AMERICAN PSYCHOLOGICAL ASSOCIATION
CFHA COLLABORATIVE FAMILY HEALTHCARE ASSOCIATION

Families, Systems, & Health

© 2021 American Psychological Association
ISSN: 1939-0602

2021, Vol. 39, No. 2, 259–268
https://doi.org/10.1037/fsh0000608

Measurement-Based Care in Integrated Health Care: A Randomized Clinical Trial

Barry L. Duncan, PsyD[1], Robert J. Reese, PhD[2], Alexander J. Lengerich, PhD[3], Brian DeSantis, PsyD[4], Cassie V. Comeau, PhD[5], and Yajaira Johnson-Esparza, PhD[6]

[1] Better Outcomes Now, West Palm Beach, Florida, United States
[2] Department of Special Education, Rehabilitation, and Counseling, Auburn University
[3] James A. Haley Veterans' Hospital, Tampa, Florida, United States
[4] Peak Vista Community Health Centers, Colorado Springs, Colorado, United States
[5] Summit Community Care Clinic, Frisco, Colorado, United States
[6] Salud Family Health Clinics, Fort Collins, Colorado, United States

Introduction: Many suggest that the next step for integrated care is widespread implementation of measurement-based care (MBC). Although the measures most associated with MBC are standardized, no randomized clinical trial has demonstrated their use to improve psychotherapeutic outcomes with embedded behavioral health providers in integrated care. Two evidence-based MBC systems have been studied in a variety of behavioral health environments, but neither system has been investigated in integrated health care. Addressing this gap in the literature, the present study evaluated the use of MBC, specifically the Partners for Change Outcome Management System, in three integrated care sites. ***Method:*** Using a randomized design within routine care, treatment as usual (TAU; n = 133) was compared using the Outcome Rating Scale (ORS) and Patient Health Questionnaire–9 (PHQ–9) with a feedback condition (n = 147) in which behavioral health providers had access to patient-generated outcome (ORS only) and alliance information at each session. ***Results:*** Patients in the feedback condition demonstrated significantly more improvement than those in the TAU condition posttreatment on the ORS. Patients in the feedback

Barry L. Duncan, PsyD https://orcid.org/0000-0001-8371-2289

Robert J. Reese, PhD https://orcid.org/0000-0002-3658-2968

Brian DeSantis, PsyD https://orcid.org/0000-0002-3410-3046

Yajaira Johnson-Esparza, PhD https://orcid.org/0000-0002-4166-8539

Cassie V. Comeau is now in Private Practice, Dillon, Colorado. Yajaira Johnson-Esparza is now at the University of Texas Health San Antonio, San Antonio, Texas.

Barry L. Duncan, PsyD, is a coholder of the copyright of the Partners for Change Outcome Management System (PCOMS) instruments. The measures are free for individuals, but Duncan receives royalties from licenses issued to groups and organizations. In addition, the web-based application of PCOMS, BetterOutcomes Now.com, is a commercial product, and he receives profits based on sales. The authors acknowledge the staff and behavioral health providers at the three study sites for not only their commitment to this study but also, more importantly, their dedication to their patients.

Barry L. Duncan, PsyD, Brian DeSantis, PsyD, and Robert J. Reese, PhD, conceptualized, designed, and administered the study with the support of Alexander J. Lengerich, PhD. Brian DeSantis, PsyD, Cassie V. Comeau, PhD, and Yajaira Johnson-Esparza, PhD, conducted the onsite investigation, supervised the participant behavioral health providers in the study, and participated in data curation. Robert J. Reese, PhD, Alexander J. Lengerich, PhD, and Barry L. Duncan, PsyD, conducted the formal analysis. Barry L. Duncan, PsyD, Robert J. Reese, PhD, and Alexander J. Lengerich, PhD, wrote the manuscript with critical feedback from Brian DeSantis, PsyD, Cassie V. Comeau, PhD, and Yajaira Johnson-Esparza, PhD.

Correspondence concerning this article should be addressed to Barry L. Duncan, PsyD, Better Outcomes Now, PsyD, Better Outcomes Now, 816 Avon Road, West Palm Beach, FL 33401, United States. Email: barryduncan@betteroutcomesnow.com

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

condition also achieved significantly more clinically significant change as measured by both the ORS and PHQ–9. Feedback condition patients also attended significantly more sessions and dropped out significantly less that TAU patients. ***Discussion:*** Although our findings need to be replicated, this study offers evidence that the improved outcomes and reduced dropouts associated with MBC in traditional behavioral health centers also occur in integrated care settings.

Public Significance Statement
Previous studies about measurement-based care, a feedback intervention designed to identify patients not responding to psychotherapy to enable providers to restore treatment to a positive trajectory, have been conducted in traditional behavioral health settings. This study is the first to demonstrate the feedback effects of improved outcomes and efficiency within the workflow demands of integrated care.

Keywords: integrated care, measurement-based care (MBC), Partners for Change Outcome Management System (PCOMS), patient feedback, routine outcome monitoring (ROM)

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Many have suggested that the next step for integrated care is the implementation of measurement-based care (MBC; Fortney et al., 2017; Kearney et al., 2015; King et al., 2018; Lewis et al., 2019; Peterson et al., 2018). MBC is often defined with four components (Kearney et al., 2015; Peterson et al., 2018): (a) routine and frequent administration of outcome measures; (b) practitioner review of data; (c) patient review of data; and (d) shared decision-making and collaborative reevaluation of the treatment plan informed by data. A review by Fortney et al. (2017) concluded: "There is mounting empirical evidence from trials that . . . patients randomly assigned to MBC have better outcomes than patients randomly assigned to usual care" (p. 186).

