

Outcomes of Technology-Enabled Psychotherapy Practice at Scale

Nicholas R. Forand, Jasmine Nettiksimmons, Margaret Anton, Raven Truxson,
Karl Vanderwood, Brandn Green

Submitted to: JMIR Formative Research
on: April 19, 2025

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript.....	5
Supplementary Files.....	41

Preprint
JMIR Publications

Outcomes of Technology-Enabled Psychotherapy Practice at Scale

Nicholas R. Forand^{1*} PhD, ABPP; Jasmine Nettiksimmons^{1*} PhD, MA, AB; Margaret Anton^{1*} PhD, BA; Raven Truxson^{2*} MPH, BS; Karl Vanderwood^{2*} PhD, MPH, BS; Brandn Green^{2*} PhD, MDiv, BA

¹ Two Chairs San Francisco US

² JG Research & Evaluation Bozeman US

*these authors contributed equally

Corresponding Author:

Jasmine Nettiksimmons PhD, MA, AB

Two Chairs
500 Davis St
San Francisco
US

Abstract

Background: Mental health conditions account for significant distress, burden, and societal costs. Despite efforts to implement evidence-based practices, access to high quality mental health treatment in general practice remains limited, and clinical outcomes sub-optimal. Measurement-based care (MBC) is a transtheoretical and transdiagnostic strategy that has the potential, when implemented effectively, to improve the quality of care. Digital tools can also support clinicians by alleviating administrative tasks and providing in-the-moment performance data and clinical decision support. In this study, we examine the clinical outcomes of a technology-enabled psychotherapy practice, where clinicians are supported by a suite of innovations including an MBC platform, clinical decision support tools, and tools designed to alleviate administrative burden.

Objective: The current study examines client retention and depression and anxiety outcomes within a technology enabled psychotherapy practice.

Methods: This retrospective cohort study examines 2,984 adults who initiated mental health treatment with Two Chairs, a hybrid technology enabled behavioral health provider, between January 1 to June 30, 2024. Rates of reliable change, recovery, remission, and magnitude and trajectory of symptom change in depression and anxiety symptoms were assessed using the PHQ-9 and GAD-7.

Results: The population demonstrated high rates of retention in care (89.9%), as well as high rates of MBC survey completion (96.3%). From baseline to the 12th session, patients showed significant symptom improvements in depression and anxiety, achieving high rates of reliable improvement (65.8%) and recovery (53.2%). Aggregate clinical outcomes continued to improve up to the point of termination. Pre to post-treatment effect sizes were large (all Cohen's d's > 0.9).

Conclusions: This study demonstrates how technology-enabled measurement-based care and clinical decision support systems may drive high quality patient outcomes in mental health. Implications for healthcare costs and value-based payment models are discussed.

(JMIR Preprints 19/04/2025:76264)

DOI: <https://doi.org/10.2196/preprints.76264>

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.

Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain visible to all users.
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <http://www.jmir.org/preprint/76264>, the full manuscript will be available to all users.
No. Please do not make my accepted manuscript PDF available to anyone. I understand that if I later pay to participate in <https://www.jmir.org/preprint/76264>, the full manuscript will be available to all users.



Original Manuscript

Outcomes of Technology-Enabled Psychotherapy Practice at Scale

Authors: Nicholas R. Forand^{1,2}, Jasmine Nettiksimmons¹, Margaret Anton¹, Raven Truxson³, Karl Vanderwood,³ Brandn Green³

¹Two Chairs

²Donald and Barbara Zucker School of Medicine at Hofstra/Northwell

³JG Research and Evaluation

Abstract

Background:

Mental health conditions account for significant distress, burden, and societal costs. Despite efforts to implement evidence-based practices, access to high quality mental health treatment in general practice remains limited, and clinical outcomes sub-optimal. Measurement-based care (MBC) is a transtheoretical and transdiagnostic strategy that has the potential, when implemented effectively, to improve the quality of care. Digital tools can also support clinicians by alleviating administrative tasks and providing in-the-moment performance data and clinical decision support. In this study, we examine the clinical outcomes of a technology-enabled psychotherapy practice, where clinicians are supported by a suite of innovations including an MBC platform, clinical decision support tools, and tools designed to alleviate administrative

burden.

Objective:

The current study examines client retention and depression and anxiety outcomes within a technology enabled psychotherapy practice.

Methods:

This retrospective cohort study examines 2,984 adults who initiated mental health treatment with Two Chairs, a hybrid technology enabled behavioral health provider, between January 1 to June 30, 2024. Rates of reliable change, recovery, remission, and magnitude and trajectory of symptom change in depression and anxiety symptoms were assessed using the PHQ-9 and GAD-7.

Results:

The population demonstrated high rates of retention in care (89.9%), as well as high rates of MBC survey completion (96.3%). From baseline to the 12th session, patients showed significant symptom improvements in depression and anxiety, achieving high rates of reliable improvement (65.8%) and recovery (53.2%). Aggregate clinical outcomes continued to improve up to the point of termination. Pre to post-treatment effect sizes were large (all Cohen's d 's > 0.9).

Conclusions:

This study demonstrates how technology-enabled measurement-based care and clinical decision support systems may drive high quality patient outcomes in mental health. Implications for healthcare costs and value-based payment models are discussed.

Keywords:

Depression; anxiety; measurement-based care; psychotherapy

Introduction

Mental health conditions are a significant public health concern, contributing substantially to the

global burden of disease.(1) The increased prevalence of mental health disorders has a clear relationship on disability, reduced quality of life, and associated increases in avoidable economic costs, including costs incurred by payers while managing the care of individuals with these conditions.(2,3) The economic burden of mental illness is profound, affecting both direct healthcare expenditures and indirect costs connected to loss of productivity in the workforce and increased demand for social supports. An analysis of the change in the economic burden of major depressive disorder between 2010 and 2018 describes a global increase from \$236.6 billion to \$326.2 billion.(4) These broad and significant personal, social, and economic implications underscore the urgency of addressing mental health challenges.(5) There are two core opportunities for improving care, specifically improved access to care(6) and improved quality of care.(7) These opportunities may be addressed through measurement-based care models that leverage technology-based innovations to support high-quality and accessible therapeutic care.

Access to quality care

Despite a substantial increase in the percentage of US adults receiving mental health services over the past 25 years, access to effective, high quality mental health care remains problematic.(5) Studies examining early termination of psychotherapy treatment suggest that a significant portion of patients drop out of therapy prematurely.(6,8) One study found that 73.5% of patients in a health system setting dropped out of care prior to their fourth session, suggesting that the majority of patients received an inadequate dose of treatment.(8) High levels of treatment dropout have been attributed to low therapeutic alliance and dissatisfaction with the quality of care.(8,9)

A systematic review of usual care practices highlights that many mental health interventions also lack a structured, evidence-based approach, leading to inconsistent treatment delivery and variable patient outcomes.(10) Outcomes in regular care practices reflect these problems, with rates of

improvement being up to 20-30% lower in these settings than in clinical trials of structured, evidence based interventions.(12–14)

Although there has been a growing recognition of mental illness as a critical health issue, the quality of care provided to individuals with mental disorders remains suboptimal due to limited quality standards, weak therapeutic alliance, and underutilization of evidence-based approaches, all of which further contribute to the burden on individuals, families and healthcare systems even among those who initiate care.

Barriers to effectively implementing evidence-based practices

Although evidence based practices (EBPs) are known to improve patient outcomes, the implementation of these practices in mental health care can be hindered by cost, complexity, and challenges in sustainment of the program model over time.(15,16) While standard models of scaling EBPs can be effective in the short-to-medium term, they often require significant financial investment and logistical coordination, limiting their long-term viability.(15) The implementation science literature has demonstrated that achieving meaningful improvements in mental health care through EBPs can take years.(17) Several real-world examples indicate that successful implementations can take from three to seven years before measurable positive outcomes are observed, underscoring the prolonged timeline required for systems wide change.(17–19) Further, given that many EBP's focus on only a narrow set of patients or conditions (e.g., CBT for patients diagnosed with panic disorder), they often leave gaps in training for clinicians or have circumscribed impact on patient outcomes.

