Psychometric assessment of the *Primary Care Behavioral Health Provider Adherence Questionnaire (PPAQ)*

Gregory P Beehler, PhD, MA,^{1,2,3} Jennifer S Funderburk, PhD,^{4,5,6} Kyle Possemato, PhD,^{4,5} Katherine M Dollar, PhD^{1,3}

¹VA Center for Integrated Healthcare,
VA WNY Healthcare System, 3495
Bailey Ave, Buffalo, NY 14215, USA

²School of Nursing,
University at Buffalo, The State
University of New York, Buffalo,
NY, USA

³School of Public Health and
Health Professions,
University at Buffalo, The State
University of New York, Buffalo,
NY, USA

⁴VA Center for Integrated
Healthcare,
Syracuse VA Medical Center,
Syracuse, NY, USA

⁵Department of Psychology,
Syracuse University, Syracuse,
NY, USA

⁶Department of Psychiatry,
University of Rochester, Rochester,
NY, USA

Correspondence to: G P Beehler
gregory.beehler@va.gov

Cite this as: *TBM* 2013;3:379–391
doi: 10.1007/s13142-013-0216-1

Abstract

Adherence to protocol among behavioral health providers working in co-located, collaborative care or Primary Care Behavioral Health settings has rarely been assessed due to limited measurement options. Development of psychometrically sound measures of provider fidelity may improve the translation of these service delivery models into every day practice. One hundred seventy-three integrated behavioral health providers in VA primary care clinics responded to an online questionnaire to assess the reliability and validity of the Primary Care Behavioral Health Provider Adherence Questionnaire (PPAQ). Psychometric assessment resulted in a reliable 48-item measure with two subscales that specified essential and prohibited provider behaviors. The PPAQ demonstrated strong convergent and divergent validity when compared to another measure of health care integration. Known-group comparisons provided partial support for criterion validity. The PPAQ is a reliable and valid self-report of behavioral health provider fidelity with implications for improving provider training, program monitoring, and clinical research.

Keywords

Guideline adherence, Mental health services, Primary health care, Program evaluation

INTRODUCTION

Behavioral health is being recognized as a crucial component of the Primary Care Medical Home [1]. A variety of health service delivery models integrating behavioral health into primary care have been described to date [2]. One example is the Primary Care-Mental Health Integration (PC-MHI) program within the Veterans Healthcare Administration [3–5] which consists of co-located, collaborative care (CCC) and care management. CCC embeds behavioral health providers (BHPs) within primary care clinics to support the primary care team in the identification and treatment of common behavioral and psychiatric conditions [6]. BHPs provide functional assessments, brief interventions, and referral to specialty mental health services [5]. Care management employs protocol-driven, diagnosis-specific interventions often delivered by a nurse to

Implications

Policy: Administrators tasked with development of policies and procedures related to health care integration might consider employing the PPAQ to improve uniformity in behavioral health service provision within medical home models.

Research: The application of the PPAQ to future effectiveness research can begin to address the paucity of literature regarding if and how behavioral health provider protocol adherence is associated with clinical outcomes.

Practice: When used for training and program development initiatives, the PPAQ has the potential to improve behavioral health provider capacity to deliver evidence-based, patient-centered care consistent with the medical home philosophy.

support the primary care providers' (PCPs) interventions for conditions such as depression [5].

CCC is just one name for a group of highly similar models of integrated health care outside of the VA, including Primary Care Behavioral Health which has become widely used in other large healthcare systems such as the United States Air Force and Kaiser Permanente [7, 8]. Despite widespread adoption across health care systems, the evidence for CCC (and Primary Care Behavioral Health) has been slow to accumulate. Key findings from studies conducted in the Department of Defense family medicine clinics have suggested that patients show significant decreases in mental health symptoms and improved functioning in relation to brief treatment of less than four encounters [9, 10]. Most recently, research has shown that outcomes associated with these brief treatments are sustained at two years following treatment [11]. Despite the value of such foundational works regarding CCC effectiveness, these observational studies have lacked control groups, taken place in a limited number of clinic locations, and have not considered how BHP behavior or context factors might act as moderators of patient outcomes.