MBC in integrated care typically involves the nine-item Patient Health Questionnaire (PHQ-9; Kroenke & Spitzer, 2002), the Generalized Anxiety Disorder (GAD-7; Spitzer et al., 2006), and the Brief Addiction Module (BAM; Cacciola et al., 2013). Although these measures are standardized, no randomized clinical trial (RCT) has demonstrated their use to improve psychotherapeutic outcomes with embedded behavioral health providers in integrated care (Peterson et al., 2018). For example, the recent review advocating MBC (Fortney et al., 2017) cited fifteen supporting RCTs. None included the PHQ-9, GAD-7, or BAM. Three of the fifteen, although demonstrating support for MBC, did not include one or more of the four components listed above, usually frequent administration and patient collaboration.

The majority of the evidence for MBC, called routine outcome monitoring or systematic client feedback in the psychotherapy literature, comes from two systems with comprehensive protocols and predictive algorithms based on large databases: The Outcome Questionnaire 45.2 (OQ; Lambert, 2015) System and the Partners for Change Outcome Management System (PCOMS; Duncan & Reese, 2015). In the Fortney et al. (2017) review, twelve of the fifteen RCTs listed in support of MBC evaluated these two systems. Both were included in the Substance Abuse and Mental Health Administration's National Registry of Evidence-Based Programs and Practices.

Lambert's OQ System is the pioneer of MBC with a proven track record of improving outcomes. Inspired by Lambert, PCOMS (Duncan, 2014; Duncan & Sparks, 2002) incorporates the four components of MBC defined above and employs two, four-item scales, one focusing on outcome and the other on the therapeutic alliance. There are eight RCTs that support the efficacy of PCOMS over treatment as usual (TAU; Anker et al., 2009; Brattland et al., 2018; Cooper et al., 2021; Reese et al., 2009, 2010; Schuman et al., 2015; She et al., 2018; Slone et al., 2015). In a meta-analysis of PCOMS, Lambert et al. (2018) found an average effect size (SMD) of .40, with significant heterogeneity across outcomes. In addition to its empirical support, the PCOMS outcome measure is ultrabrief (four items), is validated as a screener in integrated care (DeSantis et al., 2017) as well as an outcome measure, and may offer a more feasible strategy for the workflow demands of integrated care.

Given the heterogeneity of findings, that all studies have not demonstrated a feedback effect

(Østergård et al., 2020), and the need for "high-intensity implementation" (Peterson et al., 2018, p. 20), several cautions are warranted regarding the effectiveness of PCOMS (Duncan & Sparks, 2020). First, the feedback effect seems dependent on high adherence and fidelity monitoring (Cooper et al., 2021). Second, although notable exceptions exist (e.g., Brattland et al., 2018), most PCOMS studies have used only the intervention measure combined with more "real world" outcomes like dropout and number of sessions. Finally, all studies finding a feedback effect averaged four or more sessions in the experimental condition.

The two evidence-based MBC systems have been studied in a variety of behavioral health environments, but neither has been investigated in an integrated health care setting. Addressing this gap in the literature as well as the noted cautions, the purpose of this study was to evaluate the use of MBC, specifically PCOMS, in an integrated care setting. Using a randomized design within routine care, TAU was compared with a feedback condition in which clinicians had access to patient-generated outcome (ORS only) and alliance information at each session. First, we hypothesized that patients in the feedback condition would exhibit greater pre-/posttreatment gains on ORS and PHQ-9 scores compared with patients in the TAU condition. Second, we hypothesized that more patients in the MBC condition would experience clinically significant change on both measures. Finally, we hypothesized that the average number sessions would be higher and the premature termination rate lower in the feedback condition.

Method

Participants

This study (1006145-2) was approved by the University of Northern Colorado Institutional Review Board, and written informed consent was obtained from all participants.

Data were collected on patients at three Federally Qualified Health Centers (FQHC) in Colorado. Patients were required to be at least 18 years of age and had not received behavioral health services in the past three months. As shown in Figure 1, a total of 304 patients enrolled in the study. Nineteen participants were dropped for enrollment or protocol errors, leaving 285 participants. Patients ranged in age from 18 to 79 with a mean age of 37.87 (SD = 13.77); 69.5% were female (n = 198), and 30.5% were male (n = 87). The sample identified as White (69.8%, n = 199), followed by Hispanic/Latinx (24.9%, n = 71), African American (2.1%, n = 6), Asian/Pacific Islander (.7%, n = 2), Biracial (.7%, n = 2), and Other (1.8%, n = 5).

Also shown in Figure 1, 55 (19.3%) participants attended only one session (TAU = 37 and Feedback = 18), and 13 did not have an initial valid score on the ORS (>32) or were missing an initial or post score on one of the outcome measures (ORS or the PHQ-9). A post score was defined as a score for the last session attended, planned or unplanned. These patients, therefore, could not be included in the final sample analyses. The final sample consisted of 211 patients for the ORS and 204 patients for the PHQ-9 for pre–post comparisons across treatment conditions. There was no significant difference in pre-ORS score between patients who attended one session (M = 19.87, SD = 8.93) and those who attended more (M = 17.93, SD = 8.16), $t(278) = 1.50$, $p = .13$. There was a significant difference in pre–PHQ score between patients who attended one session ($M = 9.28$, $SD = 5.46$) and those attending more ($M = 11.54$, $SD = 6.55$), $t(277) = 2.33$, $p = .02$.