In light of the need for meaningful, sustained engagement to a care model, the durability of programs

is a real challenge to ensuring access to quality care over time.(20,21) Without ongoing fidelity monitoring, supervision, and institutional support, the effectiveness of these interventions can wane. (21)

Potential of measurement-based care

Measurement-based care (MBC), sometimes referred to as routine outcome monitoring or feedback informed care, is an evidence-based practice that includes the collection, sharing, and use of patient-reported data to collaboratively guide clinical decision-making.(22,23) When implemented effectively, it has been demonstrated to improve clinical outcomes.(24) Unlike other evidence-based mental health interventions which are often designed to target a specific clinical diagnosis or concerns in targeted populations, MBC is both transtheoretical and transdiagnostic. MBC implementation is therefore not as susceptible to some of the common implementation and scalability challenges of other evidence-based practices which may only be applicable to a small number of patients and/or clinicians. Research indicates that MBC enhances therapy outcomes by enabling clinicians to identify and address care delivery and alliance challenges in real time, leading to reduced dropout,(25) improved therapeutic alliance, faster symptom resolution, and a more personalized, effective approach than standard care.(23) Because MBC is both transtheoretical and transdiagnostic, it may be a valuable intervention for addressing both limited access to care and limited access to quality care at scale.(24)

Despite these advantages, implementation of MBC is not without its challenges. At the clinician level, time and administrative burden associated with selecting, delivering, scoring and tracking scores over time may inhibit uptake.(26) Without support, clinicians may also struggle with generating insights from MBC data or tracking and identifying trends. Patients presenting with complex or unique presentations have been identified by clinicians as additional barriers when

considering implementation of MBC.(26) At the organizational level, leadership alignment, funding, implementation of new technology, training, and supervision are cited as limitations to adoption.(26) Barriers to implementation of MBC may be reflected in the rates of clinicians who use these tools. Fewer than 20% of clinicians indicate using MBC in their practice,(24) and even when measures are mandated by the clinical service, rates of completion can be as low as 15%.(27)

Technology as a solution to MBC barriers

Clinical decision support systems (CDSS) and other clinical dashboard applications offer a technology-enabled solution to address many of the identified challenges associated with successfully implementing MBC.(22,23) CDSS is a form of technology supporting care that can be integrated into healthcare settings to improve patient outcomes and adjust clinician burdens by providing clinicians with evidence-based recommendations at the point of care.(24) CDSS systems analyze patient data and suggest appropriate interventions and follow-up, thereby enhancing clinical efficiency and reducing clinical errors.(25) By integrating these systems, CDSS clinicians are able to make informed decisions that align with the latest clinical guidelines leading to improved patient outcomes.(24)

Technology can also alleviate administrative tasks, allowing clinicians to focus more on patient care. (26) Additionally, digital tools can facilitate real-time performance feedback, enhancing visibility and accountability across all levels of care.(27) Clinicians can monitor patient progress, symptom changes and treatment responses.(28) In some instances, patients can also track their own progress, empowering them in their own care.(29) Systematic collection and feedback of data can also be leveraged for quality improvement efforts, such as monitoring clinician or practice performance and identifying areas for remediation or professional development.(28) Across this broad set of supports,

technological innovations in care delivery, patient tracking, and analysis have demonstrated capacity to increase sustainability of services through enhanced quality and patient engagement, while also decreasing administrative burdens associated with documenting the provision of care.

The Current Study

Two of the core challenges perpetuating the mental health care crisis in the United States are access to care and the quality of care. Implementation of EBP's are complex and may only meet the needs of a subset of patients. MBC and CDSS offer promise to improve the scalability of high-quality care in general practice; however, there have been few "real world" demonstrations of the feasibility or effectiveness of this approach. The aim of the current study is to describe the following features of a technology-enabled psychotherapy practice supported by MBC and CDSS: 1) patient retention; 2) patient MBC completion; and 3) clinical outcomes, including depression and anxiety symptom reduction among patients who engaged in weekly, outpatient psychotherapy, either virtual or in-person.

Methods

Study design

This is a retrospective cohort analysis focused on adult patients who enrolled in mental health treatment with Two Chairs in the first half of 2024. All study procedures were submitted to the Sterling institutional review board, Atlanta, Georgia, United States, and deemed exempt.

Study population

Patients were enrolled in behavioral health services at Two Chairs, a technology-enabled behavioral health company founded in 2017 that provides psychotherapy to self-pay, commercially insured, and

publicly insured individuals. During the period of data collection, services were available in California, Washington, and Florida. Patients served by Two Chairs are aged ≥ 18 or older and receive outpatient psychotherapy for anxiety, depression, and related conditions.

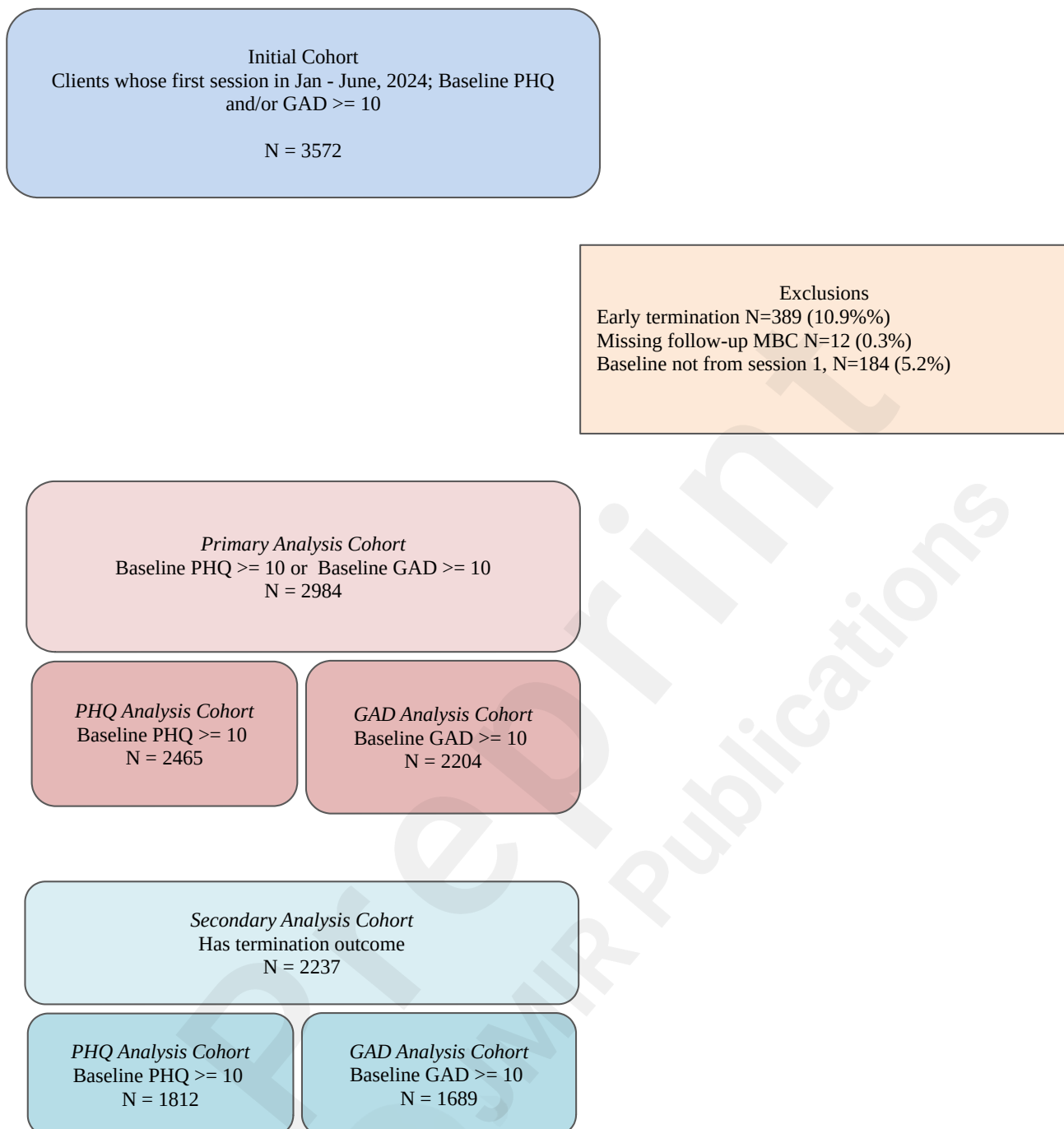
At intake patients complete a digital assessment that includes the PHQ-9 and GAD-7. Patients then attend an initial clinical interview, after which they are assigned a preliminary diagnosis and subsequently matched with an individual clinician. During onboarding 98% of all patients are screened for anxiety and depression and 45% and 48% of clients screen positive, respectively. Of those with a positive screen 84% started care within 30 days. Of clients who receive a diagnosis of a depressive disorder, 99% have at least one follow-up PHQ-9 assessment and of those who receive a diagnosis of anxiety disorder, 99% have at least one follow-up GAD-7 assessment.