Alongside the paucity of literature regarding effectiveness, CCC implementation has also rarely

been described. With the rapid expansion of CCC and related models across federal and non-federal health care systems and little evidence-based guidance on how to implement CCC, there is potential for substantial heterogeneity in the manner by which CCC is delivered. Attention to fidelity, or the degree to which providers implement programs as intended [12], is vital in ensuring that interventions are delivered in an optimal manner. Although system resource constraints and local contextual factors can be highly influential, protocol adherence, or a provider's utilization of specific procedures and engagement in specific tasks and activities [13], is a crucial measure of implementation fidelity. A major challenge to assessing CCC fidelity, however, is that a psychometrically sound and provider-friendly measure of protocol adherence has not been available.

To address the lack of appropriate instrumentation to quantify BHP adherence to CCC, we recently developed the *Primary Care Behavioral Health Provider Adherence Questionnaire* (PPAQ) as the first self-report measure of BHP protocol adherence to CCC models [14]. In sum, the PPAQ was developed following a fidelity of implementation approach [12, 13, 15] in which 56 items reflecting BHP protocol adherence were generated from a review of scientific literature, unpublished clinical practice guidelines from clinics using a CCC platform, clinical and administrative experience of the research team, and informal polling of CCC content experts. The draft survey was evaluated by 25 CCC experts from VA, Department of Defense, and community and academic health center settings using modified Delphi methodology. Experts rated each item as reflecting a behavior that was *essential* (i.e., considered to be highly reflective of the CCC model and required for good practice), *compatible* (i.e., acceptable when working in a CCC model, but not required or specific to CCC), or *prohibited* (i.e., behaviors that should clearly be avoided because they are inconsistent with CCC practice). Across three rounds of data collection, a total of 54 items (93 %) met a high level of consensus (≥ 80 %) among experts. The final items were compiled, and a five-point Likert-type response scale ranging from "never" to "always" was applied. Although several approaches to scoring were considered, we chose to apply the simplest approach of summing items into three subscales: PPAQ-E (38 essential items), PPAQ-C (6 compatible items), and PPAQ-P (10 prohibited items).

The use of the Delphi method to develop the PPAQ ensured that appropriate content validity was achieved [16], but the PPAQ had yet to be subjected to further psychometric assessment. Therefore, this paper reports the findings of the current study aimed at assessing the reliability and validity of the PPAQ. Our ultimate goals were to provide the instrumentation necessary to (a) improve training, ongoing implementation, and sustainability of CCC in VA and non-VA settings and (b) quantify BHP fidelity for use in CCC effectiveness studies. Due to the

challenge of identifying the large number of BHPs necessary for some approaches to validation, such as factor analysis, we identified methods for assessing criterion-related validity (known-groups comparisons and convergent and discriminant validity) as viable options. The known-groups method assesses differences in mean scores across groups hypothesized to score low and high on a trait or factor [17]. Hypotheses are stated a priori regarding the predicted direction of scores across groups. To identify our comparison groups for this study, we generated six hypotheses in relation to conceptually important provider and setting characteristics from available literature. Note that no single hypothesis carried greater weight than the others.

Hypothesis 1 Although CCC and care management roles can overlap within PC-MHI, we felt that the disease-focused, algorithm-based procedures that guide care management [18, 19] would provide a logical comparison group for CCC providers. Our first hypothesis was that higher fidelity would be demonstrated among BHPs who endorsed working exclusively in the CCC role compared to those working in either care management-only or CCC and care management combined.

Hypothesis 2 Implementation of PC-MHI may vary across VA medical centers (VAMCs) and community-based outpatient primary care clinics (CBOCs) in relation to the range of mental health services offered and the frequency of routine mental health screening conducted [20]. Assuming there are fewer resources available at CBOCs compared to VAMCs which may result in BHPs being called upon to provide a wide scope of services that may not coincide with CCC principles, we hypothesized that BHPs working in CCC settings at VAMCs would evidence higher fidelity scores than BHPs working in CCC settings at VA CBOCs.

Hypothesis 3 Although integration of BHPs in VA primary care was relatively recent [4], we hypothesized that BHPs with a longer history of providing CCC services would have higher fidelity ratings due to mastery over CCC-specific roles and duties compared to those BHPs with a shorter history of providing CCC services.

