Following validation of the AUDIT-C as a brief screen for the full spectrum of alcohol misuse (15), VHA changed the performance measure to require the AUDIT-C (or full AUDIT, which comprises it) for alcohol misuse screening. In late 2003, a computerized clinical reminder was disseminated widely to facilitate scoring and documentation with the AUDIT-C. Revised targets for screening rates were set at 82% (successful) and 89% (exceptional) with patients considered adequately screened if there was 1) documentation of AUDIT-C screening, 2) consistent chart documentation of no alcohol use, or 3) evidence of recent substance use disorder treatment involvement. Based on standardized medical record reviews using the revised performance measure, all 21 regional networks exceeded the exceptional standard in 2005 with a national average of 93% (11).

Subsequent analyses showed inconsistencies between screening results as documented in the medical record and responses to mailed patient surveys after outpatient visits (e.g., the medical record documented no drinking but the patient self-reported drinking on the mailed questionnaire) (16). These results prompted refinements of the performance measure that now requires documentation of AUDIT-C item responses and total score (to facilitate risk stratification) and that all patients be asked at least the first question of the AUDIT-C to establish non-drinker vs. drinker status. Along with screening for depression, Post-Traumatic Stress Disorder, and suicide risk, alcohol misuse screening with the AUDIT-C is a major component of the incentivized performance system for FY08. Notably, FY08 is the first time there has been a performance measure for appropriate follow-up of alcohol misuse screening results. This measure is now monitored as one of the selected "Transformational Measures," and this will likely follow a similar incremental course in improving performance and refining the metric over time.

LINKING PROCESS MEASURES TO DISTAL OUTCOMES

Direct monitoring of post-treatment outcomes by clinicians was once attempted in VHA. But because the time to track and re-interview former patients competed with clinician's efforts to treat current cases, follow-up rates never approached 50% at any facility and were more commonly around 20% (17, 18). Further, the validity and reliability of measurement across sites was often in question, and the program was widely resented as a drain on clinicians' time. A subsequent centralized post-treatment outcome monitoring system done by a core evaluation team resulted in substantially better, if not outstanding, follow-up rates (e.g., 65%) and also established consistent reliability and validity of measurement by using a small team of well-trained interviewers (19). Yet this approach would be very costly to implement for all patients, and further did not solve the basic

dilemma that post-treatment outcome assessments are not of help to clinicians during the process of treatment (20).

For these reasons, it is highly desirable to rely on more easily gathered, rapidly reportable process measures such as those presented in Table 1. Such measures are typically designed by expert panels, who operationalize the processes of care shown in the treatment research literature to predict better patient outcomes (21).The often unstated assumption is that performance measures constructed in this way will serve at least two functions: a) to incentivize care that produces better clinical outcomes for individual patients and b) to distinguish facility-level quality, that is, to identify high and low performing facilities as gauged by average outcomes.

Because process measures usually have high face validity, the extent to which they serve these functions is often left unchecked. However, the path from nuanced research studies or clinical practice guidelines to a performance measure that serves its intended functions is fraught with hazard (22). For logistic and cost reasons, designers of performance measures often limit themselves to data elements available in administrative databases (e.g., utilization records, diagnostic information, and procedure codes). Just because treatment research literature unequivocally supports a specific process of care need not imply that the process of care can be easily measured using currently available administrative data (23).

Relatedly, the untested assumption is often made that individual patient-level associations between processes of care and outcomes will generalize to the facility-level. Once aggregated to the facility-level, the meaning of the performance measure changes from the effect of individual patients meeting the criterion to the effect of being at a facility with a specific rate of meeting the criterion. Although a treatment retention performance measure, for example, may predict better outcomes for individual patients, being at facilities where patients tend to engage in treatment longer may be packaged with other factors that may cause average patient outcomes to be worse (e.g., higher baseline severity) or better (e.g., excellent follow-up case management) than predicted with individual level data. Perhaps surprisingly, performance measures that tap processes of care that are reasonable approximations to those indicated in the treatment literature may predict neither patient nor facility-level outcomes. This problem is not unique to substance use disorder treatment, but has been recognized in multiple medical specialties (24). Therefore it is critical to conduct post-construction validation of performance measures before implementing them, a step that is rarely taken.

