

Opportunities to Improve Measurement-Based Care Practices in Mental Health Care Systems: A Case Example of Electronic Mental Health Screening and Measurement

Paul R. King, PhD,
and Gregory P. Beehler, PhD, MA
Veterans Affairs Center for Integrated Healthcare,
Buffalo, New York, and University at Buffalo, The
State University of New York

Michael Wade, MS
Veterans Affairs Center for Integrated Healthcare,
Syracuse, New York

Laura J. Buchholz, PhD
Veterans Affairs Center for Integrated Healthcare,
Buffalo, New York, and University of Tampa

Jennifer S. Funderburk, PhD
Veterans Affairs Center for Integrated Healthcare,
Syracuse, New York, and Syracuse University

Kaitlin R. Lilienthal, PhD
Veterans Affairs Center for Integrated Healthcare,
Syracuse, New York, and Cornell University

Christina L. Vair, PhD
Veterans Affairs Center for Integrated Healthcare,
Buffalo, New York, and W.G. (Bill) Hefner
Veterans Affairs Medical Center, Salisbury,
North Carolina

Introduction: Measurement-based care (MBC) involves the systematic collection of data to inform clinical decision-making and monitor treatment outcomes. In addition to benefitting patients and providers, data on MBC implementation can also be used to inform quality improvement efforts within existing health care systems. **Method:** The

Paul R. King, PhD, Veterans Affairs Center for Integrated Healthcare, Buffalo, New York, and Department of Counseling, School, and Educational Psychology, University at Buffalo, The State University of New York; Gregory P. Beehler, PhD, MA, Veterans Affairs Center for Integrated Healthcare, Buffalo, and School of Public Health and Health Professions, University at Buffalo, The State University of New York; Michael Wade, MS, Veterans Affairs Center for Integrated Healthcare, Syracuse, New York; Laura J. Buchholz, PhD, Veterans Affairs Center for Integrated Healthcare, and Department of Psychology, University of Tampa; Jennifer S. Funderburk, PhD, Veterans Affairs Center for Integrated Healthcare, and Department of Psychology, Syracuse University; Kaitlin R. Lilienthal, PhD, Veterans Affairs Center for Integrated Healthcare, and Gannett Health Services, Cornell University; Christina L. Vair, PhD, Veterans Affairs Center for Integrated Healthcare, and W. G. (Bill) Hefner Veterans Affairs Medical Center, Salisbury, North Carolina.

This study was supported by funding from the Veterans Affairs (VA) Mental Health Quality Enhancement Research Initiative (QLP 55-035), as well as with resources and use of facilities at the VA Center for Integrated Healthcare and the VA Western New York Healthcare System. Manuscript preparation was supported in part by the VA

Office of Academic Affiliations, Advanced Fellowship Program in Mental Illness Research and Treatment. The views expressed in this article are those of the authors and do not represent the position or policy of the Department of Veterans Affairs or the United States Government. There are no conflicts of interest to report. Laura J. Buchholz's work on this project began during her postdoctoral fellowship at the VA Center for Integrated Healthcare in Buffalo, New York, and continued through her appointment at the University of Tampa. Kaitlin R. Lilienthal's work on this project began during her postdoctoral fellowship at the VA Center for Integrated Healthcare in Syracuse, New York, and continued through her appointment at Cornell University's Gannett Health Services. Christina L. Vair's work on this project began during her tenure at the VA Center for Integrated Healthcare in Buffalo, and continued through her appointment at the W. G. (Bill) Hefner VA Medical Center.

Correspondence concerning this article should be addressed to Paul R. King, PhD, Veterans Affairs Center for Integrated Healthcare, Veterans Affairs Western New York Healthcare System, 3495 Bailey Avenue (116N), Buffalo, NY 14215. E-mail: paul.king2@va.gov

method was retrospective chart review. We collected data on electronic mental health (MH) screens and symptom measures recorded by MH providers. Patients were 28,376 veterans who received MH services in a northeastern region. **Results:** Although rates varied by MH condition and clinic type, screening for alcohol misuse, depression, and posttraumatic stress disorder appeared to occur with regularity. MH symptom measurement was less frequent than screening but included measures of alcohol and substance use, posttraumatic stress disorder, depression, and suicidal ideation. Patient demographics (e.g., age, military service era, sex, MH diagnosis) and frequency of clinic contact emerged as significant predictors of symptom measurement. **Discussion:** In this article, we illustrate how data on MH screening and measurement can be organized, analyzed, and interpreted to identify opportunities to enhance MBC practices in MH care. We conclude with a discussion of how large data set analyses can contribute to programmatic MBC initiatives.

Keywords: integrated care, measurement-based care, mental health

Supplemental materials: <http://dx.doi.org/10.1037/fsh0000379.supp>

Measurement-based care (MBC) refers to the use of screening tools and outcome measures to identify the presence of prevalent conditions, monitor changes in symptoms over time, and guide treatment planning (Scott & Lewis, 2015). Incorporating MBC in mental health (MH) treatment facilitates patient engagement (Eisen, Dickey, & Sederer, 2000), quantification of symptom severity (Dowrick et al., 2009), and symptom improvement (Bickman, Kelley, Breda, de Andrade, & Riemer, 2011). MBC can also aid the diagnostic process (Sapyta, Riemer, & Bickman, 2005), support identification of suitable interventions (Scott & Lewis, 2015), and improve interprofessional communication and collaboration, in turn positively influencing patient outcomes (Katon et al., 2010).