Initial eligibility criteria for the study required patients to have initiated mental health treatment with Two Chairs with a first treatment session between January 1 and June 30, 2024 and a baseline PHQ-9 or GAD-7 score of 10 or greater. Patients reporting active suicidal ideation or otherwise deemed inappropriate for outpatient psychotherapy at intake were referred to health plan or community resources and not included in this study population. The cutoff for data included in this study is February 28, 2025, regardless of whether or not a patient had completed their episode of care by that date. A total of 3,572 patients met the initial eligibility criteria.

Patients were excluded from the primary analytic cohort if they terminated from care early (defined as fewer than four sessions), were missing MBC assessment at session 1, or lacked any MBC assessment follow-up between sessions 4 and 12 (see Figure 1). After these criteria were applied, the primary analytic cohort was $n=2,984$. Within the primary cohort, inclusion criteria for

PHQ-9 outcomes (baseline PHQ-9 ≥ 10) was met for $n=2,465$, and inclusion criteria for GAD-7 outcomes was met for $n=2,204$ (baseline GAD-7 ≥ 10). A secondary cohort included a subset of patients from the primary cohort who had completed their course of care during this timeframe ($n=2,237$; PHQ-9 $n=1,812$; GAD-7 $n=1,689$). See Figure 1.

Figure 1: Participant flow.



Study interventions

Patients enrolled in Two Chairs receive weekly 45-55 minute teletherapy or in-person mental health services delivered by licensed clinicians. Clinicians use a range of treatment modalities to address

patients' needs, including but not limited to cognitive behavioral therapy, acceptance and commitment therapy, time limited dynamic psychotherapy, and somatic therapy. As part of their practice at Two Chairs, clinicians are required to use MBC to evaluate patients' symptom presentation, make care decisions, and track progress in care. Prior to each session, patients complete measures of depression (PHQ-9) and anxiety (GAD-7) symptoms, mental health quality of life, and therapeutic alliance. For these analyses, only the PHQ-9 and GAD-7 results are reported. Clinicians use these measures to inform care and are trained to discuss the data with patients to inform collaborative decision-making around treatment progress.

The technological supports offered to support MBC and CDSS by Two Chairs include: 1) automated electronic questionnaire delivery to patients; 2) automated scoring and interpretation of patient-reported symptoms over time; 3) clinician alerting for potential safety and risk concerns like endorsement of suicidal ideation; and 4) clinical decision support features, including indications of clinically meaningful symptom changes or changes in alliance, standardized recommendations for clinician actions, alerts related to patient-endorsed suicidal ideation, and holistic summaries of change across all measures (see Figures 2-5). Patients also provided a summary of their symptoms with highlights of meaningful symptom change to allow them to reflect on their care progress and feel empowered to engage in care decisions (see Figure 5). Clinicians have access to patient goals, diagnoses, and number of sessions on each patients' individual page within the platform.

Clinicians are further supported by digital tools that are intended to relieve administrative burden, such as calendar management and automated scheduling of new patients. Finally, clinicians have access to a number of performance metrics including their aggregate adherence to MBC surveys, patient attendance, and utilization. Clinical managers have a similar suite of tools that allows them to track the overall performance of their clinical teams.

Figure 2: Clinician Dashboard.

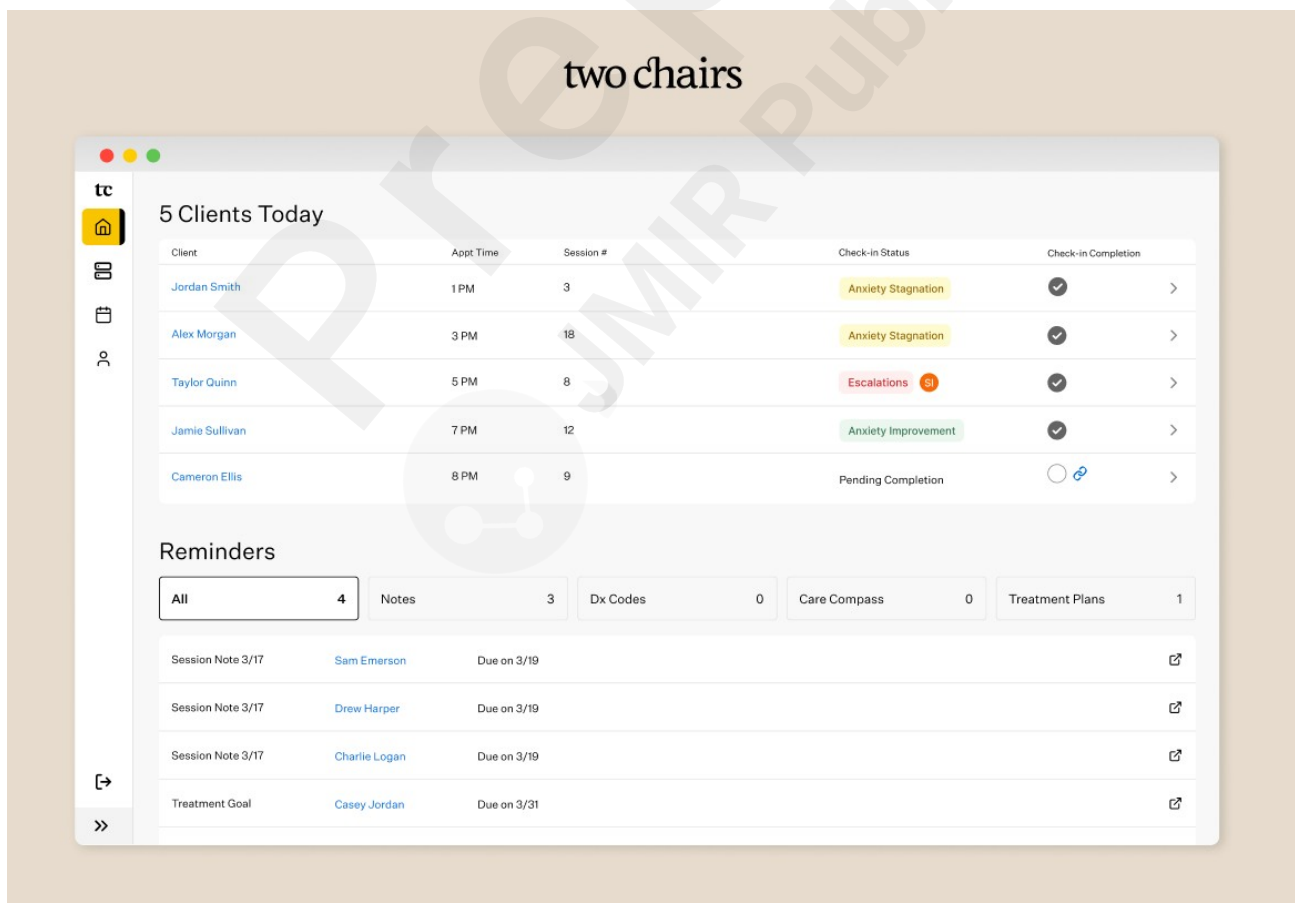


Figure 2. In this view, clinicians can view their upcoming patients for the day, their session number, and a summary of key changes on all four MBC assessments (PHQ9, GAD7, MHQoL, and alliance). They can also view reminders of any upcoming administrative tasks. Patient names are fabricated.