Our own efforts to validate the association between the VHA Continuity of Care (CoC) performance measure and patient- and facility-level outcomes did not yield encouraging results (13).The CoC measure was quite poor at predicting individual patient outcomes, and was actually negatively correlated with facility-level outcomes. Although post-hoc analysis identified ways to improve the predictive validity of the measure, these findings were, to put it mildly, rather humbling.

The implications of these analyses are clear. Although VHA has been able to show substantial increases in continuity of care since the inception of the performance measure in FY04 (27% to 44%), the reasonable-seeming and face-valid CoC performance measure is not closely linked with outcomes at follow-up. Although any one study has limitations and can be criticized, no other studies directly evaluating the CoC measure exist, especially at the facility-level. The use of quality measures with unknown or poor validity has many serious risks, including incentivizing poor care, deemphasizing issues not covered by the measure, or directing clinical energy and attention to a treatment process that is not part of the causal chain by which the treatment produces its intended effects. Performance measures are interventions targeted at the system and facility, and they should be held to the same evidence-based standards as clinical interventions at the pateint level. That is, they should be directly evaluated for their intended purpose before wide-spread implementation.

DIRECTIONS FOR THE FUTURE

Currently, performance measures are designed by expert panels that imperfectly operationalize nuanced patient-level evidence into measures intended to serve several, sometime unspecified, goals. This design process could be improved by better specification of the purpose of the measures (e.g., to guide the care of individual patients or assess the quality of facilities?), and greater use of existing data on which patterns of care best predict outcome. Evaluators could then conduct prospective analyses under real-world care conditions to determine whether changes in performance predict changes in outcomes.

Other strategies might also improve process performance measures (e.g., direct survey of patients about processes of care) in VHA and outside of it as well. Although performance measures are usually designed to be as simple and universal as possible, it may be useful to examine whether certain patient characteristics (e.g., treatment history, co-morbid diagnoses, demographics, treatment preferences) and setting factors (inpatient vs. outpatient delivery) can be used to refine the specifications such that power to predict outcome is maximized. Other possible refinements involve the use of other clinically relevant data from the medical record, such as patient or clinician assessments of need for treatment or documentation of informal care (e.g., 12-step group involvement). Such information could be used to better define the denominator of a retention-based performance measure so that only patients who need and stand to benefit from more treatment are encouraged to stay involved.

When validity and feasibility conflict in the design of performance measures, the latter usually carries the day. Gathering detailed process or outcome data on tens of thousand of patients is simply impossible without substantial investment (23).However, sampling strategies could reduce the burden of this task in those cases where the goals of a performance measure are to distinguish and improve quality at the facility level.

In addition to refining current process measures of care, VHA is also experimenting with real-time, brief assessments of patient progress during treatment. If feasible and clinically acceptable, concurrent treatment monitoring would provide a good guide to quality as well as being of use in clinical care, without the enormous cost associated with locating and re-interviewing former patients.

We hope this article provides managers, clinicians, and purchasers of substance use disorder services with useful information about the VHA performance management system. As indicated, such systems can increase the frequency with which clinicians apply evidence-based practices. However, the addiction field clearly needs more research to guide the selection of such measures so that they are validated against the criteria of interest whether that involves some evidence-based standard of quality or prediction of distal outcome. That, in our view, is the chief challenge of performance measurement for substance use disorder services, not just in the VHA but in the U.S. health care system as a whole.

ACKNOWLEDGMENTS

This study was supported by grants from the Department of Veterans Affairs Health Services Research and Development Service and Veterans Affairs Quality Enhancement Research Initiative. Any opinions expressed are those of the authors and do not necessarily reflect VA policy positions.

Declaration of Interest

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the article.