Despite the utility and benefit of MBC, evidence suggests it is often underused by MH providers (Gilbody, House, & Sheldon, 2002). Potential contributors may include practical concerns (e.g., time constraints), a lack of organizational support (e.g., financial barriers), and perceived patient complexity (Scott & Lewis, 2015). Low use of measurement has potential negative implications for patients, providers, and administrators (Institute of Medicine, 2001), including the potential for disparities in care among individuals who do not receive measurement, decreased service quality, delayed treatment decisions, and reduced efficiency in clinic operations. Further, a lack of patient-level measurement data can hinder program administrators' efforts in evaluating the effectiveness of clinic processes and outcomes, including identification of service gaps.

The use of health information technology, specifically software embedded in electronic medical records (EMRs), is one potential solution to address the low rates of MBC reported in MH (Beehler et al., 2017; Haberer, Trabin, & Klinkman, 2013; Scott & Lewis, 2015), as well as to evaluate planned changes in clinical operations in MH care systems. Electronic administration of screening tools and symptom measures offers several practical advantages, which may be imperative for busy providers who want to improve their clinical practice. For example, the use of MH software programs can facilitate self-administration by patients, as well as expedite administration, scoring, and interpretation of results by providers. Additionally, electronic chart documentation has the potential to facilitate communication among providers and increase contact between members of interdisciplinary care teams to coordinate patient care. At an administrative level, EMR data may be exported into analyzable data sets to investigate patterns of program outcomes, determine where quality improvement (QI) efforts are needed, and evaluate progress on existing QI initiatives to promote evidence-based practices.

The Veterans Health Administration (VHA) is one example of an integrated health care system with an advanced EMR capable of supporting data-driven QI. In this article, we describe the current state of MBC across diverse settings in one VHA region and discuss what analysis of large clinical and administrative data sets can contribute to programmatic MBC initiatives. Specifically, we use EMR data to illus-

trate how information on MH instrumentation can be organized, analyzed, and interpreted to identify potential targets for future QI initiatives.

Method

Sample and Procedure

We retrieved VHA EMR data for all veterans with at least one MH appointment in 2014 in a northeastern region. Data included all MH screening tools and symptom measures recorded in Mental Health Assistant, a MBC software package embedded in the EMR. The study team included six psychologists and one biostatistician. The Institutional Review Board at the Veterans Affairs Western New York Healthcare System approved all study activities.

Study Variables

Patient characteristics. Demographics included age, gender, marital status, service era, service-connected disability status,¹ and race and ethnicity (as available). Encounter data included clinic stop codes, appointment dates, and International Classification of Diseases-9 diagnostic codes recorded for each visit. We defined MH encounters per VHA 2014 stop code parameters (see the [online supplemental materials](#)), and organized each into one of four core outpatient clinic classifications: primary care–MH integration (PC-MHI), specialty MH, substance use disorder (SUD), and posttraumatic stress disorder (PTSD). Excluding assessment-based clinics (e.g., neuropsychology), which were not a part of this investigation, these settings accounted for nearly all (98%) outpatient encounters.

Diagnostic information. We aggregated encounter diagnostic codes into *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (American Psychiatric Association, 2013) taxonomic groups, including alcohol/SUDs, anxiety disorders (e.g., generalized anxiety, phobias), mood disorders (e.g., depression, bipolar disorder), personality disorders, PTSD and other trauma/stressor-related disorders (e.g., adjustment disorders), neurocognitive disorders, and schizophrenia/other psychotic disorders. We also compiled a final composite category of “other” MH disorders, which included

low base-rate conditions recorded in this sample (e.g., attention deficit/hyperactivity).

MH instruments. We retrieved names and dates of administration for validated MH screening tools and symptom measures from the EMR. Our organizational structure included five screening and nine symptom measurement domains, which mapped on to the diagnostic categories outlined here.

MH screening tools. For the purposes of this study, we defined a MH screening tool as brief instruments designed to identify risk for a MH condition. For instance, we identified the two-item Patient Health Questionnaire (Kroenke, Spitzer, & Williams, 2003) as a depression screening tool. Screening domains included alcohol/SUDs, depression/suicidal ideation, cognitive function, PTSD, and a miscellaneous category.

MH symptom measures. We defined MH symptom measures as standardized instruments designed to aid initial assessment and/or routine outcome monitoring. Specific domains included alcohol/SUDs, anxiety, depression/suicidal ideation, personality, and PTSD symptoms. General domains included health/quality of life/disability measures (e.g., Short Form-36 Health Survey; Ware & Sherbourne, 1992), cross-cutting/transdiagnostic measures (e.g., Brief Symptom Inventory; Derogatis, 1993), clinician-rated scales (e.g., Abnormal Involuntary Movement Scale; Guy, 1976), and one miscellaneous category that did not fit the aforementioned taxonomies (e.g., Zarit Burden Interview; Zarit, Reever, & Bach-Peterson, 1980).

Analysis

Our primary aim was to provide an example of how information on MH instrumentation can be organized, analyzed, and interpreted in support of programmatic initiatives to improve MBC adoption across outpatient MH settings. To this end, we used descriptive statistics to summarize demographic variables, encounter characteristics, and frequencies of screening and measurement. We calculated chi-square sta-

¹ Service-connected disabilities are objectively linked to military service, either in origin or exacerbation of an underlying condition (Veterans Benefits Administration, 2013). The overall percentage of disability is based on an algorithm that estimates functional impairment; higher percentages (maximum = 100%) indicate a greater degree of global impairment.

tistics and risk ratios (*RRs*) to explore the relative likelihood of screening and measurement in any clinic setting. Incorporating patient-level predictors of measurement (i.e., gender, diagnosis, length of care episode) identified in prior research (Beehler, King, Vair, Gass, & Funderburk, 2016), we then computed separate logistic regression models using penalized maximum likelihood estimation (Allison, 2008; Firth, 1993; Heinze & Schemper, 2002). Finally, we generated corresponding odds ratios (*ORs*) and confidence intervals to estimate the likelihood of electronic MH measurement of patient-reported outcomes based on clinic and patient-level predictors. All analyses were performed with SAS 9.4, with $\alpha = .05$.

