A Cluster Randomized Trial of Primary Care Practice Redesign to Integrate Behavioral Health for Those Who Need It Most: Patients With Multiple Chronic Conditions

Benjamin Littenberg, MD; Jessica Clifton, PhD; Abigail M. Crocker, PhD; Laura-Mae Baldwin, MD, MPH; Levi N. Bonnell, PhD; Ryan E. Breshears, PhD; Peter Callas, PhD; Prama Chakravarti, PhD; Kelly Clark/Keefe, EdD; Deborah J. Cohen, PhD; Frank V. deGruy, MD; Lauren Eidt-Pearson, MSW, LICSW; William Elder, PhD; Chester Fox, MD; Sylvie Frisbie; Katie Hekman, RN; Juvena Hitt, MPH; Jennifer Jewiss, EdD; David C. Kaelber, MD, PhD, MPH; Kairn Stetler Kelley, PhD; Rodger Kessler, PhD; Jennifer B. O'Rourke-Lavoie; George S. Leibowitz, PhD; C. R. Macchi, PhD; Matthew P. Martin, PhD; Mark McGovern, PhD; Brenda Mollis, MPH, MPA, MA; Daniel Mullin, PsyD, MPH; Zsolt Nagykaldi, PhD; Lisa W. Natkin, PhD; Wilson Pace, MD; Richard G. Pinckney, MD, MPH; Douglas Pomeroy; Paula Reynolds; Gail L. Rose, PhD; Sarah Hudson Scholle, DrPH, MPH; William J. Sieber, PhD; Jeni Soucie, MS; Terry Stancin, PhD; Kurt C. Stange, MD, PhD; Kari A. Stephens, PhD; Kathryn Teng, MD; Elizabeth Needham Waddell, PhD; Constance van Eeghen, DrPH, MHSA, MBA

MORE ONLINE
www.annfammed.org

Conflicts of interest: authors report none.

CORRESPONDING AUTHOR

Benjamin Littenberg
Division of General Internal Medicine Research
Department of Medicine
Robert Larner MD College of Medicine
at the University of Vermont
University of Vermont
89 Beaumont Ave
Burlington, VT 05405
benjamin.littenberg@uvm.edu

ABSTRACT

PURPOSE Patient outcomes can improve when primary care and behavioral health providers use a collaborative system of care, but integrating these services is difficult. We tested the effectiveness of a practice intervention for improving patient outcomes by enhancing integrated behavioral health (IBH) activities.

METHODS We conducted a pragmatic, cluster randomized controlled trial. The intervention combined practice redesign, quality improvement coaching, provider and staff education, and collaborative learning. At baseline and 2 years, staff at 42 primary care practices completed the Practice Integration Profile (PIP) as a measure of IBH. Adult patients with multiple chronic medical and behavioral conditions completed the Patient-Reported Outcomes Measurement Information System (PROMIS-29) survey. Primary outcomes were the change in 8 PROMIS-29 domain scores. Secondary outcomes included change in level of integration.

RESULTS Intervention assignment had no effect on change in outcomes reported by 2,426 patients who completed both baseline and 2-year surveys. Practices assigned to the intervention improved PIP workflow scores but not PIP total scores. Baseline PIP total score was significantly associated with patient-reported function, independent of intervention. Active practices that completed intervention workbooks (n = 13) improved patient-reported outcomes and practice integration ($P \leq .05$) compared with other active practices (n = 7).

CONCLUSION Intervention assignment had no effect on change in patient outcomes; however, we did observe improved patient outcomes among practices that entered the study with greater IBH. We also observed more improvement of integration and patient outcomes among active practices that completed the intervention compared to active practices that did not. Additional research is needed to understand how implementation efforts to enhance IBH can best reach patients.

INTRODUCTION

Comprehensive primary care for patients with chronic diseases that might include behavioral factors require targeted behavioral health services.[1] Providing care for mental health, substance use disorders, and health behavior changes, and attending to psychosocial factors could decrease the negative health effects of tobacco, diet, physical inactivity, alcohol, and other substances, which together account for 38% of all US deaths,[2,3] and prevent additional morbidity from treatment nonadherence, insomnia, anxiety, depression, and stress.[4-6] Coordinating behavioral health and primary care services has shown improved outcomes for patients with depression (eg, Collaborative Care Model[7]) and patients with mixed depression and anxiety.[8,9] Integrating primary and behavioral health care management across multiple chronic conditions (MCCs) has shown greater effectiveness than integration focused on a single diagnosis and could benefit patients, families, and the health care system.[10-16]

Integrated primary care and behavioral health (IBH) is "the care that results from a practice team of primary care and behavioral health clinicians, working together with patients and families, using a systematic and cost-effective approach to provide patient-centered care for a defined population."[17] Integration efforts can include population management, protocol-driven evidence-supported care, guideline-based external referrals, systematic needs identification, team-based care,

and practice-wide approaches to patient engagement.[17] Less intensive integration can include working with a Behavioral Health (BH) provider in the same building (ie, colocation) without shared scheduling and record systems.[17] Integration of these services is known to be efficacious, but quantifying integration is complex, involving evaluation across domains of clinical services, workflow, workspace, integration of providers, patient identification procedures, and methods of engaging patients.[18-21] More intensive integration efforts are hypothesized to result in better patient outcomes[17] but have not been studied in broadly defined populations.

Practices that attempt to increase IBH often encounter barriers arising from health care system complexity.[1] Structured implementation processes can address such barriers[22] and some, such as implementation toolkits, have demonstrated innovation uptake in health care.[23,24] Study authors (C.vE., R.K., B.L.) developed early versions of such an intervention to help practices enhance IBH by engaging frontline workers in redesigning processes while eliminating inefficient work. This approach was successful in establishing onsite BH in primary care, increasing care initiation, and decreasing time to first BH visit.[25] The research team expanded on that approach, incorporating more implementation best practice strategies[26,27] to create the current version of the integrated care toolkit tested in the present study.

The purpose of this initiative was to study the effect of a single approach to IBH improvement on outcomes for primary care patients experiencing MCCs and, secondarily, to study the effect on primary care efforts to enhance integration. We evaluated the real-world effectiveness of toolkit access and support in enhancing IBH and outcomes reported by patients with MCCs. We also explored associations between IBH and patient outcomes, toolkit completion and change in IBH, and toolkit completion and change in patient outcomes.

METHODS

Design

Our multidisciplinary research team included members in the fields of psychology, family medicine, internal medicine, clinical and translational science, education, art, biostatistics, communication, public health, bioinformatics, and health care operations, with a variety of backgrounds including clinicians, patients, family members, leaders, and researchers from across the United States. We performed a large-scale, pragmatic, cluster randomized, clinical effectiveness trial. Pragmatic trials speed the translation of research into practice[28] by testing interventions under usual, rather than ideal, conditions.[29] Primary care practices were randomized to the active or control group. We used an intention-to-treat analysis and conducted a difference-in-differences analysis of BH integration and patient-reported outcomes during the period 2018 to 2020, including the onset of the coronavirus disease 2019 (COVID-19) pandemic. Study methods[30] and qualitative findings[31] are reported elsewhere and are registered at ClinicalTrials.gov (NCT02868983). The University of Vermont Committee on Human Research in the Medical Sciences (#16-554) and institutional review boards at other participating locations approved the protocol.

Procedures

We recruited primary care practices during the period 2016 to 2018 that had a colocated behavioral health provider (BHP) (ie, psychologist, social worker, or licensed counselor) of ≥0.5 full-time equivalent, capable of billing insurers, using electronic health records, and with scores <75 on the Practice Integration Profile 1.0 (PIP), a measure of IBH.[32] The sites included internal medicine, family medicine, small and large groups, for-profit, academic, and safety-net clinics recruited from practice networks across 13 US states using professional networks, association listservs, and conference presentations. Measures of practice integration were collected from 42 randomized practices.

We recruited patient participants from the 41 practices able to provide data on eligible patients. A third party reviewed electronic health records to identify adult patients from each practice with ≥1 chronic medical condition (arthritis; obstructive lung disease including emphysema, chronic bronchitis, or asthma; nongestational diabetes; or heart disease manifested as heart failure or hypertension), ≥1 chronic BH condition (diagnosis related to mood including anxiety or depression; chronic pain including headache, migraine, neuralgia, fibromyalgia, or chronic musculoskeletal pain; insomnia; irritable bowel syndrome; or substance misuse including substance use disorder, tobacco use, or problem drinking), or 3 medical conditions. Patients were randomly selected and contacted by e-mail, telephone, or postal mail with an invitation to join the study.