Hypotheses 4 and 5 In a recent qualitative study, Beehler and Wray [21] found local models of CCC were highly influenced by two factors: BHP knowledge and skills related to CCC and BHP perception of local resource availability to

enact CCC. We hypothesized that BHPs who self-reported higher levels of knowledge and skills related to CCC would show higher levels of fidelity. We also hypothesized that those BHPs who self-reported greater access to material and organizational resources to deliver CCC would also show greater fidelity.

Hypothesis 6 Cognitive-behavioral therapy (CBT) is a time-limited, structured approach that is highly compatible with CCC, and the majority of techniques recommended for delivery in CCC settings are consistent with CBT [22]. Thus, we hypothesized that BHPs who endorse a CBT orientation over other forms of psychotherapy would demonstrate higher fidelity.

The impact of BHP discipline on CCC practice has not been described in the literature previously. Given that not all BHPs engage in standardized training for CCC, discipline-specific practice patterns may be an influential source of variability in behavior with unclear impact on fidelity. Therefore, as a seventh comparison, we examined the associations between BHP discipline (psychologist, social worker, psychiatrist, and nurse) and fidelity score without stating a specific hypothesis.

Finally, to assess convergent and discriminant validity, we aimed to compare PPAQ scores with the scores from the Level of Integration Measure [23]. The Level of Integration Measure (LIM) is a self-report measure of the degree of clinical integration in CCC settings. Although not specifically designed as a measure of fidelity, the LIM taps into related constructs, such as provider clinical practices, beliefs, and commitment to CCC. We predicted that higher fidelity scores on the PPAQ would be positively correlated with higher LIM scores. We additionally predicted that higher endorsement of PPAQ prohibited items would be inversely correlated with LIM scores.

METHODS

Participants and procedures

The conduct of this study was approved by the VA Western New York Healthcare System Institutional Review Board. Study participants were recruited from May through July of 2012. Our aim was to invite VA BHPs who provided clinical services in primary care for at least 25 % of their duties, had an active VA email account, and with sufficient time to complete a brief (i.e., <20 min) online survey. Because there was no readily accessible national list of VA PC-MHI providers who met these criteria at the time of recruitment, potential BHPs were identified from two alternative sources: individuals

who either participated previously in PC-MHI-related research conducted by the study team or individuals who attended one of several PC-MHI educational trainings led by the study team's research center from 2009 to 2011. By combining these lists and removing duplicates or individuals known to no longer be in a BHP role in PC-MHI, we identified 580 individuals to receive invitations to participate.

BHPs were sent an email invitation that briefly introduced the purpose of the study and nature of their participation. BHPs who did not respond to the initial invite were automatically sent two additional reminders over a two-week period requesting participation, unless they declined participation by either contacting the research staff or clicking the hyperlink embedded in the email to "opt out" of participation. BHPs who wished to participate followed the link in the invitation to complete the online survey. After reviewing consent information via the web-based portal, participants provided basic background information and completed the survey. To maintain anonymity, no personal identifiers were requested from participants. However, upon completing the survey, all participants were offered an educational text as a thank you. One hundred seventy-three BHPs provided complete responses to the survey, yielding a 30 % response rate which is on par with other studies using electronic surveys [24].

MATERIALS

BHP Background Questionnaire—Developed for this study, this background questionnaire requested information regarding relevant professional background characteristics of participants, including the following: primary clinical role (CCC, care management, both CCC and care management, or other), main practice setting (VAMC v. CBOC), time in current role, psychotherapy theoretical orientation (e.g., cognitive-behavioral, person-centered, etc.), and educational background (psychologist, social worker, psychiatrist, or nurse). BHPs who endorsed being in a CCC role were also asked to rate their level of agreement on a Likert scale ranging from 0 ("strongly disagree") to 3 ("strongly agree") two additional items: "I have adequate knowledge and skills to provide care for Veterans that is consistent with the Co-located, Collaborative Care model," and "I have access to adequate resources to provide care for Veterans that is consistent with the Co-located, Collaborative Care model."