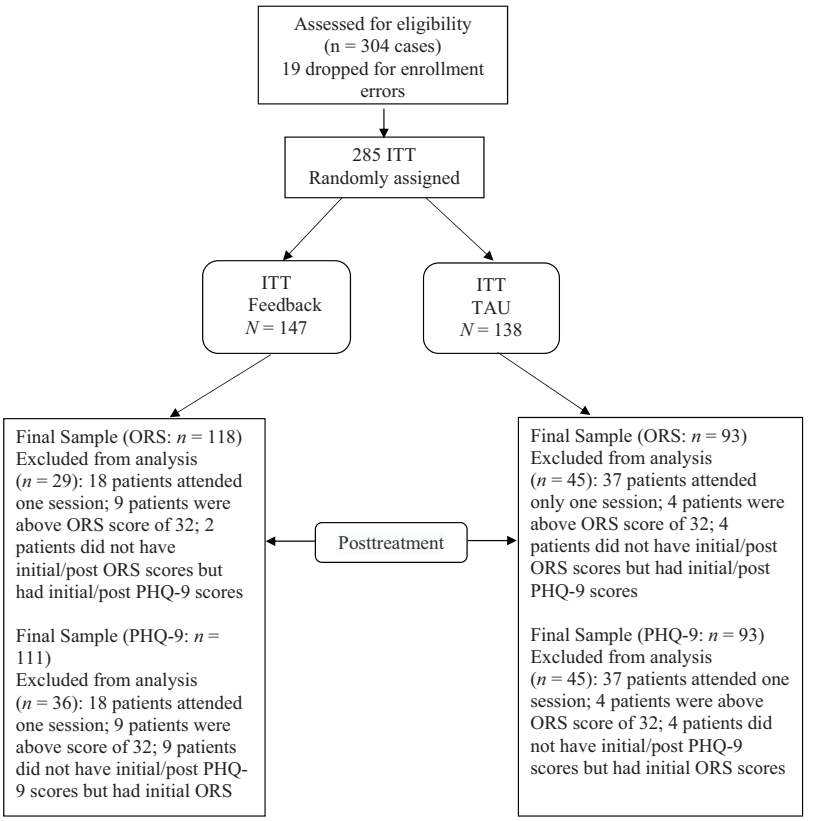
Behavioral Health Providers

Sixteen behavioral health providers (BHP) participated in the study, ranging in age from 25 to 56 with a mean of 35.50 (SD = 8.47); 14 were female (87.5%,). BHPs identified as Hispanic/Latinx (43.8%, n = 7), followed by White (31.3%, n = 5), and African American (25%, n = 4). They primarily reported using an integrative approach grounded in cognitive behavioral therapy (68.75%, n = 11), followed by eclectic (18.75%, n = 3), brief/strategic/reality (6.3%, n = 1), and integrative (6.3%, n = 1). Six BHPs were students (37.5%), four were counselors (25%, n = 4), three were postdoctoral fellows (18.8%, n = 3), two were social workers (12.5%) and one was a psychologist (6.3%).

BHPs received PCOMS training consisting of three 20-minute webinars and a one-hour question and answer session. They also had access to online training materials. BHPs received weekly PCOMS supervision of experimental patients by psychologists trained in PCOMS supervision. In the TAU condition, BHPs received weekly supervision without PCOMS data. In addition to PCOMS supervision, BHP fidelity to patient feedback was assessed

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Figure 1
Participant Flow Into Treatment Conditions and Data Analysis



Note. ITT = intent-to-treat; TAU = treatment-as-usual condition; PHQ-9 = Patient Health Questionnaire – 9; ORS = Outcome Rating Scale.

using the PCOMS Provider Adherence Scale (Duncan, 2011).

Outcome Measures

The Outcome Rating Scale (ORS; Miller et al., 2003) is a visual analog scale consisting of four 10-cm lines, corresponding to four domains (individual, interpersonal, social, overall). The ORS provides a functional rather than symptom-based assessment that casts a wider net in screening and treating patient problems (DeSantis et al., 2017). Its brevity encourages a real-time scoring and discussion of patient goals, as well as immediate feedback on treatment progress. Patients place a mark on each line to represent their functioning in each domain or a touch or click on an electronic device. A centimeter ruler is used to measure the distance to the nearest millimeter from the left end of the scale to the patient's mark on each line, or it is automatically scored by a web system. Scores range from 0 to 40, with lower scores signaling higher distress.

Multiple studies have demonstrated the reliability and validity of the ORS (Duncan & Reese, 2015). Validation (Bringhurst et al., 2006; Campbell & Hemsley, 2009; Miller et al., 2003) and clinical studies (e.g., Anker et al., 2009; She et al., 2018) have yielded coefficient alphas ranging from .82 to .92. For the current study, the coefficient alpha was .81. Concurrent validity of the ORS is demonstrated by moderately strong correlations with other measures, including the OQ45 and PHQ-9 (Campbell & Hemsley, 2009; DeSantis et al., 2017). Finally, hundreds of thousands of administrations have determined cutoff and reliable change norms as well as algorithms for expected treatment responses (ETR; Duncan & Sparks, 2018).

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

The PHQ-9 (Kroenke & Spitzer, 2002) is a widely used nine-item depression screener. Studies have demonstrated that the PHQ-9 generates reliable and valid scores, with a Cronbach's alpha ranging from .86 to .89. For the current study, the coefficient alpha was .85.