Results

Sample Characteristics

Patients ($N = 28,376$) were primarily male (89.8%), married (45.6%), and not of Hispanic/Latino origin (94.2%). Most were Persian Gulf (38.5%) or Vietnam-era veterans (37.9%), with a mean age of 53.5 years ($SD = 15.8$). Data on race were not consistently recorded (71.1% missing).² The majority of patients (68.4%) evidenced service-connected disabilities, with overall disability ratings often exceeding 70% (60.3%), suggesting a relatively high degree of functional impairment. Trauma/stressor-related disorders, including PTSD, were the most common diagnostic classification (15.1%), followed by mood disorders (12.9%). Whereas more than half of patients had comorbid MH conditions (52.7%), only 1.8% had no MH diagnosis listed. The most common comorbidities were trauma/stressor-related disorders and mood disorders (25.3%), anxiety and mood disorders (13.7%), and mood and alcohol/SUDs (12.9%). A total of 237,326 MH appointments were recorded and were predominantly individual (79.0%) versus group (21.0%) encounters. In addition, participants were more likely to be seen in specialty MH (49.8%) versus PC-MHI (18.0%) or other settings (15.8% SUD, and 14.3% PTSD; see the online supplemental materials).

Prevalence of MH Screening

Table 1 summarizes MH screens administered during the study period. Of the 340,626

screens recorded, most were for alcohol/SUDs (59.8%), with 80.9% of patients evidencing at least one completed screen. Depression/suicidal ideation comprised the next most frequent category (23.4%), with 42.3% of patients receiving at least one screen. PTSD screening tools followed (16.2% of all administrations), documented in 25.4% of patients. Cognitive screens were among the rarest administrations (0.5% of all screens, and 0.9% of all patients). All screens were more likely to be administered in PC-MHI versus other settings ($RR = 1.11$ – 1.36 , $p < .001$ – $.02$), though rates of cognitive screens remained low even after restricting comparisons to patients aged 65 or older (range = 2.0–4.6% of patients per setting).

Prevalence of MH Symptom Measurement

As noted in Table 2, clinician-rated scales (e.g., the Abnormal Involuntary Movement Scale) were most frequently recorded (27.5%). Measures of PTSD (23.6%), alcohol/SUDs (23.5%), and depression/suicidal ideation (19.6%) were the next most common, though the relative proportion of patients with a completed measure was low (7.2%, 4.3%, and 7.9%, respectively). Personality (2.7%), anxiety (2.4%), and other measures were the least frequently documented (0.0%–0.7% of all administrations); 1–1.4% of patients (respectively) had a completed personality or anxiety measure recorded. Rates of MH symptom measures varied by setting, though rates were generally low overall. Measurement of alcohol/SUD symptoms was more common in SUD-specific settings than other settings combined (32.3% vs. 0.9%; $RR = 36.08$), $\chi^2(1) = 6573.88$, $p < .0001$. Similarly, measurement of PTSD symptoms was more common in PTSD-specific treatment settings than others (24.7% vs. 4.9%; $RR = 5.03$), $\chi^2(1) = 1685.26$, $p < .0001$.

Predictors of MH Measurement

Several patient-level factors predicted the likelihood of symptom measurement (see Table

² High rates of missing race/ethnicity data have been demonstrated in other VHA research studies (e.g., Long, Bamba, Ling, & Shea, 2006), as well as within other health-care systems (e.g., Williams, Walker, & Egede, 2016). Research is needed to examine the effects of these factors on rates of measurement.

Table 1
Electronic Mental Health Screening by Category and Clinic Type

Screen category and name (<i>N</i> = 340,626 administrations) ^a	<i>n</i>	%		
Alcohol/SUD				
Alcohol Use Disorders Identification Test—Consumption	203,801	59.8		
Cut Down-Annoyed-Guilty-Eye Opener Screen	10	.0		
Depression				
Two-Item Patient Health Questionnaire	79,778	23.4		
Cognitive				
Blessed Orientation Memory Concentration Test	1,407	.4		
Mini-Cog	12	.0		
Montreal Cognitive Assessment	14	.0		
St Louis University Mental Status Exam	304	.1		
PTSD				
Primary Care PTSD Screen	55,297	16.2		
Screening rate by clinic type (<i>N</i> = 28,376 patients) ^{b,c}				
Screen type	PC-MHI (<i>n</i> = 9,720)	Specialty-MH (<i>n</i> = 21,135)	PTSD (<i>n</i> = 3,236)	SUD (<i>n</i> = 3,046)
Alcohol/SUD	86.4%	80.4%	84.6%	81.6%
Depression	48.1%	38.3%	41.1%	43.5%
Cognitive				
All ages	1.1%	.8%	.8%	.7%
Age ≥65	3.3% ^d	2.8% ^e	2.0% ^f	4.6% ^g
PTSD	30.8%	24.1%	16.8%	36.1%

Note. PC-MHI = primary care–mental health integration; PTSD = posttraumatic stress disorder; Specialty-MH = specialty mental health; SUD = substance use disorder.

^a The miscellaneous screen category (*N* = 3; e.g., Zarit Burden Interview) was excluded from the table due to overall small cell sizes. ^b 0% rate values were observed in each clinic category. ^c Row sums exceed the total *N*, as patients may have had clinical encounters in more than one clinic. ^d *n* = 2,764. ^e *n* = 5,720. ^f *n* = 1,116. ^g *n* = 326.