PPAQ—The PPAQ, described above, is a 54-item scale of BHP protocol adherence to the CCC model [14]. This instrument uses a five-point Likert-type response scale ranging from "never" to "always." The PPAQ is scored by summing items into three subscales: PPAQ-E (38 essential items), PPAQ-C (6 compatible items), and PPAQ-P (10 prohibited items). Higher PPAQ-E scores indicate high fidelity whereas

higher PPAQ-P scores suggest areas of low fidelity (i.e., frequent engagement in prohibited behaviors). Higher scores on the PPAQ-C are interpreted in relation to the PPAQ-E and PPAQ-P scores. For instance, if higher PPAQ-C scores accompany poor fidelity scores on the other subscales, BHPs might need guidance in reorienting their focus from compatible behaviors to essential behaviors.

LIM—Recently developed as part of an evaluation of the dissemination of integrated primary care across four clinics in New Hampshire [23], the LIM measures the level of behavioral and physical health integration in primary care. This scale was developed based on current literature, consultation with experts in the field, and feedback from clinicians. Respondents rate the 35 items on a five-point Likert-type scale from “strongly disagree” to “strongly agree.” Higher scores on the LIM reflect higher levels of integration. The LIM also includes six subscale domains: clinic system integration, beliefs and commitment, clinical practice, interdisciplinary alliance, training and consultation, and leadership. Preliminary psychometric evaluation was conducted with the LIM using data collected from 118 multidisciplinary primary care team members across four non-VA-integrated care clinics. Reliability assessment confirmed strong internal consistency for the full 35-item LIM (Cronbach’s alpha of 0.95) and adequate internal consistency for its six subscales (alphas ranging from 0.71 to 0.92). Although construct validity of the LIM has not been formally assessed to date, it was developed for use in integrated care settings highly similar to that of the VA BHPs in this study, making it the best available option given the dearth of instruments specific to these health care settings.

Psychometric analysis and statistical procedures

We assessed reliability and validity of the PPAQ based on procedures described by Spector [25] and Netemeyer and colleagues [17]. Item analysis was first conducted using item-based statistics and calculation of internal consistency of the PPAQ subscales. Internal consistency was assessed via Cronbach’s coefficient alpha [26]. Guidelines for interpreting coefficient alpha suggest that coefficients in the range of 0.70 or greater are acceptable [25]. We first calculated alpha coefficients for each subscale (PPAQ-E, PPAQ-C, and PPAQ-P). Corrected item-total statistics were evaluated with correlations of <0.30 indicating items that were candidates for elimination [25, 27]. However, we retained items that were deemed critical for preservation of content validity if elimination resulted in only marginal improvements in reliability.

The hypotheses that drove the known-group comparisons between PPAQ subscale scores and participants’ self-reported background characteristics were assessed using independent sample *t* tests for dichotomous variables and ANOVA for categorical variables. The Scheffe post hoc test applied to significant omnibus *F* tests. After computing Cronbach’s alpha for the LIM

to confirm that this measure generated reliable observations in the current study, Pearson’s *r* was used to assess the correlation between the PPAQ and the LIM.

RESULTS

Participant characteristics

As shown in Table 1, most participants identified being in a CCC-only role (48 %) or a CCC role combined with care management (43 %). Just over 1 % of respondents were in care manager-only roles. Participants reported the number of years in their current PC-MHI role as follows: two or fewer years (29 %), three to four years (36 %), and five or more years (35 %). The single largest group of BHPs was doctoral level psychologists (46 %), and the majority of BHPs endorsed a CBT orientation (60 %). Among participants endorsing a CCC role, about 26 % felt they lacked the *resources* to provide CCC. In contrast, 94 % agreed with the statement that they had adequate *knowledge and skills* to provide CCC. Although this was encouraging news regarding BHPs’ perceptions of CCC-related skills, because of limited variability, we were unable to include this variable in subsequent analyses.