Intervention

Practices randomized to the active group were required to assemble a team with a facilitator. The intervention provided access to and support for workbooks to guide the planning, redesign, and implementation activities of a quality improvement (QI) project; online education tailored to clinic roles (physician, BHP, nurse, etc); an online learning community; and remote coaching for the team facilitator and QI team. Workbooks included specific tactics, such as care management, use of registries and population reports, and team-based care, known to be beneficial to delivering IBH to patients with MCCs. Given the pragmatic nature of the study, each QI team determined when to start and which elements of the intervention to use within an observation period from 2018 to 2020. An example of the toolkit can be found at https://sites.google.com/view/ibhpc/home.

Measures

Degree of practice integration was measured using the PIP, a survey of staff perceptions of integration[32-34]

(Supplemental Appendix). Practice Integration Profile total score (range 0-100) is the unweighted average of the following 6 survey domains: workflow, clinical services, workspace, integration of providers, patient identification, and patient engagement. A medical provider, BHP, administrator, and ≥1 additional practice member rated their practice's performance using the PIP at baseline, midpoint, and 2 years. In addition, coaches monitored active practices' progress and determined which practices completed workbook activities through implementation of ≥1 suggested practice change in ≤24 months.

Patients who agreed to participate completed the Patient-Reported Outcomes Measurement Information System[35] (PROMIS-29) online survey at baseline, midpoint, and 2 years. The PROMIS-29 survey includes 8 domains (physical function, anxiety, depression, fatigue, sleep disturbance, social functioning, pain intensity, and pain interference) and 2 summary scores.[36] Other measures included change in

empathy,[37] medication adherence,[38] self-reported health care utilization,[39] time lost due to disability,[40] functional capacity,[41] patient centeredness,[42] depression,[43] anxiety,[44] asthma symptoms,[45,46] substance use disorder,[47] and problem drinking.[48] All measures had published indications of adequate validity and reliability.[32,33,35] For all measures, baseline data were collected before the COVID-19 pandemic; final measures were collected after the pandemic start.

Analysis

Intervention Effect on Patients' PROMIS-29 Scores

Primary outcomes were the changes (baseline to 2-year) in the 8 PROMIS-29 domain scores. The 2 PROMIS-29 summary scores were secondary outcomes. For each domain, we built a multivariate mixed linear regression model with change in outcome from baseline to 2 years as the dependent variable. Practice assignment to the intervention was the independent variable, modeled as a fixed effect, as were

Table 1. Patient Characteristics at Baseline

	Total		Control		Active		P Value
	No. (%)	Mean (SD)	No. (%)	Mean (SD)	No. (%)	Mean (SD)	
n[a]	2,426		1,459		967		
Age,[b] y		61.9 (13.1)		61.9 (13.4)		61.9 (12.7)	.49
Sex[b]							.15
Female	1,588 (65.5)		970 (66.5)		618 (63.9)		
Male	832 (34.3)		483 (33.1)		349 (36.1)		
Race[b]							.01
White	1,859 (76.8)		1,157 (79.3)		702 (72.9)		
Black or African American	281 (11.6)		154 (10.6)		127 (13.2)		
Asian	71 (2.9)		35 (2.4)		36 (3.7)		
Native Hawaiian/Other Pacific Islander	37 (1.5)		17 (1.2)		20 (2.1)		
American Indian or Alaskan Native	21 (0.9)		13 (0.9)		8 (0.8)		
Other	153 (6.3)		83 (5.7)		70 (7.3)		
Ethnicity[b]							.14
Non-Hispanic	2,207 (91.5)		1,342 (92.4)		865 (90.1)		
Hispanic	180 (7.5)		96 (6.6)		84 (8.8)		
Prefer not to say	26 (1.1)		15 (1.0)		11 (1.1)		
Marital status[b]							.13
Married	1,082 (44.8)		621 (42.7)		461 (47.9)		
Divorced	517 (21.4)		318 (21.9)		199 (20.7)		
Never married	405 (16.8)		249 (17.1)		156 (16.2)		
Widowed	290 (12.0)		183 (12.6)		107 (11.1)		
Separated	62 (2.6)		44 (3.0)		18 (1.9)		
Living as married	61 (2.5)		39 (2.7)		22 (2.3)		

continues

CARE = Consultation and Relational Empathy; DASI = Duke Activity Status Index; ED = emergency department; GAD-7 = 7-item Generalized Anxiety Disorder screener; GAIN = Global Appraisal of Individual Needs; GED = General Education Development; MMAS = Morisky Medication Adherence Scale; PHQ-9 = 9-item Patient Health Questionnaire; PROMIS-29 = Patient-Reported Outcomes Measurement Information System; SUD = substance use disorder.

Note: Wilcoxon rank-sum test used for continuous variables; χ^2 test used for categorical variables.

[a] No. varies for some characteristics.

[b] Included as possible confounding effect. Other possible included various county characteristics (age, racial composition, ethnicity, sex, income, education, employment, population density, urban/rural status).

[c] Greater score indicates better health. T-score is a measure in which 50 is the mean of a relevant population, and 10 is the SD of that population.

Table 1. Patient Characteristics at Baseline (continued)

	Total		Control		Active		P Value
	No. (%)	Mean (SD)	No. (%)	Mean (SD)	No. (%)	Mean (SD)	
Employment[b]							.11
Retired	864 (36.6)		525 (36.8)		339 (36.2)		
Disabled	627 (26.5)		372 (26.1)		255 (27.2)		
Full-time	486 (20.6)		274 (19.2)		212 (22.6)		
Part-time	205 (8.7)		136 (9.5)		69 (7.4)		
Homemaker	94 (4.0)		60 (4.2)		34 (3.6)		
Unemployed/looking	68 (2.9)		43 (3.0)		25 (2.7)		
Student	15 (0.6)		13 (0.9)		2 (0.2)		
Other	3 (0.1)		2 (0.1)		1 (0.1)		
Annual household income,[b] $							
< 15,000	681 (29.5)		416 (30.0)		265 (28.7)		.04
15,000-29,999	509 (22.0)		308 (22.2)		201 (21.8)		
30,000-44,999	281 (12.2)		180 (13.0)		101 (10.9)		
45,000-59,999	186 (8.0)		116 (8.4)		70 (7.6)		
60,000-74,999	190 (8.2)		121 (8.7)		69 (7.5)		
75,000-99,999	188 (8.1)		103 (7.4)		85 (9.2)		
> 100,000	277 (12.0)		144 (10.4)		133 (14.4)		
Education[b]							.92
< Grade 9	56 (2.4)		32 (2.2)		24 (2.5)		
Grades 9-12, no diploma	208 (8.8)		123 (8.6)		85 (9.0)		
High school graduate (including GED)	1,004 (42.2)		608 (42.6)		396 (41.7)		
Associate degree	378 (15.9)		222 (15.5)		156 (16.4)		
Bachelor's degree	387 (16.3)		229 (16.0)		158 (16.6)		
Graduate or professional degree	344 (14.5)		214 (15.0)		130 (13.7)		
Chronic conditions[b]							
Arthritis	1,023 (42.2)		613 (42.0)		410 (42.4)		.85
Asthma	545 (22.5)		316 (21.7)		229 (23.7)		.24
Chronic obstructive pulmonary disease	330 (13.6)		214 (14.7)		116 (12.0)		.06
Chronic pain	2,037 (84.0)		1,213 (83.1)		824 (85.2)		.17
Diabetes (nongestational)	1,075 (44.3)		671 (46.0)		404 (41.8)		.04
Heart failure	188 (7.8)		119 (8.2)		69 (7.1)		.36
Hypertension	2,012 (82.9)		1,207 (82.7)		805 (83.3)		.74
Irritable bowel syndrome	102 (4.2)		56 (3.8)		46 (4.8)		.27
Anxiety	830 (34.2)		491 (33.7)		339 (35.1)		.48
Depression	1,136 (46.8)		697 (47.8)		439 (45.4)		.25
Insomnia	570 (23.5)		339 (23.2)		231 (23.9)		.71
Any substance use disorder	559 (23.0)		339 (23.2)		220 (22.8)		.78
Tobacco use	445 (18.3)		274 (18.8)		171 (17.7)		.49
Alcohol use disorder	155 (6.4)		93 (6.4)		62 (6.4)		.97
Mean number of chronic conditions[b]		4.4 (1.6)		4.4 (1.6)		4.3 (1.6)	.6

continues

CARE = Consultation and Relational Empathy; DASI = Duke Activity Status Index; ED = emergency department; GAD-7 = 7-item Generalized Anxiety Disorder screener; GAIN = Global Appraisal of Individual Needs; GED = General Education Development; MMAS = Morisky Medication Adherence Scale; PHQ-9 = 9-item Patient Health Questionnaire; PROMIS-29 = Patient-Reported Outcomes Measurement Information System; SUD = substance use disorder.