Figure 3: View of GAD-7 Scores on Patient Page.

two chairs

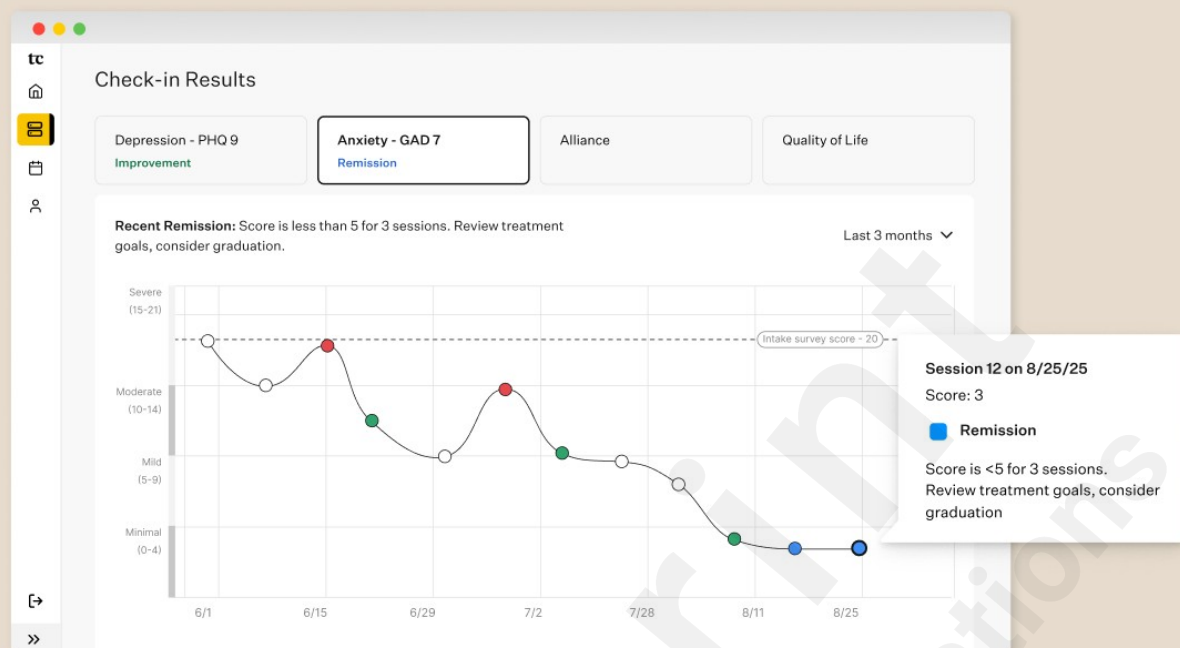


Figure 3. Graph of GAD-7 scores by session. Points are color-coded based on the magnitude of symptom changes week to week, e.g., a red dot equals reliable deterioration whereas a green dot equals reliable improvement. Hovering over a dot will show the score for that session and provide clinical decision support.

Figure 4: View of PHQ-9 Scores on Patient Page.

two chairs

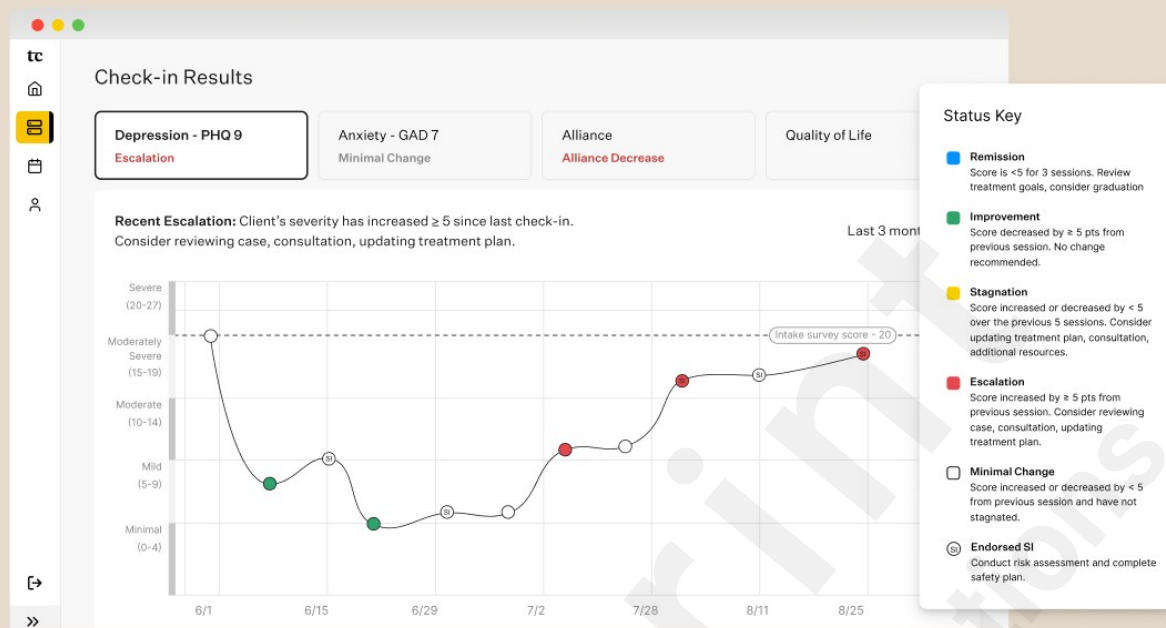


Figure 4. Graph of PHQ-9 scores by session. Points are color-coded based on the magnitude of symptom changes week to week, e.g., a red dot equals reliable deterioration whereas a green dot equals reliable improvement. Dots with SI in them indicate that a client has screened positive for suicidal ideation on the PHQ-9. In this view, the full key describing the color coding and decision support is displayed.

Figure 5: View of Patient-facing MBC Feedback and Clinician Tabular view.