Reliability analysis

Table 2 displays the PPAQ item means, standard deviations, and corrected item-total statistics. The coefficient alpha for the PPAQ-E suggested excellent internal consistency (alpha=0.92) with only two items (PPAQ2, PPAQ5) identified as sources of low reliability. Because removal of these items did not appreciably increase the PPAQ-E reliability, we chose to retain them to preserve content validity. The potential range of scores on the PPAQ-E was 38 to 190, with higher scores indicating greater fidelity. The current sample showed moderate levels of fidelity to essential elements of CCC with a mean of 144 (SD=19). In contrast, the coefficient alpha for the PPAQ-C was poor at 0.36. Most items had low corrected item-total coefficients, yet removal of the lowest items did not sufficiently improve the overall alpha. Thus, we chose to eliminate the six-item PPAQ-C from further psychometric analysis. The PPAQ-P achieved acceptable reliability with a coefficient alpha of 0.70. Similar to the PPAQ-E, two items (PPAQ17, PPAQ50) in the PPAQ-P had low correlations, but removal did not sufficiently improve the overall alpha. Thus, these items were retained. The potential range of scores on the PPAQ-P was 10 to 50 with higher scores indicating lower fidelity (i.e., engaging in prohibited items with greater frequency). The current sample reported a low to moderate level of prohibited elements with a mean of 21 (SD=5).

Known-groups validity

Table 3 displays the findings from known-groups analyses. We were unable to assess our first hypothesis as intended regarding PC-MHI role due to the very low number of participants who were in care manager-only roles. Alternatively, in comparing those BHPs in CCC-only roles to all other BHPs, we failed to find a

Table 1 | Background characteristics of survey participants, $n=173$

Characteristic	<i>n</i> (%)
Current PC-MHI role	
CCC-only	83 (48)
Both CCC and care management	75 (43)
Care management only	2 (1)
Other ^a	13 (8)
Years in current PC-MHI role	
2 or fewer	50 (28.9)
3–4	62 (35.8)
5 or more	61 (35.3)
Educational background	
PhD/PsyD	80 (46)
MSW/LCSW	41 (24)
MD/DO	19 (11)
NP/APN	19 (11)
RN	14 (8)
Psychotherapy theoretical orientation	
Cognitive behavioral	103 (60)
Humanistic or client-centered	25 (15)
Interpersonal	14 (8)
Psychoanalytic/psychodynamic	9 (5)
Systems or family systems	5 (3)
Other ^b	17 (10)
I have adequate knowledge and skills to provide CCC ^c	
Strongly disagree	0 (0)
Disagree	9 (6)
Agree	63 (41)
Strongly agree	81 (53)
I have adequate resources to provide CCC ^c	
Strongly disagree	0 (0)
Disagree	39 (26)
Agree	69 (45)
Strongly agree	45 (29)

^a Includes respondents who did not consider themselves to be specifically PC-MHI staff^b Includes those respondents endorsing eclectic orientations or a wide variety of single therapeutic orientations (i.e., existential)^c Responses are from those participants who endorsed CCC-only or both CCC and care management roles ($n=153$)

significant difference between groups. Likely, the lack of difference was due to the majority of BHPs in the comparison group engaged in both CCC and care management roles. In terms of practice setting, only PPAQ-P scores varied significantly, with higher scores (i.e., lower fidelity) among providers who were practicing outside of VAMCs, as predicted. Although there was no difference between participants on PPAQ-P scores in relation to psychotherapy orientation, those who endorsed a CBT orientation averaged eight points higher on the PPAQ-E, yielding significantly higher fidelity scores among this group. Associations between time in role and the PPAQ were limited to a nonsignificant trend for higher PPAQ-P scores with longer time in role. BHPs with a social work background had significantly higher PPAQ-P

scores than psychologists, physicians, or nurses. Among BHPs in CCC-only or combined with care management roles, those who strongly agreed that they had sufficient resources to enact a CCC model showed higher PPAQ-E scores.

Convergent and discriminant validity

Table 4 displays the results of correlations between the PPAQ-E, PPAQ-P, and LIM total and subscales. The nonsignificant correlation between the PPAQ-E and PPAQ-P ($r=0.10$, $p=.197$) reinforced that these two subscales are distinct dimensions of the PPAQ and that BHP protocol adherence is comprised of both essential and prohibited behaviors. As predicted, the PPAQ-E was significantly moderately corre-

Table 2 | PPAQ item means (SD) and corrected item-total correlations for PPAQ-E, PPAQ-C, and PPAQ-P items subscales