Note: Wilcoxon rank-sum test used for continuous variables; χ^2 test used for categorical variables.

[a] No. varies for some characteristics.

[b] Included as possible confounding effect. Other possible confounders included various county characteristics (age, racial composition, ethnicity, sex, income, education, employment, population density, urban/rural status).

[c] Greater score indicates better health. T-score is a measure in which 50 is the mean of a relevant population, and 10 is the SD of that population.

Table 1. Patient Characteristics at Baseline (continued)

	Total		Control		Active		P Value
	No. (%)	Mean (SD)	No. (%)	Mean (SD)	No. (%)	Mean (SD)	
Neighborhood characteristics[b] (home census tract)							
Social Deprivation Index (higher score indicates more deprivation)		52.6 (27.7)		53.5 (26.9)		51.2 (29.0)	.05
Rural	477 (19.8)		341 (23.4)		136 (14.3)		<.001
Population density, persons per square mile		3,768 (6,656)		2,978 (3,315)		4,980 (9,645)	<.001
Primary outcomes – PROMIS-29 t-scores[c]							
Anxiety		53.9 (10.0)		53.8 (9.9)		54.1 (10.1)	.34
Depression		52.7 (9.8)		52.7 (9.8)		52.9 (9.7)	.46
Fatigue		52.6 (10.4)		52.6 (10.4)		52.5 (10.4)	.66
Sleep disturbance		53.1 (8.9)		52.9 (8.8)		53.3 (8.9)	.26
Pain interference		58.1 (10.0)		58.0 (10.1)		58.2 (10.0)	.54
Pain intensity		4.5 (2.8)		4.4 (2.7)		4.5 (2.8)	.58
Social participation[c]		48.3 (9.9)		48.4 (9.8)		48.1 (10.0)	.65
Physical function[c]		43.6 (9.4)		43.5 (9.4)		43.7 (9.3)	.49
Secondary outcomes							
PROMIS-29 Physical Health Summary t-score[c]		45.9 (9.4)		45.8 (9.5)		46.0 (9.3)	.7
PROMIS-29 Mental Health Summary t-score[c]		50.3 (8.8)		50.4 (8.8)		50.2 (8.9)	.67
CARE total score[c]		4.3 (0.9)		4.3 (0.9)		4.3 (0.9)	.03
MMAS total score[c]		1.0 (1.1)		1.0 (1.1)		1.0 (1.1)	.57
ED visits in past year		1.0 (1.8)		1.0 (1.7)		1.1 (1.8)	.94
Health care visits in past month		2.5 (2.7)		2.4 (2.6)		2.6 (2.8)	.05
Hospital days in past year		1.0 (2.4)		1.0 (2.3)		1.0 (2.5)	.28
Restricted activity days		1.1 (2.5)		1.1 (2.5)		1.1 (2.5)	.86
Metabolic equivalents (DASI)[c]		6.4 (2.0)		6.4 (1.9)		6.4 (2.0)	.96
PHQ-9 total score		6.4 (6.1)		6.3 (6.1)		6.5 (6.1)	.48
GAD-7 total score		4.5 (5.2)		4.4 (5.2)		4.7 (5.3)	.17
Asthma Symptom Utility Index		0.8 (0.2)		0.8 (0.2)		0.8 (0.2)	.57
GAIN Lifetime SUD screener		1.7 (1.9)		1.7 (1.9)		1.8 (1.9)	.55
GAIN 1 Year SUD screener		0.6 (1.2)		0.6 (1.1)		0.6 (1.2)	.57
GAIN 3 Month SUD screener		0.5 (1.0)		0.5 (1.0)		0.5 (1.0)	.62
GAIN 1 Month SUD screener		0.5 (0.9)		0.4 (0.9)		0.5 (1.0)	.63
Unhealthy alcohol use	171 (9.6)		109 (9.7)		62 (9.6)		.94
Patient-Centeredness Index[c]		83.9 (15.9)		84.0 (15.9)		83.8 (15.7)	.62

CARE = Consultation and Relational Empathy; DASI = Duke Activity Status Index; ED = emergency department; GAD-7 = 7-item Generalized Anxiety Disorder screener; GAIN = Global Appraisal of Individual Needs; GED = General Education Development; MMAS = Morisky Medication Adherence Scale; PHQ-9 = 9-item Patient Health Questionnaire; PROMIS-29 = Patient-Reported Outcomes Measurement Information System; SUD = substance use disorder.

Note: Wilcoxon rank-sum test used for continuous variables; χ[2] test used for categorical variables.

[a] No. varies for some characteristics.

[b] Included as possible confounding effect. Other possible confounders included various county characteristics (age, racial composition, ethnicity, sex, income, education, employment, population density, urban/rural status).

[c] Greater score indicates better health. T-score is a measure in which 50 is the mean of a relevant population, and 10 is the SD of that population.

baseline level of the outcome measure and potential confounders (see below). A practice identifier was included as a random effect. We did not look for trends at midpoint because the intervention had not been in place long enough to affect outcomes.

We examined the possible confounding effects of 29 characteristics (Tables 1 and 2) on the relation between the intervention and each outcome. We compared the coefficient on the predictor from a mixed model with no fixed effect covariates to that from a model with a single covariate.

Table 2. Practice Characteristics at Baseline

	Total n = 42		Control n = 22		Active n = 20		P
	No. (%)	Mean (SD)	No. (%)	Mean (SD)	No. (%)	Mean (SD)	Value
Practice specialty[a]							.92
Internal medicine	7 (17)		4 (18)		3 (15)		
Family medicine	20 (48)		11 (50)		9 (45)		
Mixed	15 (36)		7 (32)		8 (40)		
Organization type[a,b]							
Community health center	15 (36)		7 (32)		8 (40)		.75
Hospital	20 (48)		10 (45)		10 (50)		> .99
Private	4 (10)		3 (14)		1 (5)		.61
Academic	19 (45)		9 (41)		10 (50)		.76
Number of patient visits per year[a]		27,039 (19,536)		30,104 (24,210)		23,668 (12,363)	.51
Baseline primary care provider FTE[a]		6.0 (3.2)		6.1 (3.6)		5.9 (2.7)	.96
Baseline BH provider FTE[a]		1.5 (1.1)		1.3 (0.7)		1.7 (1.4)	.27
Patients cared for by the practice per year[a]		9,285 (5,066)		9,419 (5,599)		9,138 (4,549)	.94
Resident training site[a]	16 (38)		7 (32)		9 (45)		.53
Nonprofit[a]	37 (88)		18 (82)		19 (95)		.35
Tenure of on-site BH services, y		6.3 (6.4)		6.6 (7.4)		5.9 (5.3)	.86
Geographic region							.9
Pacific Northwest	3 (7)		2 (9)		1 (5)		
Mountain	8 (19)		4 (18)		4 (20)		
South	8 (19)		4 (18)		4 (20)		
New England	9 (21)		6 (27)		3 (15)		
Mid-Atlantic & Great Lakes	6 (14)		3 (14)		3 (15)		
West Coast & Hawaii	8 (19)		3 (14)		5 (25)		
Urban by RUCA	35 (83)		17 (77)		18 (90)		.41
County social deprivation index		45 (22)		44 (21)		46 (23)	.76
PIP baseline total median[a]		59 (14)		59 (17)		59 (11)	.96

BH = behavioral health; FTE = full-time equivalent employee; PIP = Practice Integration Profile; RUCA = rural-urban commuting area.

Note: *P* values calculated by Wilcoxon rank-sum test or Fisher exact test.

[a] Included as possible confounding effect.

[b] Multiple types were possible; percentages may not sum to 100%.

For each analysis, we retained potential confounders associated with both outcome and predictor with $P < .10$ and also retained those affecting the association between the predictor and outcome.