REFERENCES

1. Kerr EA, Fleming B. Making performance indicators work: Experiences of US Veterans Health Administration. British Medical Journal 2007; 335:971–973.
2. Oliver A. The Veterans Health Administration: An American success story? Milbank Quarterly 2007; 851:5–35.
3. Wagner TH, Harris KM, Federman B, Dai L, Luna Y, Humphreys K. Prevalence of drug, alcohol and cigarette use among veterans and comparable non-veterans from the National Survey on Drug Use and Health. Psychological Services 2007; 4:149–157.
4. Seal KH, Bertenthal D, Miner CR, Sen S, Marmar C. Bringing the war back home: mental health disorders among 103,788 US veterans returning from Iraq and Afghanistan seen at Department of Veterans Affairs facilities. Archives of Internal Medicine 2007; 167:476–482.
5. Tracy S, Morales A, Trafton J. The Department of Veterans Affairs Substance Abuse Treatment System: Results of the 2006 Drug and Alcohol Program Survey. Palo Alto, CA: Program Evaluation and Resource Center, 2007.
6. Dalton A, McKellar JD. Health Services for VA Substance Use Disorder Patients: Comparison of Utilization in Fiscal Years 2006, 2005, and 2002. Palo Alto, CA: Program Evaluation and Resource Center, 2007.
7. Perlin JB, Kolodner RM, Roswell RH. The Veterans Health Administration: Quality, value, accountability, and information as transforming strategies for patient-centered care. American Journal of Managed Care 2004; 11: 828–36.
8. Rosenheck R. Mental and substance use health services for veterans: Experience with performance evaluation in the Department of Veterans Affairs. In: Improving the Quality of Health Care for Mental and Substance-Use Conditions: Quality Chasm Series. Washington, DC: Institute of Medicine, 2004.
9. Demakis JG, McQueen L, Kizer KW, Feussner JR. Quality Enhancement Research Initiative QUERI: A collaboration between research and clinical practice. Medical Care 2002; 386(Suppl 1):I17–25.
10. Francis J, Perlin JB. Improving performance through knowledge translation in the Veterans Health Administration. The Journal of Continuing Education in the Health Professions 2006; 26:63–71.
11. Bradley KA, Williams EC, Achtmeyer CE, Collins B, Volpp B, Kivlahan DR. Implementation of evidence-based alcohol screening and counseling in the Veterans Affairs Healthcare System. American Journal of Managed Care 2006; 12:597–606.
12. Sherman SE. A Framework for Tobacco Control: Lessons learned from Veterans Health Administration. British Medical Journal 2008; 336:1016–1019.
13. Harris AS, Humphreys K, Bowe T, Kivlahan D, Finney J.W. Measuring the quality of substance use disorder treatment: Assessing the VA continuity of care performance measure. Journal of Substance Abuse Treatment 2009, 36:294–305.
14. Barry KL, Blow FC, Willenbring ML, McCormick R, Brockmann LM, Visnic S. Use of alcohol screening and brief interventions in primary care settings: Implementation and barriers. Substance Abuse 2004; 25:27–36.
15. Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA. The AUDIT alcohol consumption questions AUDIT-C: An effective brief screening test for problem drinking: Ambulatory Care Quality Improvement Project ACQUIP: Alcohol Use Disorders Identification Test. Archives of Internal Medicine 1998; 158:1789–1795.
16. Hawkins E, Kivlahan DR, Williams EC, Wright SM, Craig T, Bradley KA. Examining quality issues in alcohol misuse screening. Substance Abuse 2007; 28:53–65.
17. Moos RH, Finney JW, Suchinsky R. Outcomes Monitoring for Patients with Substance Use Disorders: IV. Cohort 2 Patients' 6–12 Month Treatment Outcomes. Palo Alto, CA: Program Evaluation and Resource Center, 2000.
18. Otilingam PG, Ritsher JB, Finney JW, Moos RH, Suchinsky R. Outcomes Monitoring for Patients with Substance Use Disorders: V. Cohort 3 Patients' Characteristics, Treatment, and Treatment Outcomes. Palo Alto, CA: Program Evaluation and Resource Center, 2002.
19. Tiet QQ, Byrnes HF, Barnett P, Finney JW. A practical system for monitoring the outcomes of substance use disorder patients. Journal of Substance Abuse Treatment 2006; 30:337–347.
20. McLellan T, McKay JR, Forman R, Cacciola J, Kemp J. Reconsidering the evaluation of addiction treatment: from retrospective follow-up to concurrent recovery monitoring. Addiction 2005; 100:447–458.
21. Lilford RJ, Brown CA, Nicholl J. Use of process measures to monitor the quality of clinical practice. British Medical Journal 2007; 335:648–650.
22. Walter LC, Davidowitz NP, Heineken PA, Covinsky KE. Pitfalls of converting practice guidelines into quality measures: Lessons learned from a VA performance measure. Journal of the American Medical Association 2004; 291:2466–2470.
23. Hayward RA. Performance measurement in search of a path. New England Journal of Medicine 2007; 356:951–953.
24. Werner RM, Bradlow ET. Relationship between Medicare's hospital compare performance measure and mortality rates. Journal of the American Medical Association 2006; 296:2694–2702.