Item	Item content	Mean (SD)	Corrected item-total correlation
Essential Items Subscale (PPAQ-E)			
1	During clinical encounters with patients, I see patients for 30 min or less.	3.43 (0.83)	0.33
2	I manage patients reporting mild and moderate symptoms in primary care, and I refer those with more severe symptoms to specialty mental health services when possible.	3.85 (1.12)	0.27
3	During patient appointments, I discuss barriers to implementing a plan or adhering to treatment recommendations.	4.02 (0.85)	0.43
5	I accept referrals for patients with common mental health problems (i.e., depression, anxiety, etc.).	4.64 (0.79)	0.29
6	During clinical encounters with a patient, I implement behavioral and/or cognitive interventions.	4.21 (0.79)	0.35
7	In introducing my role in the clinic to patients, I explain that I want to get an idea of what is and what is not working for the patient and then together develop a plan to help them manage their concerns.	4.19 (0.86)	0.58
8	During clinical encounters with patients, I triage patients to determine if they can be treated in primary care or should be referred to a specialty mental health or a community agency.	4.35 (1.07)	0.41
9	I accept referrals for patients who might benefit from brief, targeted behavioral health interventions for chronic pain.	3.73 (1.31)	0.46
10	I accept referrals for patients who might benefit from brief, targeted behavioral health interventions for adjustment to illness (i.e., diabetes, heart disease, spinal cord injury, TBI, etc.).	3.74 (1.25)	0.54
11	My progress notes in the shared medical record include focused recommendations for the primary care provider and/or primary care team.	3.46 (0.99)	0.49
12	I meet briefly with primary care staff as a team to provide both a behavioral health perspective and behavioral data.	2.75 (1.07)	0.48
13	My progress notes include focused recommendations for the patient.	4.27 (0.82)	0.60
14	I consult with various members of the primary care team (i.e., pharmacist, dietician) in addition to the primary care provider about behavioral aspects of medical conditions (i.e., medications that cause nightmares).	3.29 (1.03)	0.52
15	At follow-up encounters with patients, I inquire about progress on goals or action plans set at the previous appointment.	4.43 (0.73)	0.51
16	I administer one or more brief validated measures (e.g., Patient Health Questionnaire-9 or PHQ-9) for an initial screening of symptoms of interest, or I review these findings if measures were administered by other primary care staff.	3.90 (1.18)	0.35
19	Following patient appointments, I provide feedback to primary care providers (based on their preferred method of communication) within 1 business day of an initial appointment.	3.73 (1.07)	0.56
20	During clinical encounters with patients, I clarify, confirm, and discuss the patient's concerns.	4.74 (0.56)	0.40
21	My progress notes include brief clinical impressions of the patient's presenting problem(s).	4.60 (0.70)	0.38
24	In introducing my role in the clinic to patients, I explain that our appointments typically will be 30 min or less.	3.10 (1.31)	0.44
25	I provide suicide risk assessment for primary care patients in crisis and refer to a higher level of care as indicated.	4.45 (0.87)	0.31

Table 2 | (continued)