The study was designed for 90% power to detect differences as small as 2.5 points between active and control arms in any of the 8 PROMIS-29 domain scales, consistent with minimally important differences specified for other PROMIS instruments of 2 to 8 points.[49,50] We used the Bonferroni correction to maintain $\alpha = 0.05$ as we examined the effect of the intervention on each of the 8 PROMIS-29 domains, resulting in a threshold of $P = .00625$.[51] Additional models explored the association between the intervention and patient outcomes in subgroups based on patient characteristics, presence of qualifying conditions, number of qualifying conditions, neighborhood characteristics, and practice characteristics. These analyses were exploratory rather than hypothesis testing and were not adjusted for multiple comparisons.

Intervention Effect on Practices' PIP Scores

We tested whether the change in PIP total score differed between active and control practices. In a multivariate mixed linear regression model, change in PIP total score (baseline to 2-year) was the dependent variable, intervention assignment was the predictor, and covariates were selected from potential confounders identified using the procedure described above. Change in PIP domain scores was also explored. No random effects were modeled, and no correction for multiple comparisons was applied. With a sample size of 42 practices and an SD of 14.2 points, the study had 80% power to detect an effect of 12.7 points.
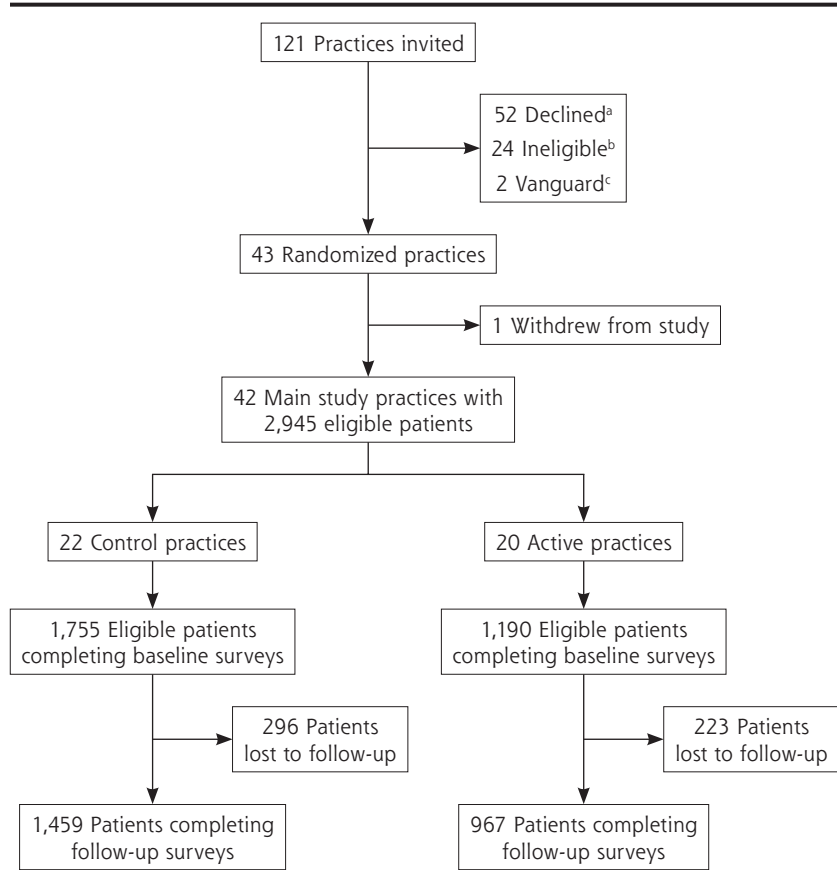
Exploratory Analyses

To better understand our findings, we built several mixed linear models to explore associations not part of the original study design. We did not correct for multiple comparisons but tested for confounding as described above and included potential confounders in our models. To evaluate the association between practices' baseline IBH on patient outcomes, we modeled each baseline and 2-year PROMIS-29 domain and summary score as dependent variables, baseline PIP total score as a fixed effect, and practice as a random effect. Noting that active practices varied in their fidelity using the intervention, we explored the association between intervention workbook completion and changes in IBH and patient outcomes. To evaluate the relation between intervention completion and change in IBH, we modeled change in PIP scores, which were adjusted for baseline PIP scores, as dependent variables, completion status as a fixed independent effect, and practice as a random effect. To determine if there was an association between intervention completion and change in patient outcomes, we modeled changes in patient outcomes, also adjusted for baseline outcomes, as dependent variables, completion status as a fixed independent effect, and practice as a random effect. All data were analyzed using Stata 17 (StataCorp LLC).

RESULTS

Of 504,827 patient records reviewed electronically, 133,041 met the inclusion criteria, of which 66,804 were recruited and 4,407 responded from practices that stayed in the study. A total of 2,945 eligible patients did not drop out before randomization and completed the baseline PROMIS-29 survey; 2,426 (82%) completed follow-up 2 years later (Figure 1). The final sample (Table 1) had a greater prevalence of chronic disease and older average age (62 years vs 47 years) than the general population of US primary care patients, as described by the National Ambulatory Medical Care Survey,[52] but was generally similar in race, ethnicity, gender, and rurality characteristics. Of 121 practices invited to participate, 52 declined, 24 were ineligible, 2 served as nonrandomized pilot sites (vanguard sites), and 1 dropped out before data collection, leaving 42 practices in the study. There were no significant differences in practice characteristics between active and control arms (Table 2).

Intervention Effect on Patients' PROMIS-29 Scores

There was no difference in the amount of change in any of the 8 PROMIS-29 domains reported by patients receiving care from active practices compared with patients from control practices (Table 3 and Figure 2). Although mental health summary scores (adjusted for baseline and population density) improved in both groups (from 50.17 to 50.64 [$\Delta = 0.47$; 95% CI, 0.10-0.84; $P = .012$] in the active group and 50.36 to 50.77 [$\Delta = 0.41$; 95% CI, 0.11-0.70; $P = .006$] in the control group), the difference in improvement between the 2 groups was not statistically significant ($\Delta = -0.06$; 95% CI, -0.51 to 0.39; $P = .79$) or clinically meaningful. Exploratory models based on patient subgroups revealed no consistent patterns among subgroups or outcomes.

Intervention Effect on Practices' PIP Scores

Practice Integration Profile total score improved significantly during the study period for both active and control practices, but there was no statistically significant difference in improvement seen between the 2 groups (Table 3). Active practices (n = 20) improved their average PIP total scores from 59.4 to 71.6 ($\Delta = 12.2$; 95% CI, 7.0-17.4;

Figure 1. CONSORT participant flow diagram.

CONSORT = Consolidated Standards of Reporting Trials.

[a] Eligible but chose not to participate.
[b] Wanted to participate but not eligible.
[c] Eligible but volunteered to pilot the intervention before study start.

Table 3. Adjusted Effects of Assignment to Active Group on Outcomes

Outcome	No.[a]	Effect	95% CI	P Value
Primary outcomes				
Anxiety	2,426	0.08	− 0.53 to 0.69	.79
Depression	2,426	0.21	− 0.37 to 0.79	.48
Fatigue	2,426	0.07	− 0.54 to 0.68	.83
Sleep disturbance	2,426	− 0.05	− 0.58 to 0.49	.86
Pain interference	2,426	0.19	− 0.40 to 0.79	.52
Pain intensity	2,426	0.1	− 0.08 to 0.28	.27
Social participation[b]	2,426	− 0.06	− 0.71 to 0.59	.86
Physical function[b]	2,426	0.1	− 0.39 to 0.60	.69
Secondary outcomes				
Physical health summary[b]	2,426	0.09	− 0.39 to 0.56	.72
Mental health summary[b]	2,426	− 0.06	− 0.51 to 0.39	.79
CARE total score[b]	2,383	− 0.03	− 0.11 to 0.05	.49
MMAS total score[b]	2,310	− 0.03	− 0.11 to 0.05	.46
ED visits in past year	2,369	− 0.01	− 0.14 to 0.11	.82
Health care visits in past month	2,363	− 0.02	− 0.29 to 0.26	.9
Hospital days in past year	2,363	− 0.04	− 0.29 to 0.21	.78
Utilization category	2,376	0.07	− 0.13 to 0.27	.48
Restricted activity days	2,301	0.16	− 0.02 to 0.35	.08
Metabolic equivalents (DASI)[b]	2,127	0.03	− 0.09 to 0.15	.64
PHQ-9 total score	2,290	0.2	− 0.13 to 0.54	.24
PHQ category	2,290	0.14	− 0.04 to 0.32	.12
GAD-7 total score	2,317	0.03	− 0.27 to 0.32	.86
GAD category	2,317	− 0.06	− 0.25 to 0.13	.53
Asthma symptom utility index	440	− 0.02	− 0.05 to 0.01	.17
GAIN Lifetime SUD screener	376	− 0.06	− 0.35 to 0.24	.71
GAIN 1-Year SUD screener	376	− 0.15	− 0.37 to 0.06	.16
GAIN 3-Month SUD screener	376	0	− 0.19 to 0.18	.96
GAIN 1-Month SUD screener	376	− 0.04	− 0.22 to 0.14	.65
Alcohol use category	1,747	− 0.03	− 0.36 to 0.29	.84
Patient-centeredness index[b]	2,261	− 0.8	− 2.3 to 0.7	.29
Practice Integration Profile domains				
Total[b]	42	3.8	− 3.6 to 11.2	.3
Workflow[b]	42	9.3	1.7 to 16.9	.02
Clinical services[b]	42	2.6	− 6.3 to 11.5	.56
Workspace[b]	42	3.5	− 7.0 to 14.1	.5
Integration[b]	42	3.6	− 7.9 to 15.1	.53
Patient identification[b]	42	3.1	− 3.6 to 9.8	.36
Patient engagement[b]	42	3.2	− 5.3 to 11.6	.45