3). Older age was generally associated with a stepwise decrease in the likelihood of measurement of anxiety, depression, personality traits, and PTSD. Veterans aged 65 or older were typically least likely to have documented MH measurement (*ORs* = .19–.74). The only consistent effects of service era were found in depression/suicidal ideation and PTSD, with Persian Gulf veterans more likely to evidence measurement than other cohorts. Individuals with service-connected disabilities were significantly less likely to evidence measurement of alcohol/SUDs (*OR* = .64), depression/suicidal ideation (*OR* = .72), and PTSD (*OR* = .86) than those without documented disabilities. Further, female veterans were uniformly less likely than males to evidence measurement in all categories (*ORs* = .43–.83). The number of MH encounters recorded per patient impacted the likelihood of MH measurement, with a greater number of encounters generally associated with a greater likelihood of measurement. Aside from PTSD symptom measurement, the pres-

ence of multiple MH diagnoses was also generally associated with an increased likelihood of symptom measurement. Three or more recorded MH diagnoses typically corresponded to the greatest likelihood of measurement (*ORs* = 3.1–31.8), with the most pronounced effect for measurement of alcohol/SUDs (*OR* = 31.8).

Discussion

Recent work on integrated MH care has encouraged the design and implementation of scientifically rigorous program evaluation and QI activities (e.g., Funderburk & Shepardson, 2017; Hunter et al., 2018). Given the widespread utility—but documented underuse—of MBC in MH (e.g., Gilbody et al., 2002; Scott & Lewis, 2015), we elected to focus on this practice element as one example of how large EMR data set analyses can contribute to programmatic initiatives to improve MBC. Achieving change in MBC practices would likely be advantageous across settings given its potential to

Table 2

Electronic Mental Health Symptom Measurement by Category and Clinic Type

Measure domain and name (<i>N</i> = 161,113 administrations) ^a	<i>n</i>	%
Alcohol/SUD		
Alcohol Use Disorder Identification Test	5,837	3.6
Brief Addictions Monitor	27,498	17.1
Brief Addictions Monitor-Consumption Items	6	.0
Clinical institute Withdrawal Assessment-Alcohol Revised	2,335	1.4
Clinical Opiate Withdrawal Scale	2,129	1.3
Anxiety		
Beck Anxiety Inventory	2,942	1.8
Seven-item Generalized Anxiety Disorder Scale	944	.6
Clinician-rated		
Abnormal Involuntary Movement Scale	44,382	27.4
Brief Psychiatric Ratings Scale	83	.1
Global Deterioration Scale	79	.0
Cross-cutting		
Brief Symptom Inventory	2	.0
18-item Brief Symptom Inventory	31	.0
Symptom Checklist-90 Revised	55	.0
Depression/Suicidal ideation		
Automatic Thoughts Questionnaire	6	.0
Beck Depression Inventory-2	13,163	8.2
Beck Hopelessness Scale	113	.1
Beck Scale for Suicide Ideation	2	.0
Geriatric Depression Scale	1,737	1.1
Nine-item Patient Health Questionnaire	16,385	10.2
Zung Depression Scale	128	.1
Health/Quality of life/disability		
Millon Behavioral Medicine Diagnostic	73	.0
Multidimensional Health Locus of Control-Form C	4	.0
Short Form-36 Health Survey	1	.0
World Health Organization Disability Assessment Schedule 2	970	.6
Personality		
Millon Clinical Multiaxial Inventory 3	297	.2
Minnesota Multiphasic Personality Inventory 2	2,016	1.3
Minnesota Multiphasic Personality Inventory 2-Restructured Form	1,743	1.1
Personality Assessment Inventory	231	.1
PTSD		
Mississippi Scale	4,076	2.5
PTSD Checklist-Civilian Version	22,588	14.0
PTSD Checklist-Military Version	9,709	6.0
PTSD Checklist Stressor Specific	1,652	1.0

Symptom measurement by clinic type (*N* = 28,376 patients)^{b,c}

Screen type	PC-MHI (<i>n</i> = 9,720)	Specialty-MH (<i>n</i> = 21,135)	PTSD (<i>n</i> = 3,236)	SUD (<i>n</i> = 3,046)
Alcohol/SUD	2.5%	4.3%	4.4%	32.3%
Anxiety	2.1%	1.4%	2.0%	1.4%
Clinician-rated	7.4%	14.7%	13.0%	15.7%
Cross-cutting	.1%	.1%	.0%	.0%
Depression/Suicidal ideation	9.8%	8.2%	11.6%	20.3%
Health/Quality of life/disability	.1%	.6%	.8%	.4%
Personality	.8%	1.1%	1.6%	1.6%
PTSD	8.2%	7.9%	24.7%	10.9%

Note. PC-MHI = primary care-mental health integration; PTSD = posttraumatic stress disorder; Specialty-MH = specialty mental health; SUD = substance use disorder.

^a The miscellaneous measure category (*N* = 60; e.g., Acceptance and Action Questionnaire) was excluded from the table due to overall small cell sizes. ^b 0% rate values were observed in each clinic category. ^c Row sums exceed the total *N*, as patients may have had clinical encounters in more than one clinic.