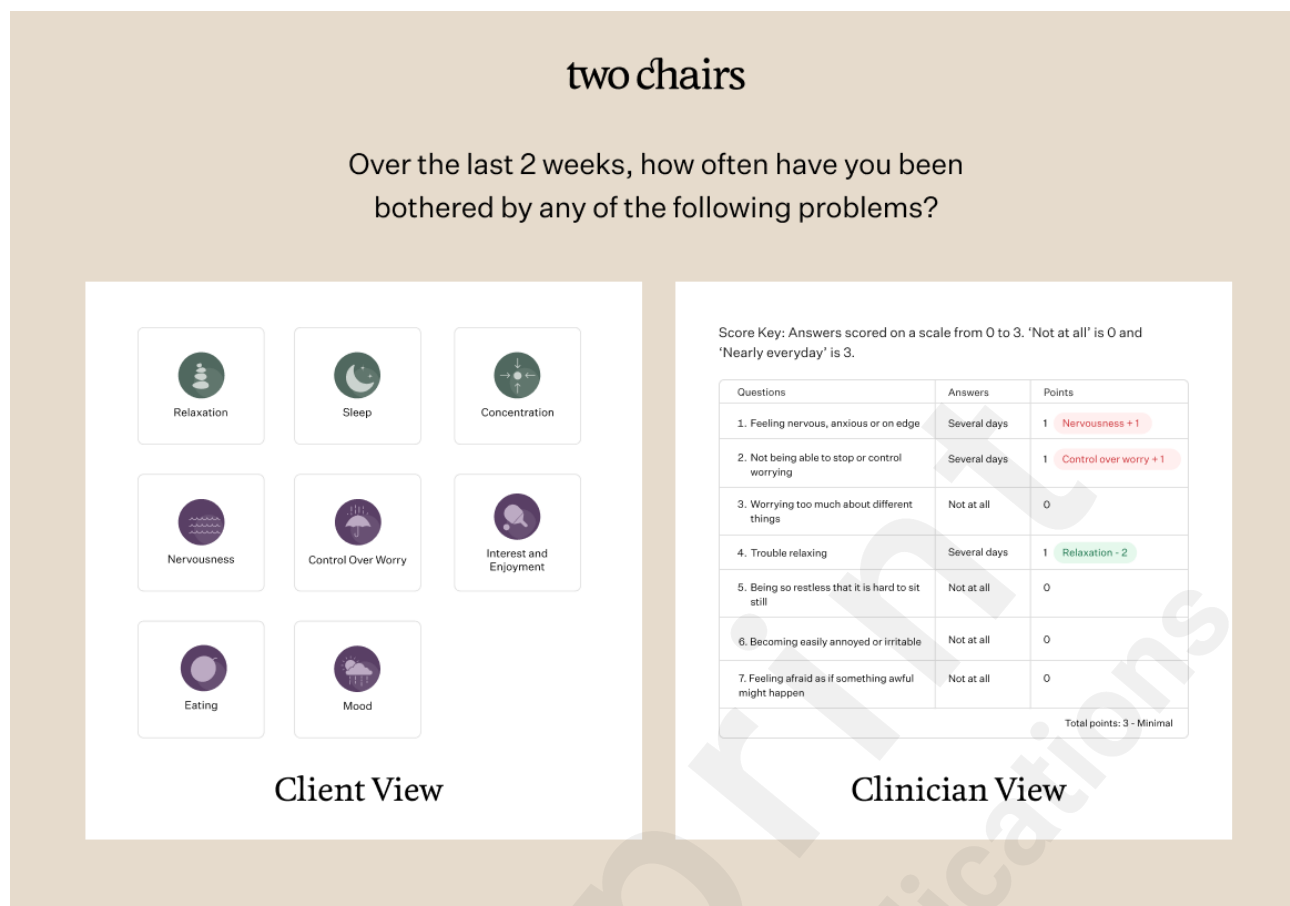


Figure 5. The left panel displays feedback on week to week symptom changes provided to patients after they complete their weekly MBC surveys. Green indicates symptoms improvements whereas purple indicates symptom worsening. The right panel shows a tabular view of the GAD-7 which provides complimentary information to the clinician on symptom level changes from the previous week.

Measurement

Patient demographics

As part of the onboarding process at Two Chairs, patients provided basic demographic information (e.g., age, race/ethnicity, gender, employment status). Data is collected as a part of routine care and patients are not required to complete all questions. As a result, data sparseness varies across variables (see Table 1).

Patient retention

Patient retention was evaluated as the proportion of patients who met the initial study inclusion criteria and were retained until session 4, qualifying for inclusion in the primary analytic cohort.

MBC completion

MBC completion was assessed as the proportion of observed sessions (up to 12) in which MBC assessments were completed.

Measures

Patients completed the *Patient Health Questionnaire-9* (PHQ-9)(30) and *Generalized Anxiety Disorder-7* (GAD-7)(31) at each session throughout care. The PHQ-9 is a 9-item self-report of depressive symptoms. Items are rated on a scale ranging from 0 (*not at all*) to 3 (*nearly every day*). Total scores range from 0 to 27, and cut-off scores for mild, moderate, moderately severe, and severe depressive symptoms are 5, 10, 15, and 20, respectively. Reliable change on the PHQ-9 was a change in score of 5 or greater.(32) The GAD-7 is a 7-item self-report measure of anxiety. Items are rated on a scale ranging from 0 (*not at all*) to 3 (*nearly every day*). Total scores range from 0 to 21, and the cut-off scores for mild, moderate, and severe anxiety symptoms are 5, 10, and 15, respectively. Reliable change on the GAD-7 was a score of 4 or greater.(33)

Clinical outcomes

For these analyses, three timepoints were defined: baseline (session 1), the clinical check-in session (session 12 or the last available MBC assessment from sessions 4-11), and at termination (last available measurement prior to termination) for patients who terminated care prior to February 28,

2025.

This study focused on characterizing depression and anxiety symptom improvement using both joint outcomes for the PHQ-9 and GAD-7 combined and independent outcomes for the PHQ-9 and GAD-7 individually. The goal of the joint measures was to holistically evaluate change in commonly co-morbid depression and anxiety,(34) similar to methods used in other settings where these measures are used concurrently.(35) Patients were included in the summaries of joint outcomes if either PHQ-9 and/or GAD-7 baseline scores ≥ 10 . For assessment-specific outcomes (PHQ-9 or GAD-7), inclusion required a baseline score ≥ 10 on the given assessment.

Independent outcomes

Assessment-specific outcomes were defined as follows: reliable improvement was defined as an improvement of 5pts on PHQ-9, 4pts on GAD-7, respectively; recovery was defined as a follow-up score < 10 ; remission was defined as a follow-up score < 5 ; and response was defined as a 50% or greater improvement in score from baseline.

Joint outcomes

For the joint measures, reliable improvement was defined as reliable improvement on *either* the PHQ-9 or GAD-7 without reliable deterioration on the other measure. Recovery was defined as *both* PHQ-9 or GAD-7 < 10 at the follow-up. Clinically meaningful improvement was defined as meeting either joint recovery or joint reliable improvement criteria. This holistic measure of change across both depression and anxiety encompasses both magnitude of change and symptom severity at the time of measurement.

Statistical analysis

The study population was characterized using means, standard deviations, and frequencies with statistical tests to evaluate differences between the primary cohort and the population who was excluded from analysis (t-tests, Chi-squared tests, Fisher exact tests for cell counts < 20) (Table 1).

Table 1. Patient Demographics and Baseline Characteristics of Patients Included and Excluded in the Primary Cohort

	Included <i>n</i> (%)	Excluded <i>n</i> (%)	
Characteristic	(<i>n</i> = 2984)	(<i>n</i> = 588)	<i>P</i> value
Age			
18-29	774 (25.9%)	134 (22.8%)	0.05
30-49	1549 (51.9%)	295 (50.2%)	
50-69	545 (18.3%)	127 (21.6%)	
70+	116 (3.9%)	32 (5.4%)	
Gender			
Woman	1793 (60.1%)	361 (61.4%)	0.006
Man	695 (23.3%)	123 (20.9%)	
Unknown	325 (10.9%)	85 (14.5%)	
Trans/nonbinary/other	171 (5.7%)	19 (3.2%)	
Race & Ethnicity			
Non-Hispanic White	1297 (43.5%)	231 (39.3%)	0.23
Hispanic	573 (19.2%)	118 (20.1%)	0.40
Unknown race/ethnicity	544 (18.2%)	141 (24.0%)	0.001
Asian	371 (12.4%)	59 (10.0%)	0.20
Black	273 (9.1%)	61 (10.4%)	0.26
Other race/ethnicity	250 (8.4%)	47 (8.0%)	1.00
American Indian/Alaska Native	80 (2.7%)	18 (3.1%)	0.49
Asian Pacific Islander	74 (1.5%)	9 (1.5%)	0.23
Primary Payor			
HMO	2316 (77.6%)	408 (69.4%)	<0.001
Public	478 (16.0%)	123 (20.9%)	
Commercial insurance	142 (4.8%)	38 (6.5%)	
Self-Pay	48 (1.6%)	19 (3.2%)	
Employment status			
Fulltime	1615 (54.1%)	293 (49.8%)	0.01
Other	774 (25.9%)	139 (23.6%)	
Unknown	315 (10.6%)	84 (14.3%)	
Disabled/Unemployed	280 (9.4%)	72 (12.2%)	
CCSR diagnosis categories			

Anxiety and Fear-Related Disorders	990 (33.2%)	178 (30.3%)	<0.001
Depressive Disorders	971 (32.5%)	176 (29.9%)	
Trauma- and Stressor-Related Disorders	422 (14.1%)	61 (10.4%)	
Adjustment Disorders	289 (9.7%)	62 (10.5%)	
Other or Primary Dx Unknowns	312 (10.5%)	111 (18.9%)	
Baseline Score	Mean (SD)	Mean (SD)	
PHQ-9	13.2 (4.6)	13.3 (5.1)	0.70
GAD-7	12 (4.2)	12.6 (4.3)	0.001

PHQ-9 and GAD-7 are described at the three time points (baseline, check-in session, and termination) with means and standard deviations. Change from baseline on PHQ-9 and GAD-7 is described by the mean difference, 95% confidence intervals, and Cohen's *d* (see Table 2). Summaries of dichotomous outcomes (reliable improvement, recovery, meaningful improvement, remission, and response) are described with frequencies and 95% confidence intervals (Table 3).

Longitudinal analysis of PHQ-9 and GAD-7 was performed using linear mixed effects models with random intercepts for clinicians and random intercepts and slopes (session number) for patients. Fixed effects included age, gender, session number, and a quadratic term for session number (Supplemental Table 1). Average effects across sessions were estimated with the most commonly occurring level of age and gender (Figure 6).