CARE = Consultation and Relational Empathy; DASI = Duke Activity Status Index; ED = emergency department; GAD = generalized anxiety disorder; GAD = 7-item Generalized Anxiety Disorder screener; GAIN = Global Appraisal of Individual Needs; MMAS = Morisky Medication Adherence Scale; PHQ = Patient Health Questionnaire; PHQ-9 = 9-item Patient Health Questionnaire; SUD = substance use disorder.

[a] N varies depending on availability of covariates.

[b] Greater score indicates better health.

than control practices (9.3 points; 95% CI, 1.7-16.9; P = .02), and all other PIP domains showed a pattern of nominally greater improvements (Table 3, Figure 2).

Baseline IBH
Regardless of assignment to active or control group, we observed a positive association between degree of integration (PIP total score) at baseline and patient mental health function (PROMIS-29 score) at baseline and 2 years (Table 4, Figure 3). Baseline anxiety, sleep disturbance, social participation, and mental health summary were significantly associated with greater PIP total score at baseline (eg, mental health summary: 0.05 points; 95% CI, 0-0.09; P = .05), with similar outcomes at 2 years (Table 4). Other PROMIS-29 scores showed a pattern of effect with nominally better patient function associated with greater PIP total score, although not significant when tested individually.

Intervention Workbook Completion
Active practices that completed intervention workbooks (n = 13) had significantly greater improvement in PIP total scores than practices that did not (n = 7) (9.9-point difference in change in PIP total score; 95% CI, 2.0-18.0; P = .02), with nominally greater changes in all PIP domain scores (Figure 4).

Patients who received care in practices that completed intervention workbooks reported significantly greater improvement in mental health summary (0.8; 95% CI, 0-1.6; P = .05) and social participation (1.4; 95% CI, 0.2-2.7; P = .02) (Figure 4) and nominally greater positive change on the remaining PROMIS-29 subscales than patients of active practices that did not complete workbooks, despite having access.
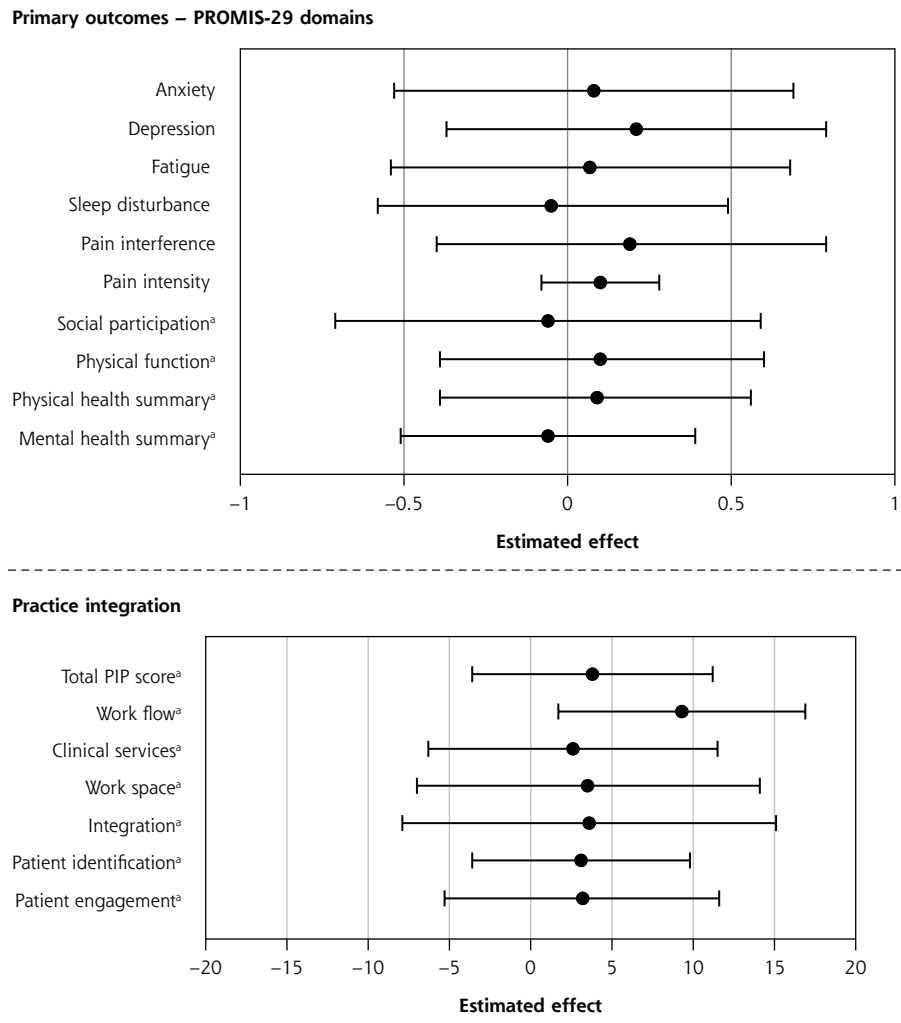
DISCUSSION
Practice access to the study intervention had no significant effect on the amount of change observed in patient-reported outcomes or PIP total score. Only the difference in PIP workflow improvement was large enough to attribute to active group assignment, which matches the structured focus of the toolkit on workflow redesign. However, study data allowed us to conduct additional analyses that could be important for future work attempting to understand how to evaluate and implement IBH.

Patients receiving care in practices that started the study with greater BH integration scores (before intervention)

P <.001); control practices (n = 22) improved from 59.4 to 67.7 (Δ = 8.3; 95% CI, 1.5-15.2; P = .027). The difference in improvement between active and control groups was not significant (Δ = 3.9; 95% CI, −3.6 to 11.2; P = .3). However, PIP workflow domain scores improved significantly more in active

Figure 2. Effect of active group assignment on patient and practice outcomes.

Primary outcomes – PROMIS-29 domains

Practice integration

PIP = Practice Integration Profile; PROMIS-29 = Patient-Reported Outcomes Measurement Information System.

Note: The center point of each bar is the mean adjusted effect (regression coefficient) of the intervention on a specific outcome, with the bars representing 95% CI. Values >0 indicate that the intervention was associated with an increase in the outcome measured. Bars that do not cross 0 indicate statistical significance (P < .05).

[a] Greater score indicates improved outcome.

reported significantly better mental health function—both at baseline and at 2 years. Patients of practices that successfully completed intervention workbooks reported significantly greater improvements in mental health after 2 years. Although not proof of cause, these observations are consistent with the expectation that patients with MCCs benefit from efforts to enhance IBH. That total PIP score was associated with patient-reported outcomes and with implementation of IBH improvements supports the continued use of PIP as a measure of integration for future IBH enhancement activities.

Although the intervention was designed to support complex practice change and adapt to local context, only 13 of 20 active practices completed the intervention during the study period. Those 13 had significantly greater improvements in IBH and patient outcomes. Because confounding factors could affect both the ability to implement and the outcomes, these observations are not evidence of causation. As noted by Walunas and others, we need more studies to understand how much intervention is sufficient for practice change[53-56] as well as how effective toolkits are in improving integration.