Table 3
Logistic Regression of Patient Demographics, Encounters, and Diagnosis as Predictors of Selected Electronic Symptom Measurement

Predictor	Alcohol/SUD measure		Anxiety measure		Depression/Suicidal ideation measure		Personality measure		PTSD measure	
	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
Age										
19–40	1.0	—	1.0	—	1.0	—	1.0	—	1.0	—
41–55	.93	[.76, 1.1]	.69	[.51, .93]*	.81	[.71, .92]*	.63	[.45, .87]*	.59	[.52, .68]*
56–64	.84	[.64, 1.1]	.55	[.34, .87]*	.78	[.64, .94]*	.52	[.29, .89]*	.52	[.41, .66]*
65 or older	.47	[.34, .66]*	.50	[.29, .84]*	.74	[.59, .91]*	.19	[.08, .40]*	.42	[.32, .56]*
Service era										
Persian Gulf	1.0	—	1.0	—	1.0	—	1.0	—	1.0	—
Korean War	.22	[.05, .68]*	.21	[.02, .86]*	.64	[.41, .96]*	.22	[.01, 1.9]	.30	[.15, .56]*
Vietnam	.84	[.64, 1.1]	.81	[.50, 1.3]	.66	[.54, .80]*	.46	[.24, .89]*	.29	[.22, .38]*
Other	.99	[.81, 1.2]	.82	[.58, 1.2]	.82	[.71, .95]*	.68	[.45, 1.0]	.32	[.26, .39]*
Service-connected disability										
No	1.0	—	1.0	—	1.0	—	1.0	—	1.0	—
Yes	.64	[.55, .74]*	.80	[.63, 1.01]	.72	[.65, .80]*	.93	[.7, 1.3]	.86	[.76, .98]*
Sex										
Male	1.0	—	1.0	—	1.0	—	1.0	—	1.0	—
Female	.43	[.32, .56]*	.67	[.46, .95]*	.83	[.72, .96]*	.43	[.26, .68]*	.57	[.48, .67]*
Total number of MH encounters										
1	1.0	—	1.0	—	1.0	—	1.0	—	1.0	—
2–3	.88	[.65, 1.2]	1.2	[.82, 1.8]	1.2	[.95, 1.3]	.97	[.55, 1.7]	1.5	[1.1, 1.9]*
4–6	1.0	[.76, 1.4]	1.2	[.77, 1.8]	1.2	[.99, 1.4]	1.2	[.67, 2.2]	2.1	[1.7, 2.7]*
7–12	1.4	[1.02, 1.9]*	1.5	[.96, 2.2]	1.3	[1.1, 1.6]*	1.3	[.73, 2.4]	3.7	[2.9, 4.7]*
≥ 13	4.5	[3.4, 5.9]*	1.6	[1.04, 2.4]*	1.9	[1.6, 2.3]*	2.3	[1.3, 4.0]*	6.5	[5.1, 8.2]*
Number of MH diagnoses recorded (ever)										
0	1.0	—	1.0	—	1.0	—	1.0	—	1.0	—
1	10.3	[4.6, 24.8]*	1.6	[.66, 3.9]	4.1	[2.8, 6.2]*	.98	[.3, 2.9]	1.8	[1.0, 3.2]
2	10.4	[5.9, 19.7]*	2.0	[.99, 4.2]	4.4	[3.2, 6.2]*	2.1	[.9, 5.1]	1.4	[.9, 2.3]
≥ 3	31.8	[17.8, 61.1]*	3.1	[1.5, 6.9]*	6.8	[4.8, 9.6]*	5.5	[2.3, 13.8]*	1.5	[.9, 2.5]

(table continues)

Table 3 (continued)

Predictor	Alcohol/SUD measure		Anxiety measure		Depression/Suicidal ideation measure		Personality measure		PTSD measure	
	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
Type of diagnosis recorded										
Other MH disorder only	1.0	—	1.0	—	1.0	—	1.0	—	1.0	—
Personality disorder only	3.6	[.02, 72]*	5.5	[.96, 23.5]	5.1	[1.4, 15.3]*	.71	[.01, 6.6]	.66	[.01, 5.9]
Anxiety disorder only	.12	[.01, 2.2]	1.4	[.6, 3.6]	1.3	[.72, 2.5]	.44	[.2, 1.2]	1.0	[.49, 2.2]
PTSD/Trauma only	.47	[.11, 4.4]	.75	[.35, 1.9]	1.6	[.97, 3.0]	.22	[.1, .5]*	3.3	[1.8, 6.7]*
Mood disorder only	.49	[.11, 4.7]	.39	[.17, 1.02]	2.0	[1.2, 3.6]*	.17	[.1, .4]*	.56	[.29, 1.17]
Neurocognitive disorder only	3.0	[.02, 57.5]	.58	[.01, 5.2]	.71	[.1, 3.0]	1.7	[.2, 9.1]	2.0	[.21, 9.3]
Alcohol/SUD only	38.8	[10.7, 341]*	.20	[.05, .66]*	2.9	[1.7, 5.4]*	.05	[.01, .22]*	.09	[.03, .27]*
Schizophrenia/Psychotic disorder only	.36	[.03, 4.5]	.04	[.01, .36]*	.7	[.34, 1.5]	.12	[.01, .55]*	.20	[.05, .63]*
Alcohol/SUD and PTSD/trauma	17.2	[4.0, 161]*	.44	[.15, 1.4]	2.5	[1.3, 4.8]*	.17	[.06, .54]*	4.1	[2.0, 9.0]*
Anxiety and mood disorder	3.0	[.71, 28.5]	.83	[.32, 2.4]	1.8	[.98, 3.4]	.16	[.06, .50]*	.69	[.33, 1.6]
Anxiety and PTSD/trauma	2.3	[.51, 22.2]	.87	[.32, 2.6]	1.8	[.95, 3.5]	.11	[.03, .37]*	4.8	[2.4, 10.5]*
Mood and PTSD/trauma	2.5	[.60, 23.7]	.47	[.19, 1.4]	1.9	[1.1, 3.7]*	.13	[.05, .39]*	3.9	[2.0, 8.4]*
Any other combination of disorders	12.9	[3.0, 120]*	.42	[.16, 1.3]	2.2	[1.2, 4.2]*	.16	[.06, .49]*	.81	[.39, 1.8]

Note. SUD = substance use disorder; PTSD = posttraumatic stress disorder; OR = odds ratio; CI = confidence interval; MH = mental health; Trauma = trauma and stressor-related disorder. Service-connected disabilities are those disabilities objectively determined to be related to military service, either in origin or an exacerbation of an underlying medical or mental health condition. Patients who were not classified as having a mental health disorder ($n = 496$; 1.8% of overall sample) were excluded from analyses that used diagnosis as a predictor. CIs that do not include 0 are considered statistically significant.