All analysis was performed using R version 4.4.2 (2024-10-31).

Results

Patient characteristics

Among the 3,572 patients who met the initial criteria for inclusion, 89.9% ($n=3,183$) were retained to the 4th session and thus eligible for the primary analytic cohort. A small number of patients were excluded due to missing follow-up MBC scores or because their session 1 baseline was unavailable, leaving 2,984 patients (83.5% of the population meeting the original inclusion criteria, Figure 1). For 71.4% of the included cohort, the clinical check-in session occurred at session 12 with an average

check-in session number of 10.8. A total of 2,237 patients from the primary cohort terminated care during the observation window and therefore have termination outcomes. During the study period, among the primary cohort, therapy was delivered by 510 clinicians, with a total of 32,348 sessions. Patients completed the Two Chairs recommended screening instruments at 96.3% of sessions.

Within the primary analytic cohort, the population was typically between the ages of 30 and 49, slightly more female, more commonly employed, and more commonly had an HMO as a payor source. Although 18% have an unknown race or ethnicity, the sample is notable for the racial diversity represented, with just 43% self-identifying as non-Hispanic White (Table 1). The excluded population was more likely to have unknown/unspecified gender, race, and employment status, all features collected from an optional patient profile, and was more likely to have public insurance and primary diagnoses other than anxiety disorders, depressive disorders, and adjustment disorders. Table 2 shows changes in depressive and anxiety symptoms from baseline to the clinical check-in session (for all individuals included in the primary cohort) and changes from baseline to termination (for individuals included in the secondary cohort). At both the check-in assessment and termination, Two Chairs patients show improvement with strong effect sizes across both cohorts and symptom measures.

Table 2. Change in MBC Scores from Baseline to Check-in (Primary Cohort) and Termination (Secondary Cohort).

	Baseline		Clinical Check-in Session			Termination Session			
	Mean		Mean	Change	Cohen's		Mean	Change	Cohen's
Measure	(SD)	N	(SD)	(95% CI)	d	N	(SD)	(95% CI)	d
PHQ-9	14.5 (3.8)	2465	9.5 (5.5)	-5 (-5.3,-4.8)	0.92	1812	7.9 (5.7)	-6.6 (-6.9,-	1.14
PHQ-9 % change				-33.9% (-35.4%,-	0.94			6.3) -45.0%,-	1.2
								46.8%,-	

				32.5%)				43.3%)	
				-5.1 (-5.3,-				-6.6 (-6.8,-	
GAD-7	13.8 (3.1)	2204	8.7 (5.1)	4.9	1.03	1689	7.2 (5.2)	6.3	1.25
GAD-7 %				-36.6% (-				-47.3% (-	
change				38.1%,-				49.0%,-	
				35.1%)	1.04			45.5%)	1.31

From baseline to the most recent check-in session, patients show on average a 5.0-point (Cohen's $d=0.92$), 34% (Cohen's $d=0.94$) decrease in PHQ-9 scores and a 5.1-point (Cohen's $d=1.03$), 37% (Cohen's $d=1.04$) decrease in GAD-7 scores. From baseline to termination, patients show on average a 6.6-point (Cohen's $d=1.14$), 45% (Cohen's $d=1.2$) decrease in PHQ-9 scores and a 6.6 (Cohen's $d=1.25$), 47% Cohen's $d=1.31$) decrease in GAD-7 scores.

Table 3 shows dichotomous outcomes at the clinical check-in session and termination session, and, where available, across assessment-specific and joint measures of achievement.

At the clinical check-in session (typically session 12), 71.3% of patients achieved clinically meaningful improvement, defined as joint recovery or joint reliable improvement. Among patients with a termination outcome available, 79.2% of patients experienced clinically meaningful improvement. A majority of patients who met the baseline criteria for PHQ-9 outcomes experienced reliable improvement (53.0%) and recovery (55.1%) on PHQ-9 at the clinical check-in. These numbers improved to 65.3% and 66.6%, respectively, among clients with available termination sessions. A majority of patients who met the baseline criteria for GAD-7 outcomes experienced reliable improvement (64.4%) and recovery (61.5%) on the GAD-7 measures at the clinical check-in session; both of these measures improved to over 70% at the termination session. Across all measures, outcomes at termination in the secondary cohort were higher than outcomes at the clinical check-in session for the primary cohort.

Table 3. Patient Response, Recovery and Remission Rates at Check-in (Primary Cohort) and Termination (Secondary Cohort).

		Clinical Check-in (Primary Cohort)		Termination (Secondary Cohort)	
	Measure	Outcomes (%)	95% CI	Outcomes (%)	95% CI
Joint Measures	Clinically Meaningful Improvement ^a	71.3	(69.7,72.9)	79.2	(77.5,80.9)
	Recovery ^b	53.2	(51.4,55.0)	65.2	(63.2,67.2)
	Reliable improvement ^c	65.8	(64.0,67.5)	74.8	(73.0,76.6)
	Remission	16.4	(15.0,17.7)	27.7	(25.9,29.6)
PHQ-9^d Measures	Reliable improvement	53.0	(51.1,55.0)	65.3	(63.1,67.5)
	Response	36.4	(34.5,38.3)	51.1	(48.8,53.4)
	Recovery	55.1	(53.2,57.1)	66.6	(64.4,68.7)
	Remission	19.5	(17.9,21.1)	32.3	(30.2,34.5)

GAD-7^e Measures	Reliable improvement	64.4	(62.4,66.4)	73.6	(71.5,75.7)
	Response	41.1	(39.0,43.1)	55.8	(53.4,58.1)
	Recovery	61.5	(59.4,63.5)	71.9	(69.7,74.0)
	Remission	20.8	(19.1,22.5)	34.0	(31.8,36.3)

^APatient met criteria for joint measures for Reliable improvement OR Recovery.

^bAt the clinical check-in session, both GAD and PHQ are < 10.

^cAt the clinical check-in session, the patient showed improvement of at least 5 points on PHQ OR 4 points on GAD without a decline of 5 or 4 points on the other measure, respectively.

^dIncludes patients with PHQ-9 $\geq$ 10 at baseline.

^eIncludes patients with GAD-7 $\geq$ 10 at baseline.

Longitudinal analysis of PHQ-9 and GAD-7 assessments at the session level estimated an approximately 1-point reduction in assessment score per attended session which is attenuated over time by approximately 0.05 points per session, accounting for the non-linear trajectory of change (see Supplemental Table 1). Figure 6 shows model-derived average estimates of PHQ-9 and GAD-7 across session number along with model parameter-based 95% confidence intervals for each estimated point. The figure demonstrates the swift early change in symptoms followed by an attenuated rate of change, with both curves dipping below the clinical threshold (score of 10) near 8 (PHQ-9) and 6 (GAD-7) sessions, respectively.

Figure 6. Model-derived estimates for average PHQ and GAD scores over time.