Limitations

The pragmatic design of the present study revealed real-world challenges to IBH. Participating practices in both groups experienced natural disasters, major work stoppages, loss of key

Table 4. Adjusted Association Between Baseline Practice Integration (PIP Total Score) and Patient Outcomes at Baseline and Follow-Up (N = 42; All Practices)

PROMIS-29 Domain	Baseline			Follow-Up		
	Effect	95% CI	P Value	Effect	95% CI	P Value
Anxiety	−0.06	−0.12 to 0	.05	−0.05	−0.09 to −0.01	.01
Depression	−0.04	−0.10 to 0.01	.1	−0.04	−0.08 to 0.01	.13
Fatigue	−0.03	−0.07 to 0.01	.1	−0.03	−0.07 to 0.01	.19
Sleep disturbance	−0.05	−0.10 to −0.01	.03	−0.02	−0.07 to 0.02	.3
Pain interference	−0.01	−0.06 to 0.04	.64	−0.03	−0.07 to 0.02	.3
Pain intensity	−0.01	−0.03 to 0.01	.52	−0.01	−0.02 to 0.01	.45
Social participation[a]	0.05	0.01 to 0.09	.02	0.06	0.02 to 0.10	.01
Physical function[a]	0.04	−0.01 to 0.08	.12	0.02	−0.03 to 0.07	.39
Physical health summary[a]	0.04	−0.01 to 0.09	.12	0.02	−0.03 to 0.07	.35
Mental health summary[a]	0.05	0 to 0.09	.05	0.04	0 to 0.09	.04

PIP = Practice Integration Profile; PROMIS-29 = Patient-Reported Outcomes Measurement Information System.

[a] Greater score indicates improved function.

Figure 3. Association of practice integration and patient outcomes.

Effect of total PIP at baseline on patient function at baseline

Linear regression coefficients and 95% CIs

Effect of total PIP at baseline on patient function at follow-up

Linear regression coefficients and 95% CIs
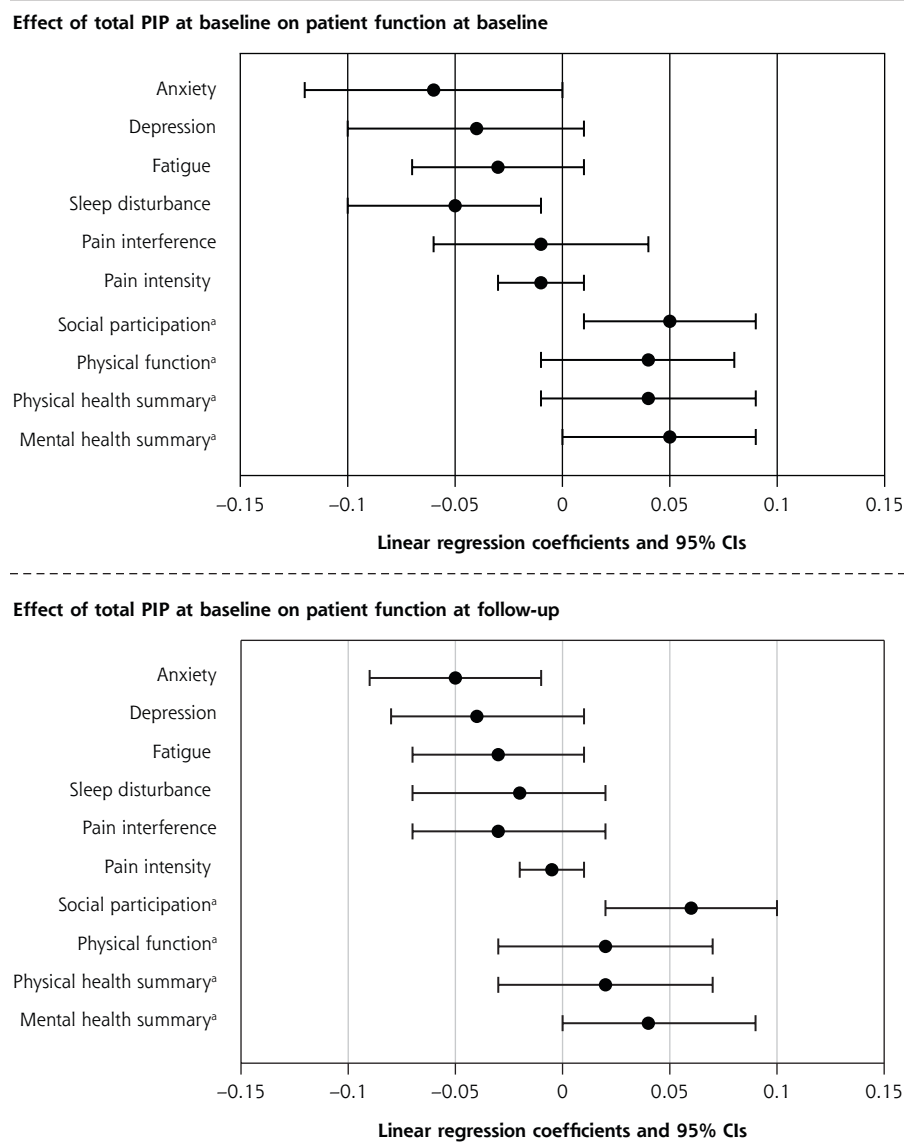
PIP = Practice Integration Profile.

Note: The center point of each bar is the mean adjusted association (regression coefficient) of the baseline median total PIP score with specific patient outcomes, adjusting for potential confounders and using a random intercept for each practice. The bars represent 95% CI. Values >0 indicate that the intervention was associated with an increase in the outcome measured. Bars that do not cross 0 indicate statistical significance (P < .05).
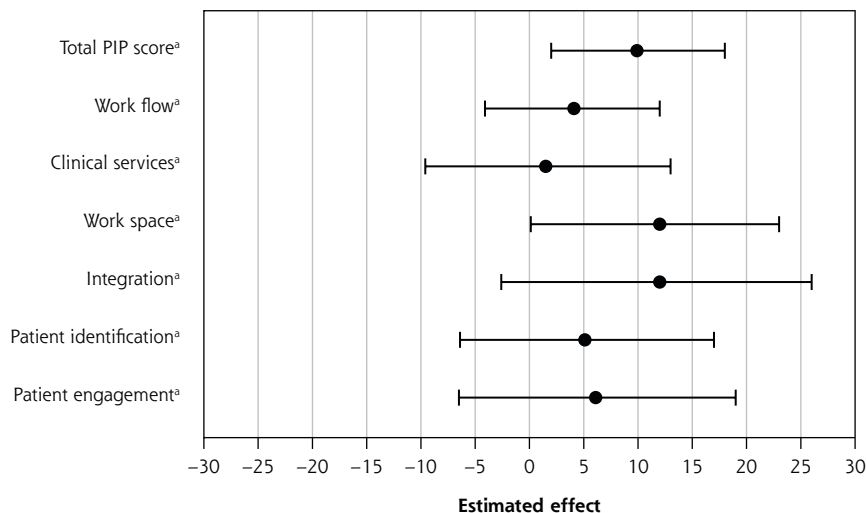
[a] Greater score indicates improved outcome.

personnel, and a pandemic. The research design allowed practices randomized to the control group to do QI work of their own (without access to the specific resources of the intervention) and allowed those in the active group to select the changes they perceived as most helpful. Too much flexibility of the intervention might have led to changes that were insufficiently strong to cause a clinic-wide effect on patient outcomes and integration. The resilience of practices in both groups, recruited based on their preexisting commitment to improve IBH, showed general improvement in integration but

might have biased the outcomes and contributed to null results for the intervention. In particular, interpreting how the COVID-19 pandemic affected IBH is complex and requires further study.

A limitation of the present study is that patients who agreed to participate might be systematically different from those who declined. In addition, researchers did not have access to individual patient electronic health records after enrollment to know which patients were provided BH services and whether those who received additional services fared differently than others. Future studies distinguishing patients by service need and service use will provide additional insights. Continued research that includes the effect of BH services with meaningful clinical and service data, and relevant patient-reported outcomes for patients with MCCs, is needed to guide selection of the best IBH strategies and models (eg, Collaborative Care Model).[7,57]

CONCLUSIONS

Although the specific practice redesign intervention studied did significantly improve the workflow of integrated care, it had little effect on other practice or patient outcomes. Greater practice IBH levels at baseline were consistently associated with better patient mental and physical function. Practices that completed the intervention improved integration and patient outcomes more than those that did not. Future practice-based research will examine relations associated with practice engagement in the intervention steps and other factors to understand how to improve outcomes via integrated medical and behavioral health services for those who need it the most—patients with MCCs.