* $p < .05$.

improve the efficiency and effectiveness of MH care, and provide specific and measurable points of reference for evaluation of program outcomes (Kearney, Wray, Dollar, & King, 2015).

Our findings related to screening were either intuitive or consistent with previous literature. For instance, screen administration was more common in PC-MHI versus other settings, which is sensible as lower steps in care may emphasize screening over more elaborate forms of evaluation (Kearney et al., 2015), and PC-MHI is a common gateway to MH services (Wray, Szymanski, Kearney, & McCarthy, 2012). Although overall screening rates varied by condition and clinic type, several (i.e., alcohol misuse, depression, and PTSD) appeared to be regularly administered by MH providers. Such observations are likely influenced by practice standards that encourage or mandate population-based screening for patients treated in the VHA (e.g., Tew, Klaus, & Oslin, 2010). For instance, annual screening for depression, alcohol misuse, and trauma symptoms is part of preprogrammed “clinical reminders” embedded in the EMR. Thus, total screening completion rates (i.e., numbers of patients ever screened at *any* VHA encounter) likely exceed our observations of rates of patients screened at MH encounters alone. While these screening tools may be administered by any clinical staff, they are often also completed by non-MH providers at other points in care, in which case positive results often trigger referral to MH providers for more thorough evaluation and treatment. Therefore, several explanations exist for the screening rates we observed, such as the following: (a) high rates of alcohol screening across settings may indicate that MH providers reliably respond to “clinical reminders” and/or that the screening has been adopted in general MH care, (b) lower rates of depression and trauma screening may indicate that screening requirements were met by other staff or that screening was not necessary (e.g., if diagnosis was preestablished), and (c) low rates of cognitive screening observed among older veterans may be influenced by the lack of strong recommendations to implement population-based screening in this age cohort (Kansagara & Freeman, 2010). Additional data capture and analysis could be planned to evaluate the extent to which such explanations may apply, for instance, to more

thoroughly investigate rates of MH screening by non-MH staff.

Our findings on symptom measurement were also consistent with previous research. Studies using PC-MHI-only samples have found similar rates of symptom measurement, with older age and female sex associated with a decreased likelihood of measurement (Beehler et al., 2016; Lilienthal et al., 2017). While we did not examine the trajectory or timeline of measurement directly, the low percentage of patients with standard measures documented suggests the possibility that routine outcome monitoring with electronic measures may not be occurring. Possible explanations include (a) providers’ time and resource limitations (Scott & Lewis, 2015), (b) uncertainty about how to use measurement in clinical decision-making (Gleacher et al., 2016), (c) an undervaluation of the benefits of MBC (Zimmerman & McGlinchey, 2008), or (d) a lack of specific guidelines/requirements to use standard measures. Coupled with the aforementioned patient-level predictors identified by our statistical models, these practical and empirically supported barriers to MBC may be targets for QI efforts.

Within the VHA, the national MBC Initiative (Department of Veterans Affairs, Veterans Health Administration (VHA), 2016) aims to implement several core MH outcome measures. In addition to this larger initiative, local QI plans can nonetheless be developed to address site-specific areas of suboptimal MBC uptake. Similar to the data capture and review procedures used in this study, administrators could review EMR data on MH screens and symptom measures to identify noteworthy trends in MBC practice within their clinic(s). Using an approach inspired by the plan–do–study–act framework (Deming, 1994), they may then outline specific hypotheses and goals for modification. For instance, given evidence of lower rates of symptom measurement among women veterans, and if such trends were also observed locally, a plausible target might be to improve rates of PTSD symptom monitoring among women veterans treated in their clinic(s). Administrators could plan to gather EMR data on women patients served by their clinic, as well as qualitative data from stakeholders in their workgroup on the perceived value and pros/cons of symptom monitoring in their practice (plan). Following an initial test of altered practices

(do), subsequent data may be collected to identify whether and how the proposed practice changes impacted patient outcomes and clinic processes (e.g., increased use of MH measures with women veterans), to include any unintentional consequences, such as disruption of work flow (study). Data collected as part of this process could then be used to develop or refine standard operating procedures that would reduce the opportunity for disparities in care by ensuring routine outcome monitoring for all patients (act; Kearney et al., 2015).

While we opted for a robust evaluation of data collected over a 1-year observation, worth mention is that real-world QI need not be complicated to be informative; data-driven QI can be based on descriptive data collected over a relatively short period of time. While it is possible that certain statistical tests (e.g., Fisher's exact test to evaluate associations in Sex $\times$ Measurement contingency tables) may enhance the quality of conclusions drawn, basic rates of completed assessments over time are simple to calculate and are generally understood by most audiences. Furthermore, undue emphasis on large and representative data sets (as opposed to small, practical samples) could emerge as a barrier to engaging in and sustaining the ongoing QI efforts necessary to promote a culture of change and rapid improvement (Etchells, Ho, & Shojania, 2016; Etchells & Woodcock, 2018).