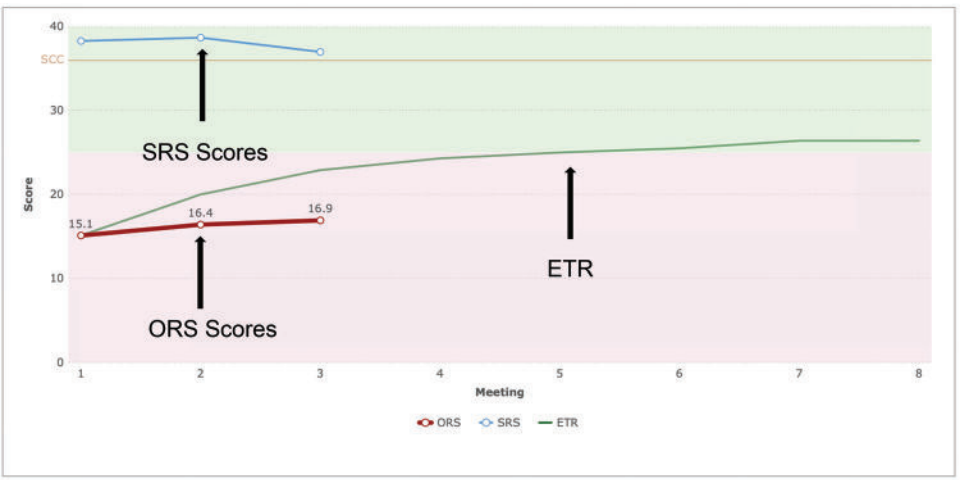
Procedure

Standard practice at each site is that BHPs are consulted by medical provider request or if screening by the PHQ-9 revealed a score in the clinical range. During the study, once the BHP was consulted, the ORS was also used to screen for possible intervention. Participants were randomized via a randomized block design to either the feedback or the TAU condition. BHPs served as their own control, with half of their patients in each condition. The FQHCs use a brief model, typically up to six sessions lasting thirty minutes. The PHQ-9 and ORS were administered via hard copy by the support staff to the TAU patients. In the experimental condition, support staff administered the PHQ-9. BHPs administered the ORS and conducted the feedback intervention via a tablet and the PCOMS web-based application, Better Outcomes Now. BHPs did not have access to data from the ORS or PHQ-9 for patients in the TAU condition; clinicians did not have access to information from the PHQ-9 for patients in the feedback condition.

BHPs in the feedback condition were instructed to follow the PCOMS evidence-based protocol (Duncan & Sparks, 2018). The ORS was administered by the BHP at the beginning of each session via the web application. An inspection of Figure 2 reveals that the web system totals and plots the patient's score on a graph that indicates progress, or lack thereof, across the course of treatment. The graph includes the expected treatment response (ETR) for the patient's intake score and provides easily understood graphics about progress relative to the ETR. Patient progress is discussed at each session. Toward the end of the session, the Session Rating Scale (SRS; Duncan et al., 2003), a four-item visual analogue alliance scale, was also administered to detect potential breaches. The SRS allows the BHP to discuss any patient concerns and how treatment may better fit patient expectations. The total score is charted on the same graph as the ORS. BHPs and patients in the feedback condition, therefore, had ongoing, real-time access to ORS and SRS scores, and ETRs.

Figure 2

The Web-Based PCOMS Software, Better Outcomes Now, Graph of Scores, and Expected Treatment Response (ETR)



First (Date)	Last (Date)	Meetings	Change	Progress
15.1 (2019-03-25)	16.9 (2020-12-15)	3	1.8	23.2%

Note. ORS = Outcome Rating Scale; SRS = Session Rating Scale; ETR = expected treatment response. Progress meter shows patient to be less than 50% of ETR, suggesting a conversation about changing therapeutic directions. See the online article for the color version of this figure.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Analytic Strategy

Hierarchical linear modeling (Raudenbush & Byrk, 2002) was used to account for the nested design (i.e., patients nested within BHPs). The first hypothesis predicted that patients in the feedback condition would demonstrate better outcomes than those in the TAU condition as measured by both the ORS and the PHQ-9 after controlling for pre-ORS and pre–PHQ-9 scores. To address this hypothesis, we first constructed a two-level model for each outcome measure (with initial ORS/PHQ-9 scores grand-mean centered to serve as a covariate) to measure the BHP-level variance. We then constructed two-level multilevel models (Feedback model) to represent the amount of variation in post-ORS/PHQ-9 scores at the patient and BHP levels.

The second hypothesis predicted that significantly more patients in the feedback condition would achieve clinically significant change compared with TAU based on pre–post change scores. To achieve clinically significant change, the patient must begin in the clinical range (ORS $<$ 25; PHQ-9 $>$ 9), achieve reliable change (six points on the ORS and five points on the PHQ-9), and complete treatment in the nonclinical range. To address this hypothesis, percentages of patients in each condition who achieved clinically significant change were calculated and chi-square analyses were conducted comparing the two conditions.

The third hypothesis predicted that patients in the feedback condition would have a lower rate of dropout and attend more sessions than those in the TAU condition. Dropout rate was calculated as the percentage of patients in each condition who did not complete treatment in the nonclinical range ($>$25 on the ORS).

Results

Pre–post ORS and PHQ-9 mean total scores and standard deviations for each treatment condition can be observed in Table 1, including the ITT and final samples. For the final sample ($N = 211$, ORS; $N = 204$, PHQ-9), independent samples t tests found that the pretreatment mean differences were not statistically significant, $t(1, 209) = 1.15$, $p > .05$, ORS and $t(1, 276) = .96$, $p > .34$, PHQ-9.