Item	Item content	Mean (SD)	Corrected item-total correlation
27	During patient appointments, I use local community resources to assist me in meeting the behavioral health needs of patients.	3.15 (0.90)	0.31
28	I provide education to the primary care team or PACT staff on behavioral health issues (e.g., presentations and handouts).	2.87 (1.02)	0.48
29	I provide advice to primary care team or PACT staff about appropriate referrals to specialty behavioral health services.	3.53 (0.96)	0.59
32	My progress notes include findings from functional assessments and brief screening instruments.	3.53 (0.92)	0.35
33	I administer one or more brief validated measures (e.g., Patient Health Questionnaire-9 or PHQ-9) for follow-up screening of symptoms of interest, or I review these findings if measures were administered by other primary care staff.	3.46 (1.13)	0.40
34	I routinely consult with PCPs to increase my knowledge about behavioral aspects of medical conditions, such as the role of anxiety in cardiac distress.	3.09 (0.97)	0.50
37	Following clinical encounters with patients, I continue to provide feedback to the PCP about follow-up appointments when needed.	3.49 (1.04)	0.60
38	During clinical encounters with patients, I work with the patient to develop a specific plan to address their presenting problem and document this plan.	4.33 (0.76)	0.49
39	I accept referrals for patients who need lifestyle interventions (e.g., tobacco cessation, weight control, stress management).	3.72 (1.24)	0.40
40	I accept referrals for patients in need of behavioral health interventions for medication issues (i.e., adherence).	3.66 (1.23)	0.51
43	I accept referrals for patients in need of behavioral health interventions for adjustment to aging and issues specific to older patients.	3.50 (1.06)	0.41
44	During a patient appointment, I provide functional assessment, focused intervention, and address disposition.	3.92 (0.96)	0.42
46	I accept referrals for patients from PCPs as a warm hand off (i.e., the PCP introduces me to the Veteran).	3.82 (1.06)	0.47
47	In introducing my role in the clinic to patients, I explain that I work with the PCPs in situations where good health care involves paying attention to physical health, habits, behaviors, emotional health, and how those things interact.	3.77 (1.07)	0.65
51	During clinical encounters with patients, I address the PCPs' reason for referral.	4.76 (0.56)	0.46
52	I employ strategies to identify and prevent exacerbation of at-risk, sub-syndromal behaviors and symptoms.	3.96 (0.87)	0.59
53	I provide information regarding a patient's symptoms and functioning to assist primary care providers (and/or clinical pharmacists, primary care psychiatrists, psychiatric nurse practitioners) in initiating or modifying common psychotropic medications, such as antidepressants.	3.88 (1.02)	0.55
54	I participate in primary care-based clinical pathways for common health conditions, such as chronic pain or comorbid depression and cardiovascular disease.	2.45 (1.33)	0.39
Compatible Items Subscale (PPAQ-C)			
4	I collaborate with primary care team staff to provide group medical visits (or shared medical appointments)	2.32 (1.35)	0.29

Table 2 | (continued)

Item	Item content	Mean (SD)	Corrected item-total correlation
	to patients.		
18	During patient appointments, I promote emotional exploration.	3.72 (0.87)	0.09
30	I conduct follow-up appointments via telephone when appropriate.	3.42 (1.09)	0.26
35	During a patient appointment, I provide supportive interventions without addressing cognitive or behavioral change.	2.60 (0.95)	0.11
42	Following patient appointments, I typically schedule follow-ups at least 2 weeks apart.	3.61 (0.88)	0.10
45	I provide brief psycho-educational groups or classes on specific topics (such as mood management, stress reduction, etc.).	2.73 (1.38)	0.16
Prohibited Items Subscale (PPAQ-P)			
17	It takes 30 min or more for me to complete all documentation following the initial appointment.	3.27 (1.11)	0.18
22	During a clinical encounter with a patient, I provide full neuropsychological, cognitive, or personality assessments.	1.64 (0.88)	0.35
23	I see patients for weekly, open-ended therapy.	1.66 (0.83)	0.44
26	I typically see patients for 50-min appointments.	2.99 (1.05)	0.44
31	I provide family or couples therapy for 10 or more appointments per episode of care.	1.49 (0.71)	0.41
36	During a patient appointment, I provide full length empirically supported treatments, such as prolonged exposure or dialectical behavior therapy.	1.69 (0.97)	0.41
41	I typically see patients for 10 or more appointments per episode of care.	2.45 (0.98)	0.41
48	I provide long-term (i.e., greater than 8 sessions) group psychotherapy.	1.64 (1.07)	0.32
49	I meet with a patient for greater than 50 min to gather a full psychosocial history and comprehensive psychiatric interview.	2.60 (1.23)	0.42
50	During a patient appointment, I typically provide medical social work services, including, but not limited to, assistance with disability claims, obtaining health insurance, and/or assisting with housing.	1.95 (1.04)	0.16

Alpha coefficients for subscales are as follows: PPAQ-E=0.92, PPAQ-C=0.36, and PPAQ-P=0.70

lated with the LIM total score ($r=0.52$, $p<0.001$) suggesting convergent validity across these two similar scales. Correlations between the PPAQ-E and LIM subscales range from 0.35 to 0.49, each highly significant ($p<0.001$). Furthermore, the PPAQ-P showed only a weak, nonsignificant correlation with the total LIM and LIM subscales.