As with any chart-based methodology, there are limitations that warrant mention. Our analytic emphasis was on identifying the rates and relative likelihood of electronic screening and measurement across outpatient MH settings as an example of how these data can be organized, analyzed, and interpreted in support of programmatic MBC initiatives. As such, we are unable to provide insight into how measurement data were integrated into individual treatment or interprofessional communication. Such questions cannot be answered by summary data alone. Because we focused solely on screening and measurement recorded in the EMR during MH encounters, our estimates do not pertain to other instances of MBC (e.g., paper-and-pencil administrations) that would be necessary for broader discussion of MBC practices. As a substantial proportion of race and ethnicity data were missing from our study sample, we were also unable to explore these factors as potential predictors of screening or measurement. Fi-

nally, our data collection spanned just one year and was limited to a single geographic region. It is therefore plausible that sample (e.g., high rates of objective disability) and other contextual factors (e.g., data collection predated a major national QI initiative on a similar topic) would further limit generalizability to other samples and settings.

Though our data are not generalizable, they nonetheless illustrate how data-driven decision-making can be employed in clinically relevant QI activities. While it is difficult to make broad recommendations regarding which outcome measures to use, under what circumstances, and how often to administer them at each step in care, clinic administrators may nonetheless benefit from conducting similarly focused investigations of current practices, gathering feedback from key stakeholders, and developing systematic plans to improve uptake of MBC. Measure recommendations and assessment schedules can be included as part of clinic-specific operating procedures that reflect local context, resources, and standards of practice. Though we assert that MBC contributes to higher-quality care, patient-level barriers remain understudied, and may represent important contributors to the rates of measurement we observed. Further work is necessary to (a) better understand patient, provider, and systems-level facilitators and barriers to electronic MH measurement and MBC; (b) develop appropriate educational/systems-level interventions to improve uptake of measurement; and (c) describe how providers use measurement data to alter treatment.

References

- Allison, P. D. (2008, March). *Convergence failures in logistic regression* (Paper 360–2008). Paper presented at the SAS Global Forum, San Antonio, TX. Retrieved from <http://www2.sas.com/proceedings/forum2008/360-2008.pdf>
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.
- Beehler, G. P., King, P. R., Vair, C. L., Gass, J., & Funderburk, J. S. (2016). Measurement of common mental health conditions in VHA co-located, collaborative care. *Journal of Clinical Psychology in Medical Settings*, 23, 378–388. <http://dx.doi.org/10.1007/s10880-016-9478-3>
- Beehler, G. P., Lilienthal, K. R., Possemato, K., Johnson, E. M., King, P. R., Shepardson, R. L., . . .

- Wray, L. O. (2017). Narrative review of provider behavior in primary care behavioral health: How process data can inform quality improvement. *Families, Systems & Health*, 35, 257–270. <http://dx.doi.org/10.1037/fsh0000263>
- Bickman, L., Kelley, S. D., Breda, C., de Andrade, A. R., & Riemer, M. (2011). Effects of routine feedback to clinicians on mental health outcomes of youths: Results of a randomized trial. *Psychiatric Services*, 62, 1423–1429. <http://dx.doi.org/10.1176/appi.ps.002052011>
- Deming, W. E. (1994). *The new economics for industry, government, education* (2nd ed.). Cambridge, MA: MIT Press.
- Department of Veterans Affairs, Veterans Health Administration (VHA). (2016). *Measurement based care (MBC) in mental health initiative*. Washington, DC: Author.
- Derogatis, L. R. (1993). *BSI: Brief Symptom Inventory: Administration, scoring, and procedure manual*. Minneapolis, MN: National Computer Systems.
- Dowrick, C., Leydon, G. M., McBride, A., Howe, A., Burgess, H., Clarke, P., . . . Kendrick, T. (2009). Patients' and doctors' views on depression severity questionnaires incentivised in U.K. quality and outcomes framework: Qualitative study. *British Medical Journal*, 338, b663. <http://dx.doi.org/10.1136/bmj.b663>
- Eisen, S. V., Dickey, B., & Sederer, L. I. (2000). A self-report symptom and problem rating scale to increase inpatients' involvement in treatment. *Psychiatric Services*, 51, 349–353. <http://dx.doi.org/10.1176/appi.ps.51.3.349>
- Etchells, E., Ho, M., & Shojania, K. G. (2016). Value of small sample sizes in rapid-cycle quality improvement projects. *British Medical Journal Quality & Safety*, 25, 202–206. <http://dx.doi.org/10.1136/bmjqs-2015-005094>
- Etchells, E., & Woodcock, T. (2018). Value of small sample sizes in rapid-cycle quality improvement projects 2: Assessing fidelity of implementation for improvement interventions. *British Medical Journal Quality & Safety*, 27, 61–65. <http://dx.doi.org/10.1136/bmjqs-2017-006963>
- Firth, D. (1993). Bias reduction of maximum likelihood estimates. *Biometrika*, 80, 27–38. <http://dx.doi.org/10.1093/biomet/80.1.27>
- Funderburk, J. S., & Shepardson, R. L. (2017). Real-world program evaluation of integrated behavioral health care: Improving scientific rigor. *Families, Systems & Health*, 35, 114–124. <http://dx.doi.org/10.1037/fsh0000253>
- Gilbody, S. M., House, A. O., & Sheldon, T. A. (2002). Psychiatrists in the U.K. do not use outcomes measures. National survey. *The British Journal of Psychiatry*, 180, 101–103. <http://dx.doi.org/10.1192/bjp.180.2.101>
- Gleacher, A. A., Olin, S. S., Nadeem, E., Pollock, M., Ringle, V., Bickman, L., . . . Hoagwood, K. (2016). Implementing a measurement feedback system in community mental health clinics: A case study of multilevel barriers and facilitators. *Administration and Policy in Mental Health*, 43, 426–440. <http://dx.doi.org/10.1007/s10488-015-0642-0>
- Guy, W. (1976). *ECDEU assessment manual for psychopharmacology: Revised* (DHEW Publication Number ADM 76–338). Rockville, MD: U.S. Department of Health, Education and Welfare, Public Health Service, Alcohol, Drug Abuse and Mental Health Administration, National Institute of Mental Health Psychopharmacology Research Branch, Division of Extramural Research Programs.
- Haberer, J. E., Trabin, T., & Klinkman, M. (2013). Furthering the reliable and valid measurement of mental health screening, diagnoses, treatment and outcomes through health information technology. *General Hospital Psychiatry*, 35, 349–353. <http://dx.doi.org/10.1016/j.genhosppsych.2013.03.009>
- Heinze, G., & Schemper, M. (2002). A solution to the problem of separation in logistic regression. *Statistics in Medicine*, 21, 2409–2419. <http://dx.doi.org/10.1002/sim.1047>
- Hunter, C. L., Funderburk, J. S., Polaha, J., Bauman, D., Goodie, J. L., & Hunter, C. M. (2018). Primary care behavioral health (PCBH) model research: Current state of the science and a call to action. *Journal of Clinical Psychology in Medical Settings*, 25, 127–156. <http://dx.doi.org/10.1007/s10880-017-9512-0>
- Institute of Medicine. (2001). *Crossing the quality chasm: A new health system for the 21st century*. Washington, DC: National Academy Press.
- Kansagara, D., & Freeman, M. (2010). A systematic evidence review of the signs and symptoms of dementia and brief cognitive tests available in VA (VA-ESP Project Number 05–225). Washington, DC: Department of Veterans Affairs.
- Katon, W. J., Lin, E. H., Von Korff, M., Ciechanowski, P., Ludman, E. J., Young, B., . . . McCulloch, D. (2010). Collaborative care for patients with depression and chronic illnesses. *The New England Journal of Medicine*, 363, 2611–2620. <http://dx.doi.org/10.1056/NEJMoa1003955>
- Kearney, L. K., Wray, L. O., Dollar, K. M., & King, P. R. (2015). Establishing measurement-based care in integrated primary care: Monitoring clinical outcomes over time. *Journal of Clinical Psychology in Medical Settings*, 22, 213–227. <http://dx.doi.org/10.1007/s10880-015-9443-6>
- Kroenke, K., Spitzer, R. L., & Williams, J. B. (2003). The Patient Health Questionnaire-2: Validity of a two-item depression screener. *Medical Care*, 41, 1284–1292. <http://dx.doi.org/10.1097/01.MLR.0000093487.78664.3C>