The results from the covariate-only model are shown in Table 2. This baseline model estimated covariance parameters to compute an intraclass coefficient (ICC) and determine the amount of variability at the BHP level. The ICC showed that roughly 4% of the variance (.044, ORS; .041, PHQ-9) was accounted for by BHP differences, which is larger compared with other PCOMS studies (Anker et al., 2009; Reese et al., 2010; She et al., 2018) that were roughly 2% or smaller. The covariate-only model indicates a significant positive slope ($\gamma_{10} = .54$, $p < .001$) between initial ORS scores and post-ORS scores across patients as well as initial PHQ-9 scores and post–PHQ-9 scores ($\gamma_{10} = -.55$, $p < .001$). Patients with an average pre-ORS score had an average post-ORS score equal to 22.60 (γ_{00}) and patients with an average pre-PHQ-9 score had an average post–PHQ-9 score equal to 8.35 (γ_{00}). The standardized mean effect size, Cohens' d (see the footnote to Table 2 for formula used), from pre- to post-ORS was $d = 1.12$ and .56 on the PHQ-9.

The Feedback model added the treatment condition (FEEDBACK) and number of sessions as fixed effects. This model evaluated the differences in slopes for each treatment condition while controlling for initial ORS and PHQ-9 scores. Regarding the

Table 1
Means and Standard Deviations for the ORS and PHQ-9 for Intent to Treat and Final Samples Per Treatment Condition

Measure	Feedback		TAU	
	ITT M (n = 147)	Final M (n = 118)	ITT M (n = 133)	Final M (n = 93)
ORS: ITT N = 280[a], Final N = 211				
Pre	18.13 (8.24)	16.40 (7.13)	18.48 (8.46)	17.57 (7.57)
Post	24.62 (9.53)	24.38 (9.68)	21.25 (8.27)	20.97 (8.13)
	ITT M (n = 144)	Final M (n = 111)	ITT M (n = 134)	Final M (n = 93)
PHQ-9: ITT N = 278[b], Final N = 204				
Pre	11.50 (6.58)	12.11 (6.38)	10.76 (6.20)	11.45 (6.36)
Post	8.48 (6.22)	8.56 (6.03)	8.73 (5.85)	8.86 (5.82)
Sessions	4.24 (4.08)	4.76 (4.19)	3.01 (2.19)	3.73 (1.96)

Note. ORS = Outcome Rating Scale; PHQ-9 = Patient Health Questionnaire-9; ITT = intent-to-treat; TAU = treatment-as-usual condition. Standard deviations are presented within parentheses.
[a] Five patients did not have initial ORS scores but had initial PHQ-9 scores. [b] Seven patients did not have initial PHQ-9 scores but had initial ORS scores.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 2
Fixed and Random Effect Estimates Predicting Post ORS and PHQ-9 Scores for Final Sample

Measure	Covariate-only model Coefficient (*SE*)	Feedback model Coefficient (*SE*)
ORS		
Fixed effects		
Intercept – *M* post-ORS (γ_{00})	22.63*** (0.77)	18.57*** (1.02)
Patient pre-ORS (γ_{10})	0.54*** (0.08)	0.58*** (0.07)
Feedback (γ_{01})		3.45*** (1.08)
Session (γ_{20})		0.55*** (0.16)
Random effects		
Patient intercept variance ($\sigma^2_{\text{Patient}}$)	64.08*** (6.38)	59.10*** (5.90)
BHPs intercept variance (σ^2_{BHPs})	2.93	0.73
Standardized effect size	1.11[a]	0.38[b]
PHQ-9		
Fixed effects		
Intercept – *M* post-PHQ-9 (γ_{00})	8.71*** (0.44)	10.18*** (0.61)
Patient pre-PHQ-9 (γ_{10})	−0.55*** (0.05)	−0.57*** (0.05)
Feedback (γ_{01})		−0.33 (0.65)
Session (γ_{20})		−0.30** (0.09)
Random effects		
Patient intercept variance ($\sigma^2_{\text{Patient}}$)	21.30*** (2.18)	20.75*** (2.15)
BHPs intercept variance (σ^2_{BHPs})	0.90	0.26
Standardized effect size	.56[a]	0.06[b]

Note. *SE* = standard error; Patient pre-ORS = patient's initial ORS score grand mean centered; Feedback = type of feedback condition (0 = treatment as usual; 1 = feedback); Patient pre-PHQ-9 = patient's initial PHQ-9 score grand mean centered [a] Cohen's $d = (M_{\text{post}} - M_{\text{pre}})/SD_{\text{pre}}$ [b] Hedges' $g = \frac{\gamma_{01}}{\sqrt{\frac{(n_{\text{TAU}}-1)S^2_{\text{TAU(post-ORS)}} + (n_{\text{Feedback}}-1)S^2_{\text{Feedback(post-ORS)}}}{N-2}}}$

** $p < .01$. *** $p < .001$.

ORS, patients in the feedback condition demonstrated significantly more improvement than TAU patients, on average 3.45 (γ_{01}) points higher. The effect size between treatment conditions after controlling for pre-ORS scores using Hedges' *g* (see the footnote to Table 2 for formula used) was .38. Regarding the PHQ-9, patients in the feedback condition did not demonstrate more improvement than patients in TAU. The effect size between treatment conditions after controlling for pre-PHQ-9 scores using Hedges' g was .06. The number of sessions attended was a statistically significant fixed effect for the ORS ($\gamma_{20} = .55$, $p < .01$) and PHQ-9 ($\gamma_{20} = -.30, p < .01$).