👆 **Read or post commentaries in response to this article.**

Author affiliations: University of Vermont, Burlington, Vermont (B.L., J.C., A.M.C., L.N.B., P.C., K.C/K., S.F., J.H., J.J., K.S.K., J.B.O-L., L.W.N., R.G.P., D.P.,

Figure 4. Association of intervention completion in active group (n = 20) and change in practice integration and patient-reported outcomes.

Patient outcomes – PROMIS-29 domains

Anxiety
Depression
Fatigue
Sleep disturbance
Pain interference
Pain intensity
Social participation[a]
Physical function[a]
Physical health summary[a]
Mental health summary[a]

Estimated effect

Practice outcomes - PIP domains

Total PIP score[a]
Work flow[a]
Clinical services[a]
Work space[a]
Integration[a]
Patient identification[a]
Patient engagement[a]

Estimated effect

PIP = Practice Integration Profile; PROMIS-29 = Patient-Reported Outcomes Measurement Information System.

Note: The center point of each bar is the mean adjusted effect (regression coefficient) of the completion of the intervention on a specific outcome, with the bars representing 95% CI. Values >0 indicate that completion of the intervention was associated with an increase in the outcome measured. Bars that do not cross 0 indicate statistical significance ($P < .05$).

[a] Greater score indicates improved outcome.

Sciences Center, Oklahoma City, Oklahoma (Z.N.); DARTNet Institute, Aurora, Colorado (W.P.); National Committee for Quality Assurance, Washington, DC (S.H.S., J.S.).

🖐 Supplemental materials

References

1. Baird M, Blount A, Brungardt S, et al. The development of joint principles: integrating behavioral health care into the patient-centered medical home. *Ann Fam Med*. 2014;12(2):183-185. 10.1370/afm.1634

2. Johnson NB, Hayes LD, Brown K, Hoo EC, Ethier KA; Centers for Disease Control and Prevention (CDC). CDC National Health Report: leading causes of morbidity and mortality and associated behavioral risk and protective factors—United States, 2005-2013. *MMWR Suppl*. 2014;63(4):3-27.

3. Mokdad AH, Marks JS, Stroup DF, Gerberding JL. Actual causes of death in the United States, 2000. *JAMA*. 2004;291(10):1238-1245. 10.1001/jama.291.10.1238

4. Druss BG, Walker ER. Mental disorders and medical comorbidity. *Synth Proj Res Synth Rep*. 2011; (21):1-26.

5. Kathol RG, McAlpine D, Kishi Y, et al. General medical and pharmacy claims expenditures in users of behavioral health services. *J Gen Intern Med*. 2005;20(2):160-167. 10.1111/j.1525-1497.2005.40099.x

P.R., G.L.R., C.vE.); Parhelia Wellness, Santa Rosa, California (J.C.); University of Washington, Seattle, Washington (L-M.B., B.M., K.A.S.); Wellstar Health System, Marietta, Georgia (R.E.B.); Prama Psychology, PLLC, Richland, Washington (P.C.); Oregon Health & Science University, Portland, Oregon (D.J.C., E.N.W.); University of Colorado School of Medicine, Aurora, Colorado (F.V.D., R.K.); UMass Chan Medical School, Worcester, Massachusetts (L.E-P., D.M.); University of Houston, Houston, Texas (W.E.); University at Buffalo, Buffalo, New York (C.F.); University of California San Diego, San Diego, California (K.H., W.J.S.); The MetroHealth System, Cleveland, Ohio (D.C.K., T.S., K.T.); Case Western Reserve University, Cleveland, Ohio (D.C.K., K.C.S.); Stony Brook University, Stony Brook, New York (G.S.L.); Arizona State University, Tempe, Arizona (C.R.M., M.P.M.); Stanford University School of Medicine, Stanford, California (M.M.); University of Oklahoma Health

6. Melek S, Norris D. *Chronic Conditions and Comorbid Psychological Disorders.* Milliman; 2008. Accessed Aug 16, 2023. https://www.cbhc.org/wp-content/uploads/2015/11/chronic-conditions-and-comorbid-RR07-01-08.pdf

7. Kroenke K, Unutzer J. Closing the false divide: sustainable approaches to integrating mental health services into primary care. *J Gen Intern Med*. 2017; 32(4):404-410. 10.1007/s11606-016-3967-9

8. Balasubramanian BA, Cohen DJ, Jetelina KK, et al. Outcomes of integrated behavioral health with primary care. *J Am Board Fam Med*. 2017;30(2):130-139. 10.3122/jabfm.2017.02.160234

9. Reed SJ, Shore KK, Tice JA. Effectiveness and value of integrating behavioral health into primary care. *JAMA Intern Med*. 2016;176(5):691-692. 10.1001/jamainternmed.2016.0804

10. Bierman AS. Preventing and managing multimorbidity by integrating behavioral health and primary care. *Health Psychol*. 2019;38(9):851-854. 10.1037/hea0000787

11. Gilmer TP, Avery M, Siantz E, et al. Evaluation of the Behavioral Health Integration and Complex Care Initiative in Medi-Cal. *Health Aff (Millwood)*. 2018; 37(9):1442-1449. 10.1377/hlthaff.2018.0372

12. Loeb DF, Bayliss EA, Candrian C, deGruy FV, Binswanger IA. Primary care providers' experiences caring for complex patients in primary care: a qualitative study. *BMC Fam Pract*. 2016;17:34. 10.1186/s12875-016-0433-z

13. Bogner HR, Morales KH, de Vries HF, Cappola AR. Integrated management of type 2 diabetes mellitus and depression treatment to improve medication adherence: a randomized controlled trial. *Ann Fam Med*. 2012;10(1):15-22. 10.1370/afm.1344

14. Katon WJ, Lin EHB, Von Korff M, et al. Collaborative care for patients with depression and chronic illnesses. *N Engl J Med*. 2010;363(27):2611-2620. 10.1056/NEJMoa1003955

15. Lin EH, Von Korff M, Ciechanowski P, et al. Treatment adjustment and medication adherence for complex patients with diabetes, heart disease, and depression: a randomized controlled trial. *Ann Fam Med*. 2012;10(1):6-14. 10.1370/afm.1343

16. Stewart JC, Perkins AJ, Callahan CM. Effect of collaborative care for depression on risk of cardiovascular events: data from the IMPACT randomized controlled trial. *Psychosom Med*. 2014;76(1):29-37. 10.1097/psy.0000000000000022

17. Peek CJ and the National Integration Academy Council. *Lexicon for Behavioral Health and Primary Care Integration: Concepts and Definitions Developed by Expert Concensus*. AHRQ Publications No.13-IP001-EF. Agency for Healthcare Research and Quality; 2013. https://integrationacademy.ahrq.gov/sites/default/files/2020-06/Lexicon.pdf?_gl=1*1nqtq3h*_ga*NTU4NjQwODM5LjE2OTQ4OTQ2OTU.*_ga_45NDTD15CJ*MTY5NDkwNDMxMC4xLjAuMTY5NDkwNDMxMC42MC4wLjA

18. Macchi CR, Kessler R, Auxier A, et al. The Practice Integration Profile: rationale, development, method, and research. *Fam Syst Health*. 2016;34(4):334-341. 10.1037/fsh0000235

19. Archer J, Bower P, Gilbody S, et al. Collaborative care for depression and anxiety problems. *Cochrane Database Syst Rev*. 2012;10:CD006525. 10.1002/14651858.CD006525.pub2

20. Hunter CL, Funderburk JS, Polaha J, Bauman D, Goodie JL, Hunter CM. Primary Care Behavioral Health (PCBH) model research: current state of the science and a call to action. *J Clin Psychol Med Settings*. 2018;25(2):127-156. 10.1007/s10880-017-9512-0

21. Moriarty AS, Coventry PA, Hudson JL, et al. The role of relapse prevention for depression in collaborative care: a systematic review. *J Affect Disord*. 2020; 265:618-644. 10.1016/j.jad.2019.11.105

22. Cohen D, McDaniel RR Jr, Crabtree BF, et al. A practice change model for quality improvement in primary care practice. *J Healthc Manag*. 2004;49(3): 155-168; discussion 169-170.