- Lilienthal, K. R., Buchholz, L. J., King, P. R., Vair, C. L., Funderburk, J. S., & Beehler, G. P. (2017). Mental health measurement among women veterans receiving co-located, collaborative care services. *Psychology Health and Medicine*, 22, 1192–1202. <http://dx.doi.org/10.1080/13548506.2017.1290809>
- Long, J. A., Bamba, M. I., Ling, B., & Shea, J. A. (2006). Missing race/ethnicity data in Veterans Health Administration based disparities research: A systematic review. *Journal of Health Care for the Poor and Underserved*, 17, 128–140. <http://dx.doi.org/10.1353/hpu.2006.0029>
- Sapyta, J., Riemer, M., & Bickman, L. (2005). Feedback to clinicians: Theory, research, and practice. *Journal of Clinical Psychology*, 61, 145–153. <http://dx.doi.org/10.1002/jclp.20107>
- Scott, K., & Lewis, C. C. (2015). Using measurement-based care to enhance any treatment. *Cognitive and Behavioral Practice*, 22, 49–59. <http://dx.doi.org/10.1016/j.cbpra.2014.01.010>
- Tew, J., Klaus, J., & Oslin, D. W. (2010). The Behavioral Health Laboratory: Building a stronger foundation for the patient-centered medical home. *Families, Systems & Health*, 28, 130–145. <http://dx.doi.org/10.1037/a0020249>
- Veterans Benefits Administration. (2013). *Disability compensation*. Retrieved from <http://www.benefits.va.gov/COMPENSATION/types-disability.asp>
- Ware, J. E., Jr., & Sherbourne, C. D. (1992). The MOS 36-item Short-Form Health Survey (SF-36): I. Conceptual framework and item selection. *Medical Care*, 30, 473–483. <http://dx.doi.org/10.1097/00005650-199206000-00002>
- Williams, J. S., Walker, R. J., & Egede, L. E. (2016). Achieving equity in an evolving healthcare system: Opportunities and challenges. *The American Journal of the Medical Sciences*, 351, 33–43. <http://dx.doi.org/10.1016/j.amjms.2015.10.012>
- Wray, L. O., Szymanski, B. R., Kearney, L. K., & McCarthy, J. F. (2012). Implementation of primary care-mental health integration services in the Veterans Health Administration: Program activity and associations with engagement in specialty mental health services. *Journal of Clinical Psychology in Medical Settings*, 19, 105–116. <http://dx.doi.org/10.1007/s10880-011-9285-9>
- Zarit, S. H., Reever, K. E., & Bach-Peterson, J. (1980). Relatives of the impaired elderly: Correlates of feelings of burden. *The Gerontologist*, 20, 649–655. <http://dx.doi.org/10.1093/geront/20.6.649>
- Zimmerman, M., & McGlinchey, J. B. (2008). Why don't psychiatrists use scales to measure outcome when treating depressed patients? *The Journal of Clinical Psychiatry*, 69, 1916–1919. <http://dx.doi.org/10.4088/JCP.v69n1209>

Received February 2, 2018

Revision received October 10, 2018

Accepted October 10, 2018 ■

E-Mail Notification of Your Latest Issue Online!

Would you like to know when the next issue of your favorite APA journal will be available online? This service is now available to you. Sign up at <https://my.apa.org/portal/alerts/> and you will be notified by e-mail when issues of interest to you become available!