Table 3 reveals that for both measures, a chi-square analysis for clinically significant change showed significant advantages in the final sample for feedback over TAU (ORS: $\chi^2[1, N = 211] = 12.03, p < .01$; PHQ-9: $\chi^2(1, N = 204) = 4.45, p = .03$). We evaluated the rates of premature termination using the ORS (post-ORS score < 25) for the ITT and final samples. The rates were significantly lower for the feedback condition compared with TAU in the ITT sample, $\chi^2(1, N = 282) = 13.05$, $p < .01$, and the final sample, $\chi^2(1, N = 211) = 8.70$, $p < .01$. Patients in the feedback condition attended significantly more sessions than patients in the TAU condition in the ITT sample, $t(1,283) =$

Table 3
Clinically Significant Change Rates Per Treatment Condition for the Final Sample

		Feedback		TAU	
Measure	Clinically Sig.	*n*	%	*n*	%
ORS (*n* = 211)	Yes	65	55.1**	29	31.2
	No	53	44.9	64	68.8
PHQ-9 (*n* = 204)	Yes	36	32.4*	18	19.4
	No	75	67.6	75	80.6

* $p < .05$. ** $p < .01$.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

3.16, $p = .002$, as well as the final sample, $t(1, 209) = 2.19$, $p = .03$. As noted, 37 patients in TAU attended only one session compared with 18 in the feedback condition. This difference is statistically significant, $\chi^2(1, N = 55) = 6.56$, $p = .01$.

Discussion

This study investigated whether an evidence-based MBC system used in behavioral health settings could improve psychotherapy outcomes in integrated care. Consistent with our first hypothesis, the feedback condition demonstrated significantly larger treatment gains compared with TAU as measured by the ORS. PHQ-9 scores, however, did not indicate a difference between TAU and the feedback condition.

Our second hypothesis was supported by both measures. The percentage of patients reaching clinically significant change using ORS scores in the final sample was 55.1% in the feedback condition versus 31.2% in the TAU group. Using PHQ-9 scores, the percentage of patients reaching clinically significant change in the final sample was 32.4% in the feedback condition versus 19.4% in the TAU group. Results were also consistent with our hypothesis regarding premature termination and number of sessions. Patients in the feedback condition dropped out less (47.5 v. 67.7%) and attended more sessions (4.76 v. 3.73) than TAU patients.

The differences between feedback and TAU on the ORS as well as the dropout and attendance findings corroborate the effects reported in other PCOMS trials (see Duncan & Reese, 2015) as well as a meta-analysis of PCOMS studies (Lambert et al., 2018). Our findings also highlight the noted cautions of fidelity, the use of only the intervention measure, and dose of treatment (Duncan & Sparks, 2020; Østergård et al., 2020). Adherence may be particularly important to the PCOMS feedback effect (Duncan & Sparks, 2019; Peterson et al., 2018). Although the adherence scale did not differentiate the BHPs, it likely provided encouragement to follow the protocol. In addition, the BHPs received weekly supervision that used ORS data to address nonresponding patients and encouraged BHPs to follow the protocol.

MBC, as intended, is not only to monitor treatment progress but also to guide decisions while providing relevant information to multiple levels of stakeholders (Kearney et al., 2015). Making data-based decisions while involving patients, basic tenets of MBC and the Collaborative Care model (CoCM; Ivbijaro et al., 2014), are perhaps the most difficult aspects to implement in the workflow demands of daily practice. Measuring outcomes, in other words, is no guarantee that providers pay attention to the data. Although PCOMS has demonstrated successful application in real-world outpatient and inpatient behavioral health settings (Reese et al., 2014, 2018), implementation with the special challenges of integrated care has not been explored. Future research efforts, both quantitative and qualitative, may shed light on the challenges of implementation as well as both clinician and patient experience of MBC.

Regarding the use of the intervention measure only, several studies have reported corroborating effects on measures other than the ORS (Anker et al., 2009; Brattland et al., 2018; Cooper et al., 2021; Schuman et al., 2015) and some have not (Davidsen et al., 2017; Janse et al., 2017; van Oenen et al., 2016). We did not find corresponding changes on PHQ-9 scores but there were significant differences in favor of PCOMS on the percentage of clinically significant change. Finally, the experimental condition in this study did surpass the proposed four-session threshold for positive results.

There are several possible explanations for our mixed results. First is that the feedback effect on the ORS does not lead to similar changes measured by the PHQ-9. The positive finding on the PHQ-9 regarding clinically significant change seems to refute this possibility. A related possibility is that the feedback group only received information from the ORS and not the PHQ-9, which resulted in the stronger findings on the ORS. Second is that 44.8% of the patients fell into the nonclinical range on the PHQ-9 (v. 22.1% on the ORS), leaving little room for measuring change. Third is that the dose of the intervention was not sufficient to realize gains on PHQ-9 scores. Finally, and related, the PHQ-9 may not be as sensitive to change in problems unrelated to depression. As a symptom-based scale, the PHQ-9 tends not to identify or measure other areas of distress (e.g., relational, social) as suggested in our data as well as a study examining the ORS and the PHQ-9 as screeners in primary care (DeSantis et al., 2017).

There are several limitations to our study. First is the high percentage of nonclinical patients included in the PHQ-9 sample. The relatively small sample represents another limitation. Finally, our lack of formal adherence assessment to the PCOMS protocol is a significant limitation. Although we solicited BHP responses to a self-report adherence scale,

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

they were uniformly scored as if the protocol was followed.