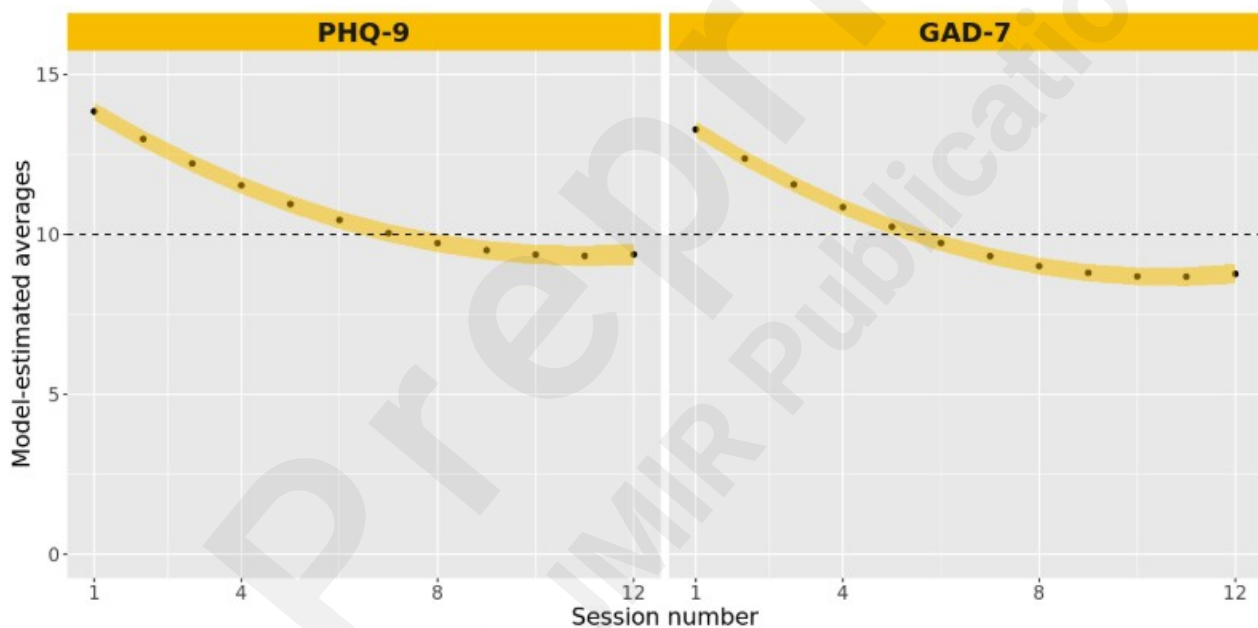


Figure 6. Covariates were assumed to be the most populous levels of age and gender. The yellow band reflects model-based 95% confidence intervals for each estimated point. The dotted line at 10 represents the clinical threshold for PHQ-9 and GAD-7.

Discussion

In this study, we examined clinical outcomes among a cohort of 2,984 patients who received

psychotherapy as delivered by Two Chairs clinicians through telehealth or in-person settings. The study was intended to characterize patient retention and outcomes in technology-enabled psychotherapy practice that utilizes measurement-based care and clinical decision support.

Our findings indicate that changes in depressive and anxiety symptoms over the course of care were meaningful. By the 12th session, 71% of patients achieved clinically meaningful improvement in their symptoms, and outcomes continued to improve up to the completion of care. On average, improvement in PHQ-9 and GAD-7 scores exceeded the threshold for reliable and clinically significant symptom improvement on these measures.(33,36) We also observed very high rates of patient retention in care and over 95% rates of completion for MBC surveys (e.g., PHQ-9, GAD-7), indicating high levels of engagement with the care mode by clinicians.

The results of this study compare favorably to benchmarks in general practice settings. Rates of dropout in this sample are nearly 6 times lower than in a standard healthcare setting (73.5% vs 10.1%),(37) and outcomes are 2 to 2.5 times better than observed samples of individuals who received usual care for anxiety and depression.(11,13)¹ Finally, adherence to MBC activities is substantially higher than in comparable settings.(24) These results suggest that the majority of individuals with elevated symptoms of depression and anxiety treated within this mode are receiving an effective dose of care and also receiving meaningful benefit, reflected in substantial reductions in clinical symptoms.

Although examined indirectly in this study, the features of the care model offered by Two Chairs suggest some explanations for the comparative differences in patient outcomes and retention. Evidence-based practices are known to be effective for improving overall outcomes in care; however,

1

the burden associated with implementation and sustainment of EBPs to fidelity likely contribute to their low uptake in clinical practice. Removing barriers to the implementation of measurement-based care via technology may account for some of the favorable patient outcomes observed in the current study. The Two Chairs platform provides clinicians with consistent, accessible, and patient-salient information. This set of clinician tools allows clinicians to more easily access and interpret information provided via MBC surveys and provides them with actionable insights that may help them focus on providing effective interventions for patients. These findings suggest, in a preliminary way, that the execution of these concepts in an integrated fashion is producing high quality patient experiences and improved outcomes. Additional research will be needed to isolate the specific impacts of each technological and practice element.

This study's findings have implications for both value-based payment models as well as standard reimbursement models of care. Psychotherapy is an effective treatment for a broad range of mental health conditions, and healthcare claims-based analyses suggest that outpatient mental health treatment can save overall healthcare costs compared to those who do not receive treatment at all. (37) Importantly, evidence suggests that cost savings improve further if individuals are retained in care past 1-2 sessions, indicating that the dose of treatment must be sufficient to produce clinically meaningful change.(37) Moreover, emerging evidence suggests that decreases in symptoms drive reductions in costs associated with mental health conditions, particularly when compared to those who stagnate or worsen over time.(4) Together, these findings highlight the importance of receiving an adequate dose of effective care.(4) Providing patients with the appropriate care, ensuring they remain engaged in care after initiation, and achieving meaningful outcomes can help reduce the overall cost and burden associated with mental health conditions.

Limitations

This study is not without limitations. The technology-based MBC clinical decision support systems

were examined in one organization rather than being examined across settings. As the improvements are process related, this may be an ongoing limitation. Second, the retrospective and non-experimental design limits our ability to examine causal pathways. Third, the Two Chairs care model supports integration of multiple types of EBPs, and this study did not look to distinguish differential patient outcomes across treatment programs. These limitations withstanding, the results do suggest a real impact on a substantial patient population.

Conclusion

The results of this retrospective cohort study suggest that an integrated platform supporting measurement based care, clinical decision support, and automation of clinically burdensome tasks can produce improvements in care engagement and clinical outcomes. Although additional research is needed to further clarify the influential pathways of the integrated model, these findings are strongly suggestive that the technology-enabled innovations being offered by Two Chairs meaningfully improve outcomes for patients.

Disclosures

Nicholas R. Forand, Jasmine Nettiksimmons, and Margaret Anton are employees of and have equity interest in Two Chairs.

Acknowledgments

We appreciated the review and feedback on this paper from our colleagues at Two Chairs (Maura Church and Marisa Perera) and JG Research & Evaluation (Kristal Jones and Carol Hardy).

Footnotes

¹ Comparisons are based on two large studies of outcomes in general practice. The first comparator drew data from a meta-analytic review of “usual care” conditions in the treatment of depression,

where rates of PHQ-9 response at treatment completion were 25% (usual care) vs 51.1% (Two Chairs) and rates of PHQ-9 reliable improvement were 31% (usual care) vs 65.3% (Two Chairs). The second comparator was from a large sample of care delivered in standard practice settings, in which rates of PHQ-9 recovery at 12 sessions were 21% (regular practice) vs. 55% (Two Chairs) and rates of GAD-7 recovery were 31% (regular practice) vs. 61% (Two Chairs). To more accurately reflect the inclusion criteria of the latter comparator sample (derived from the Blueprint Quality Index study), we included all individuals who had any follow-up data after baseline in the Two Chairs sample, in contrast to our primary analyses which included follow-up data only for those who completed at least 4 sessions.

References

1. Arias D, Saxena S, Verguet S. Quantifying the global burden of mental disorders and their economic value. *eClinicalMedicine* [Internet]. 2022 Dec 1 [cited 2025 Feb 24];54. Available from: [https://www.thelancet.com/journals/eclinm/article/PIIS2589-5370\(22\)00405-9/fulltext](https://www.thelancet.com/journals/eclinm/article/PIIS2589-5370(22)00405-9/fulltext)
2. Greenberg PE, Fournier AA, Sisitsky T, Simes M, Berman R, Koenigsberg SH, et al. The Economic Burden of Adults with Major Depressive Disorder in the United States (2010 and 2018). *Pharmacoeconomics*. 2021 Jun;39(6):653–65.
3. Cutler AJ, Keyloun KR, Higa S, Park J, Bonafede M, Gillard P, et al. Annual costs among patients with major depressive disorder and the impact of key clinical events. *JMCP*. 2022 Dec;28(12):1335–43.
4. Popkov AA, Barrett TS, Shergill A, Donohue M, Anderson RJ, Karlin BE. Association between depression symptom severity and total cost of care: Findings from a large, 2-year, claims-based, retrospective population health study. *Journal of Affective Disorders*. 2025 Jan;368:41–7.
5. Wang J, Qiu Y, Zhu X. Trends of mental health care utilization among US adults from 1999 to 2018. *BMC Psychiatry*. 2023 Sep 12;23(1):665.
6. Swift JK, Callahan JL, Cooper M, Parkin SR. The impact of accommodating client preference in psychotherapy: A meta-analysis. *J Clin Psychol*. 2018 Nov;74(11):1924–37.
7. Shimokawa K, Lambert MJ, Smart DW. Enhancing treatment outcome of patients at risk of treatment failure: meta-analytic and mega-analytic review of a psychotherapy quality assurance system. *J Consult Clin Psychol*. 2010 Jun;78(3):298–311.