23. Davis MM, Howk S, Spurlock M, McGinnis PB, Cohen DJ, Fagnan LJ. A qualitative study of clinic and community member perspectives on intervention toolkits: "unless the toolkit is used it won't help solve the problem". *BMC Health Serv Res*. 2017;17(1):497. 10.1186/s12913-017-2413-y

24. Hempel S, Miake-Lye I, Brega AG, et al. Quality improvement toolkits: recommendations for development. *Am J Med Qual*. 2019;34(6):538-544. 10.1177/1062860618822102

25. van Eeghen C, Littenberg B, Holman MD, Kessler R. Integrating behavioral health in primary care using lean workflow analysis: a case study. *J Am Board Fam Med*. 2016;29(3):385-393. 10.3122/jabfm.2016.03.150186

26. Perry CK, Damschroder LJ, Hemler JR, Woodson TT, Ono SS, Cohen DJ. Specifying and comparing implementation strategies across seven large implementation interventions: a practical application of theory. *Implement Sci*. 2019; 14(1):32. 10.1186/s13012-019-0876-4

27. Powell BJ, Waltz TJ, Chinman MJ, et al. A refined compilation of implementation strategies: results from the Expert Recommendations for Implementing Change (ERIC) project. *Implement Sci*. 2015;10:21. 10.1186/s13012-015-0209-1

28. Kessler R, Glasgow RE. A proposal to speed translation of healthcare research into practice: dramatic change is needed. *Am J Prev Med*. 2011;40(6):637-644. 10.1016/j.amepre.2011.02.023

29. Thorpe KE, Zwarenstein M, Oxman AD, et al. A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. *J Clin Epidemiol*. 2009;62(5):464-475. 10.1016/j.jclinepi.2008.12.011

30. Crocker AM, Kessler R, van Eeghen C, et al. Integrating Behavioral Health and Primary Care (IBH-PC) to improve patient-centered outcomes in adults with multiple chronic medical and behavioral health conditions: study protocol for a pragmatic cluster-randomized control trial. *Trials*. 2021;22(1):200. 10.1186/s13063-021-05133-8

31. Jewiss J, Natkin LW, Clark/Keefe K, Crocker A, Welkowitz JA. Integrating behavioural health and primary care: qualitative findings on contextual factors that influence integration. *Fam Pract*. 2023 Mar 21:cmad029. 10.1093/fampra/cmad029

32. Hitt JR, Brennhofer SA, Martin MP, et al. Further experience with the Practice Integration Profile: a measure of behavioral health and primary care integration. *J Clin Psychol Med Settings*. 2022;29(2):274-284. 10.1007/s10880-021-09806-z

33. Kessler RS, Auxier A, Hitt JR, et al. Development and validation of a measure of primary care behavioral health integration. *Fam Syst Health*. 2016;34(4): 342-356. 10.1037/fsh0000227

34. Mullin DJ, Hargreaves L, Auxier A, et al. Measuring the integration of primary care and behavioral health services. *Health Serv Res*. 2019;54(2):379-389. 10.1111/1475-6773.13117

35. Cella D, Riley W, Stone A, et al; PROMIS Cooperative Group. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005-2008. *J Clin Epidemiol*. 2010;63(11):1179-1194. 10.1016/j.jclinepi.2010.04.011

36. Hays RD, Spritzer KL, Schalet BD, Cella D. PROMIS®-29 v2.0 profile physical and mental health summary scores. *Qual Life Res*. 2018;27(7):1885-1891. 10.1007/s11136-018-1842-3

37. Mercer SW, Maxwell M, Heaney D, Watt GC. The consultation and relational empathy (CARE) measure: development and preliminary validation and reliability of an empathy-based consultation process measure. *Fam Pract*. 2004; 21(6):699-705. 10.1093/fampra/cmh621

38. Morisky DE, Green LW, Levine DM. Concurrent and predictive validity of a self-reported measure of medication adherence. *Med Care*. 1986;24(1): 67-74. 10.1097/00005650-198601000-00007

39. MacLean CD, Littenberg B, Gagnon M, Reardon M, Turner PD, Jordan C. The Vermont Diabetes Information System (VDIS): study design and subject recruitment for a cluster randomized trial of a decision support system in a regional sample of primary care practices. *Clin Trials*. 2004;1(6):532-544. 10.1191/1740774504cn051oa

40. Adams PF, Hendershot GE, Marano MA; Centers for Disease Control and Prevention/National Center for Health Statistics. Current estimates from the National Health Interview Survey, 1996. *Vital Health Stat 10*. 1999;(200):1-203.

41. Hlatky MA, Boineau RE, Higginbotham MB, et al. A brief self-administered questionnaire to determine functional capacity (the Duke Activity Status Index). *Am J Cardiol*. 1989;64(10):651-654. 10.1016/0002-9149(89)90496-7

42. Rose GL, Bonnell LN, O'Rourke-Lavoie JB, et al. Development and validation of the patient centeredness index for primary care. *J Clin Nurs*. 2022;31(23-24):3485-3497. 10.1111/jocn.16177

43. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. *JAMA*. 1999;282(18):1737-1744. 10.1001/jama.282.18.1737

44. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med*. 2006;166(10): 1092-1097. 10.1001/archinte.166.10.1092

45. Bime C, Wei CY, Holbrook JT, Sockrider MM, Revicki DA, Wise RA. Asthma symptom utility index: reliability, validity, responsiveness, and the minimal important difference in adult asthmatic patients. *J Allergy Clin Immunol*. 2012; 130(5):1078-1084. 10.1016/j.jaci.2012.07.058

46. Revicki DA, Leidy NK, Brennan-Diemer F, Sorensen S, Togias A. Integrating patient preferences into health outcomes assessment: the multiattribute Asthma Symptom Utility Index. *Chest*. 1998;114(4):998-1007. 10.1378/chest.114.4.998

47. Dennis ML, Chan YF, Funk RR. Development and validation of the GAIN Short Screener (GSS) for internalizing, externalizing and substance use disorders and crime/violence problems among adolescents and adults. *Am J Addict*. 2006; 15(Suppl 1):80-91. 10.1080/10550490601006055

48. Morean ME, DeMartini KS, Foster D, et al. The Self-Report Habit Index: assessing habitual marijuana, alcohol, e-cigarette, and cigarette use. *Drug Alcohol Depend*. 2018;186:207-214. 10.1016/j.drugalcdep.2018.01.014

49. Hays RD, Spritzer KL, Fries JF, Krishnan E. Responsiveness and minimally important difference for the patient-reported outcomes measurement information system (PROMIS) 20-item physical functioning short form in a prospective observational study of rheumatoid arthritis. *Ann Rheum Dis*. 2015;74(1):104-107. 10.1136/annrheumdis-2013-204053

50. Yost KJ, Eton DT, Garcia SF, Cella D. Minimally important differences were estimated for six Patient-Reported Outcomes Measurement Information System-Cancer scales in advanced-stage cancer patients. *J Clin Epidemiol*. 2011;64(5):507-516. 10.1016/j.jclinepi.2010.11.018

51. Abdi H. The Bonferonni and Šidák corrections for multiple comparisons. In: Salkind N, ed. *Encyclopedia of Measurement and Statistics*. Sage; 2007.

52. Centers for Disease Control and Prevention. *Ambulatory Health Care Data*. Last reviewed Dec 30, 2021. Accessed Aug 16, 2023. https://www.cdc.gov/nchs/ahcd/about_ahcd.htm

53. Cohen DJ, Sweeney SM, Miller WL, et al. Improving smoking and blood pressure outcomes: the interplay between operational changes and local context. *Ann Fam Med*. 2021;19(3):240-248. 10.1370/afm.2668

54. Keddem S, Agha AZ, Long JA, Shasha B, Hausmann LRM, Shea JA. The gears of knowledge translation: process evaluation of the dissemination and implementation of a patient engagement toolkit. *J Gen Intern Med*. 2020;35(Suppl 2):808-814. 10.1007/s11606-020-06099-0

55. Perry CK, Lindner S, Hall J, Solberg LI, Baron A, Cohen DJ. How type of practice ownership affects participation with quality improvement external facilitation: findings from EvidenceNOW. *J Gen Intern Med*. 2022;37(4):793-801. 10.1007/s11606-021-07204-7

56. Walunas TL, Ye J, Bannon J, et al. Does coaching matter? Examining the impact of specific practice facilitation strategies on implementation of quality improvement interventions in the Healthy Hearts in the Heartland study. *Implement Sci*. 2021;16(1):33. 10.1186/s13012-021-01100-8

57. Kathol RG, Hobbs Knutson K, Dehnel PJ. *Physician's Guide: Understanding and Working With Integrated Case Managers*. 1st ed. Springer; 2016.