Until this study, it was uncertain whether the feedback effect could be replicated in integrated care. PCOMS is a widely researched MBC system with a database of more than 1.5 million administrations that includes detailed protocols for practice, supervision, and implementation. Because its measures are ultrabrief, PCOMS seems well suited for the workflow demands of integrated care and already meets most if not all of Lewis et al.'s (2019) proposed research agenda to improve the integration of MBC into clinical practice. Embodying the principles of the CoCM, PCOMS could help accomplish the triple aim of improved outcomes, decreased cost, and enhanced patient experience (Berwick et al., 2008). Although preliminary and not without mixed results, this study offers some evidence, indicated by both the intervention measure (ORS) and the PHQ-9, that the improved outcomes, increased attendance, and reduced dropouts associated with MBC in mental health venues may also occur in integrated care settings.

References

Anker, M. G., Duncan, B. L., & Sparks, J. A. (2009). Using client feedback to improve couple therapy outcomes: A randomized clinical trial in a naturalistic setting. *Journal of Consulting and Clinical Psychology*, *77*(4), 693–704. https://doi.org/10.1037/a0016062

Berwick, D. M., Nolan, T. W., & Whittington, J. (2008). The triple aim: Care, health, and cost. *Health Affairs*, *27*(3), 759–769. https://doi.org/10.1377/hlthaff.27.3.759

Brattland, H., Koksvik, J. M., Burkeland, O., Gråwe, R. W., Klöckner, C., Linaker, O. M., Ryum, T., Wampold, B., Lara-Cabrera, M. L., & Iversen, V. C. (2018). The effects of Routine Outcomes Monitoring (ROM) on therapy outcome in the course of an implementation process: A randomized clinical trial. *Journal of Counseling Psychology*, *65*(5), 641–652. https://doi.org/10.1037/cou0000286

Bringhurst, D. L., Watson, C. W., Miller, S. D., & Duncan, B. L. (2006). The reliability and validity of the Outcome Rating Scale: A replication study of a brief clinical measure. *Journal of Brief Therapy*, *5*(1), 23–30.

Cacciola, J. S., Alterman, A. I., Dephilippis, D., Drapkin, M. L., Valadez, C., Fala, N. C., Oslin, D., & McKay, J. R. (2013). Development and initial evaluation of the Brief Addiction Monitor (BAM). *Journal of Substance Abuse Treatment*, *44*(3), 256–263. https://doi.org/10.1016/j.jsat.2012.07.013

Campbell, A., & Hemsley, S. (2009). Outcome rating scale and session rating scale in psychological practice: Clinical utility of ultra-brief measures. *Clinical Psychologist*, *13*(1), 1–9. https://doi.org/10.1080/13284200802676 3

Cooper, M., Duncan, B., Golden, S., & Toth, K. (2021). Systematic client feedback in therapy for children with psychological difficulties: Pilot cluster randomised controlled trial. *Counselling Psychology Quarterly*, *34*(1), Advance online publication. https://doi.org/10.1080/09515070.2019.1647142

Davidsen, A. H., Poulsen, S., Lindschou, J., Winkel, P., Tróndarson, M. F., Waaddegaard, M., & Lau, M. (2017). Feedback in group psychotherapy for eating disorders: A randomized clinical trial. *Journal of Consulting and Clinical Psychology*, *85*(5), 484–494. https://doi.org/10.1037/ccp0000173

DeSantis, B., Jackson, M., Duncan, B., & Reese, R. (2017). Casting a wider net in behavioral health screening in primary care: A preliminary study of the Outcome Rating Scale. *Primary Health Care Research and Development*, *18*(2), 188–193.

Duncan, B. (2011). *PCOMS Provider Adherence Scale*.

Duncan, B. L. (2014). *On becoming a better therapist: Evidence based practice one client at a time* (2nd ed.). American Psychological Association.

Duncan, B., Miller, S., Sparks, J., Claud, D., Reynolds, L., Brown, J., & Johnson, L. D. (2003). The Session Rating Scale: Preliminary psychometric properties of a "working" alliance measure. *Journal of Brief Therapy*, *3*(1), 3–12.

Duncan, B. L., & Reese, R. J. (2015). The partners for change outcome management system (PCOMS) revisiting the client's frame of reference. *Psychotherapy: Theory, Research & Practice*. *52*(4), 391–401. https://doi.org/10.1037/pst0000026

Duncan, B. L., & Sparks, J. A. (2002). *Heroic clients, heroic agencies: Partners for change*. Author.

Duncan, B., & Sparks, J. (2018). *The Partners for Change Outcome Management System: An integrated elearning manual for everything PCOMS*. Author. https://www.betteroutcomesnow.com

Duncan, B., & Sparks, J. (2020). When meta-analysis misleads: A critical case study of a meta-analysis of client feedback. *Psychological Services*, *17*(4), 487–496. https://doi.org/10.1037/ser0000398

Fortney, J. C., Unützer, J., Wrenn, G., Pyne, J. M., Smith, G. R., Schoenbaum, M., & Harbin, H. T. (2017). A tipping point for measurement-based care. *Psychiatric Services*, *68*(2), 179–188. https://doi.org/10.1176/appi.ps.201500439

Ivbijaro, G. O., Enum, Y., Khan, A. A., Lam, S. S., & Gabzdyl, A. (2014). Collaborative care: Models for treatment of patients with complex medical-psychiatric conditions. *Current Psychiatry Reports*, *16*(11), 506. https://doi.org/10.1007/s11920-014-0506-4