8. Simon GE, Imel ZE, Ludman EJ, Steinfeld BJ. Is dropout after a first psychotherapy visit always a bad outcome? *Psychiatr Serv.* 2012 Jul;63(7):705–7.
9. Barrett MS, Chua WJ, Crits-Christoph P, Gibbons MB, Thompson D. Early withdrawal from mental health treatment: Implications for psychotherapy practice. *Psychotherapy: Theory, Research, Practice, Training.* 2008;45(2):247–67.
10. Kolovos S, Van Tulder MW, Cuijpers P, Prigent A, Chevreur K, Riper H, et al. The effect of treatment as usual on major depressive disorder: A meta-analysis. *Journal of Affective Disorders.* 2017 Mar;210:72–81.
11. Sawchuk CN, Craner JR, Berg SL, Smyth K, Mack J, Glader M, et al. Initial outcomes of a real-world multi-site primary care psychotherapy program. *General Hospital Psychiatry.* 2018 Sep;54:5–11.
12. Cuijpers P, Karyotaki E, Ciharova M, Miguel C, Noma H, Furukawa TA. The effects of psychotherapies for depression on response, remission, reliable change, and deterioration: A meta-analysis. *Acta Psychiatr Scand.* 2021 Sep;144(3):288–99.
13. Founder DFB, October 11 CEO, 2023. Introducing the Blueprint Quality Index – A New Way to Benchmark Behavioral Healthcare Quality [Internet]. [cited 2025 Feb 24]. Available from: <https://www.blueprint.ai/blog/introducing-the-blueprint-quality-index-a-new-way-to-benchmark-behavioral-healthcare-quality>
14. McKay CE. Evidence Based Practices in Mental Health: Advantages, Disadvantages, and Research Considerations.
15. Bowser DM, Henry BF, McCollister KE. Cost analysis in implementation studies of evidence-based practices for mental health and substance use disorders: a systematic review.

- Implement Sci. 2021 Mar 12;16:26.
16. McAleavey AA, de Jong K, Nissen-Lie HA, Boswell JF, Moltu C, Lutz W. Routine Outcome Monitoring and Clinical Feedback in Psychotherapy: Recent Advances and Future Directions. *Adm Policy Ment Health*. 2024 May 1;51(3):291–305.
 17. Miller SD, Hubble MA, Chow D, Seidel J. Beyond measures and monitoring: Realizing the potential of feedback-informed treatment. *Psychotherapy*. 2015 Dec;52(4):449–57.
 18. Fixsen DL, Blase KA, Naoom SF, Wallace F. Core Implementation Components. *Research on Social Work Practice*. 2009 Sep;19(5):531–40.
 19. Roundfield KD, Lang JM. Costs to Community Mental Health Agencies to Sustain an Evidence-Based Practice. *Psychiatr Serv*. 2017 Sep 1;68(9):876–82.
 20. Edmunds JM, Beidas RS, Kendall PC. Dissemination and Implementation of Evidence-Based Practices: Training and Consultation as Implementation Strategies. *Clin Psychol (New York)*. 2013 Jun 1;20(2):152–65.
 21. Hooke GR, Savani P, Stewart B, Araujo S, Page AC. Illustrating routine outcomes monitoring at different points in a patient's journey: Inpatient then daypatient treatment of a patient with depressive and borderline symptoms. *J Clin Psychol*. 2022 Oct;78(10):2041–53.
 22. Barkham M, De Jong K, Delgadillo J, Lutz W. Routine Outcome Monitoring (ROM) and Feedback: Research Review and Recommendations. *Psychotherapy Research*. 2023 Oct 3;33(7):841–55.
 23. Lewis CC, Boyd M, Puspitasari A, Navarro E, Howard J, Kassab H, et al. Implementing Measurement-Based Care in Behavioral Health: A Review. *JAMA Psychiatry*. 2018 Dec 19;10.1001/jamapsychiatry.2018.3329.

24. Schuckard E, Miller S, Hubble M. Feedback-informed treatment: Historical and empirical foundations. In: Feedback-informed treatment in clinical practice: Reaching for excellence.
25. Scott K, Lewis CC. Using Measurement-Based Care to Enhance Any Treatment. *Cogn Behav Pract.* 2015 Feb;22(1):49–59.
26. Price MS, Christensen H, Urquhart J, Butte S, Acierno R, Little DM. Patient-facing, Semi-automated, solutions to enhance patient participation in measurement-based care practice standards in a tele-mental health specialty clinic for populations at higher risk for self-harm. *International Journal of Medical Informatics.* 2023 Sep;177:105155.
27. Tsai KH, Moskowitz AL, Brown TE, Park AL, Chorpita BF. Interpreting Progress Feedback to Guide Clinical Decision-Making in Children's Mental Health Services. *Adm Policy Ment Health.* 2016 Mar;43(2):199–206.
28. Dodson P, Haase AM, Jeffreys M, Hales C. Capturing patient experiences of care with digital technology to improve service delivery and quality of care: A scoping review. *DIGITAL HEALTH.* 2024 Sep 1;10:20552076241282900.
29. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med.* 2001 Sep;16(9):606–13.
30. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A Brief Measure for Assessing Generalized Anxiety Disorder: The GAD-7. *Arch Intern Med.* 2006 May 22;166(10):1092.
31. Kroenke K. Enhancing the clinical utility of depression screening. *CMAJ.* 2012 Feb 21;184(3):281–2.
32. Toussaint A, Hüsing P, Gumz A, Wingenfeld K, Härter M, Schramm E, et al. Sensitivity to change and minimal clinically important difference of the 7-item Generalized Anxiety Disorder

- Questionnaire (GAD-7). *Journal of Affective Disorders*. 2020 Mar;265:395–401.
33. Kalin NH. The Critical Relationship Between Anxiety and Depression. *AJP*. 2020 May 1;177(5):365–7.
34. Gyani A, Shafran R, Layard R, Clark DM. Enhancing recovery rates: Lessons from year one of IAPT. *Behaviour Research and Therapy*. 2013 Sep;51(9):597–606.
35. Brehaut E, Neupane D, Levis B, Wu Y, Sun Y, Ioannidis JPA, et al. ‘Optimal’ cutoff selection in studies of depression screening tool accuracy using the PHQ-9, EPDS, or HADS-D: A meta-research study. *Int J Methods Psych Res*. 2023 Sep;32(3):e1956.
36. Simon GE, Ludman EJ. Predictors of Early Dropout From Psychotherapy for Depression in Community Practice. 2010;61(7).
37. Bellon J, Quinlan C, Taylor B, Nemecek D, Borden E, Needs P. Association of Outpatient Behavioral Health Treatment With Medical and Pharmacy Costs in the First 27 Months Following a New Behavioral Health Diagnosis in the US. *JAMA Netw Open*. 2022 Dec 6;5(12):e2244644.

Supplemental Table

Supplemental Table 1. Fixed effects output from longitudinal models of change in sessions 1

- 12 for PHQ-9 and GAD-7.

	PHQ-9 model				GAD-7 model			
	Estimate	Std. error	z value	Pr(> z)	Estimate	Std. error	z value	Pr(> z)
Intercept	14.80	0.15	100.07	0	14.29	0.14	105.34	0
Client session number	-1.00	0.02	-41.64	0	-1.06	0.02	-43.87	0
Client session number squared	0.05	0.00	26.33	0	0.05	0.00	28.26	0
Age 18 - 29	0.55	0.19	2.84	0	0.23	0.18	1.3	0.19
Age 50 - 69	0.39	0.22	1.79	0.07	0.30	0.21	1.44	0.15
Age 70+	-0.34	0.43	-0.79	0.43	-0.48	0.41	-1.17	0.24
Gender: man	-0.06	0.19	-0.29	0.77	-0.35	0.18	-1.94	0.05
Gender: nonbinary/trans/other	-0.28	0.35	-0.8	0.42	-0.42	0.35	-1.19	0.23
Gender: unknown	-0.55	0.27	-2.06	0.04	-0.60	0.24	-2.46	0.01

Supplementary Files