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Janse, P. D., de Jong, K., van Dijk, M. K., Hutschemaekers, G. J. M., & Verbraak, M. J. P. M. (2017). Improving the efficiency of cognitive-behavioural therapy by using formal client feedback. *Psychotherapy Research*, *27*(5), 525–538. https://doi.org/10.1080/10503307.2016.1152408

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*(4), 213–227. https://doi.org/10.1007/s10880-015-9443-6

King, P. R., Beehler, B. P., Wade, M., Buchholz, L. J., Funderburk, J., Lilienthal, K. R., & Vair, C. L. (2018). Opportunities to improve measurement-based care practices in mental health care systems: A case example of electronic mental health screening and measurement. *Families, Systems & Health*, *36*(4), 427–438. https://doi.org/10.1037/fsh0000379

Kroenke, K., & Spitzer, R. L. (2002). The phq-9: A new depression diagnostic and severity measure. *Psychiatric Annals*, *32*(9), 509–515. https://doi.org/10.3928/0048-5713-20020901-06

Lambert, M. J. (2015). Progress feedback and the OQ-system: The past and the future *Psychotherapy: Theory, Research, & Practice*, *52*(4), 381–390. https://doi.org/10.1037/pst0000027

Lambert, M. J., Whipple, J. L., & Kleinstäuber, M. (2018). Collecting and delivering Progress feedback: A meta-analysis of routine outcome monitoring. *Psychotherapy: Theory, Research, & Practice*, *55*(4), 520–537. https://doi.org/10.1037/pst0000167

Lewis, C. C., Boyd, M., Puspitasari, A., Navarro, E., Howard, J., Kassab, H., Hoffman, M., Scott, K., Lyon, A., Douglas, S., Simon, G., & Kroenke, K. (2019). Implementing measurement-based care in behavioral health: A review. *JAMA Psychiatry*, *76*(3), 324–335. https://doi.org/10.1001/jamapsychiatry.2018.3329

Miller, S. D., Duncan, B. L., Brown, J., Sparks, J., & Claud, D. (2003). The outcome rating scale: A preliminary study of the reliability, validity, and feasibility of a brief visual analog measure. *Journal of Brief Therapy*, *2*(2), 91–100.

Østergård, O. K., Randa, H., & Hougaard, E. (2020). The effect of using the Partners for Change Outcome Management System as feedback tool in psychotherapy—A systematic review and meta-analysis. *Psychotherapy Research*, *30*(2), 195–212 https://doi.org/10.1080/10503307.2018.1517949

Peterson, K., Anderson, J., & Bourne, D. (2018). Evidence brief: Use of patient reported outcome measures for measurement-based care in mental health shared decision-making. *Veterans Administration Evidence-Based Synthesis Program, Project, #09-199*.

Raudenbush, S. W., & Byrk, A. S. (2002). *Hierarchical linear models* (2nd ed.). SAGE.

Reese, R. J., Duncan, B. L., Bohanske, R. T., Owen, J. J., & Minami, T. (2014). Benchmarking outcomes in a public behavioral health setting: Feedback as a quality improvement strategy. *Journal of Consulting and Clinical Psychology*, *82*(4), 731–742. https://doi.org/10.1037/a0036915

Reese, R. J., Duncan, B. L., Kodet, J., Brown, H., Meiller, C., Farook, M., Lengerich, A., Vasilj, I., Hong, S., & Bohanske, R. (2018). Patient feedback as a quality improvement strategy in an acute care, inpatient unit: An investigation of outcome and readmission rates. *Psychological Services*, *15*(4), 470–476. https://doi.org/10.1037/ser0000163

Reese, R. J., Norsworthy, L. A., & Rowlands, S. R. (2009). Does a continuous feedback system improve psychotherapy outcome? *Psychotherapy: Theory, Research, & Practice*, *46*(4), 418–431. https://doi.org/10.1037/a0017901

Reese, R. J., Toland, M. D., Slone, N. C., & Norsworthy, L. A. (2010). Effect of client feedback on couple psychotherapy outcomes. *Psychotherapy: Theory, Research, & Practice*, *47*(4), 616–630. https://doi.org/10.1037/a0021182

Schuman, D. L., Slone, N. C., Reese, R. J., & Duncan, B. L. (2015). Using client feedback to improve outcomes in group psychotherapy with soldiers referred for substance abuse treatment. *Psychotherapy Research*, *25*(4), 396–407. https://doi.org/10.1080/10503307.2014.900875

She, Z., Duncan, B. L., Reese, R. J., Sun, Q., Shi, Y., Jiang, G., Wu, C., & Clements, A. L. (2018). Client feedback in China: A randomized clinical trial in a college counseling center. *Journal of Counseling Psychology*, *65*(6), 727–737. https://doi.org/10.1037/cou0000300

Slone, N. C., Reese, R. J., Mathews-Duvall, S., & Kodet, J. (2015). Evaluating the efficacy of client feedback in group psychotherapy. *Group Dynamics*, *19*(2), 122–136. https://doi.org/10.1037/gdn0000026

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder—The GAD-7. *Archives of Internal Medicine*, *166*(10), 1092–1097. https://doi.org/10.1001/archinte.166.10.1092

Van Oenen, F. J., Schipper, S., van, R., Schoevers, R., Visch, I., Peen, J., & Dekker, J. (2016). Feedback-informed treatment in emergency psychiatry; a randomised controlled trial. *BMC Psychiatry*, *16*, 110. https://doi.org/10.1186/s12888-016-0811-z

Received August 26, 2020
Revision received February 11, 2021
Accepted February 28, 2021 ■

This document is copyrighted by the American Psychological Association or one of its allied publishers.
This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.