DISCUSSION

Psychometric assessment of the PPAQ resulted in a 48-item, reliable measure of CCC BHP protocol adherence high in content validity. Assessment of

criterion validity received partial support using known-groups comparisons and strong support from convergent and divergent validity with the LIM. Our findings also provide empirical support for the conceptual frameworks that guided this study that emphasize measuring the dose of intervention content by first identifying its critical components. Although breaking down complex interventions and health care platforms into their constituent parts runs the risk of omissions or oversimplification, the benefit of this approach is that it can identify elements that should form the basis of more precise and rigorous fidelity measurement, an often

Table 3 | Known-groups comparisons of BHP and practice setting characteristics in relation to the PPAQ subscales

PPAQ subscale	PC-MHI role M (SD)		<i>t</i> (<i>df</i>)	<i>p</i>	Orientation, <i>M</i> (SD)		<i>t</i> (<i>df</i>)	<i>p</i>	Setting, <i>M</i> (SD)		<i>t</i> (<i>df</i>)	<i>p</i>	PPAQ subscale educational background, <i>M</i> (SD)	
	Other				Other				VAMC				PhD	
	CCC-only	Other			CBT					other				SW
Essential	143.2 (14.4)	144.3 (21.8)	0.415 (171)	0.679	147.2 (15.3)	138.7 (21.7)	−3.04 (171)	0.003	144.0 (19.9)	143.3 (16.1)	−23 (171)	0.812	143.5 (16.3)	145.0 (19.7)
Prohibited	21.4 (4.8)	21.4 (5.4)	0.020 (171)	0.984	21.4 (5.4)	21.3 (4.7)	−0.112 (171)	0.911	20.5 (4.9)	23.0 (5.1)	3.27 (171)	0.001	20.7 (4.9)	23.9 (5.02)
PPAQ subscale educational background, <i>M</i> (SD)	<i>F</i> (<i>df</i>)	<i>p</i>	Time in role, in years, <i>M</i> (SD)		<i>F</i> (<i>df</i>)		<i>p</i>	Adequate CCC resources, <i>M</i> (SD)		<i>F</i> (<i>df</i>)		<i>p</i>		
MD/nurse			≤2		3–4		5+	Disagree		Agree		Strongly agree		
143.2 (21.1)	0.118 (2, 170)	0.889	142.2 (17.4)	143.2 (16.8)	145.7 (21.2)	0.541 (2, 170)	0.583	145.8 (15.9)	143.2 (15.7)	151.1 (13.5)	3.66 (2, 150)	0.028 ^b		
20.4 (4.9)	6.85 (2, 170)	0.001 ^a	20.3 (4.8)	21.3 (5.1)	22.3 (5.3)	2.22 (2, 170)	0.112	22.2 (5.0)	21.6 (4.6)	20.5 (5.5)	1.19 (2, 150)	0.307		
The values presented in bold italics are p-values.														
Two-group comparisons assessed by <i>t</i> tests and multiple group comparisons assessed by analysis of variance with the Scheffe post hoc test applied to significant omnibus <i>F</i> tests														
<i>PC-MHI</i> Primary Care-Mental Health Integration, <i>CCC</i> co-located, collaborative care, <i>CBT</i> cognitive behavioral therapy, <i>VAMC</i> Veterans Affairs Medical Center, <i>SW</i> social worker														
SWaPhD, MD/nurse, <i>p</i> = 0.01														
^a Agree < strongly agree														

The values presented in bold italics are *p*-values.Two-group comparisons assessed by *t* tests and multiple group comparisons assessed by analysis of variance with the Scheffe post hoc test applied to significant omnibus *F* tests*PC-MHI* Primary Care-Mental Health Integration, *CCC* co-located, collaborative care, *CBT* cognitive behavioral therapy, *VAMC* Veterans Affairs Medical Center, *SW* social worker^a SW>PhD, MD/nurse, *p*=0.01^b Agree < strongly agree

Table 4 | Coefficient alphas and intercorrelations of the PPAQ and LIM

Measure	PPAQ-E	PPAQ-P	LIM	SIS	BCS	IPS	RSS	TSS	LSS
PPAQ-E	0.92								
PPAQ-P	0.10	0.70							
LIM	0.52*	0.06	0.94						
SIS	0.48*	0.03	0.88*	0.82					
BCS	0.41*	0.03	0.77*	0.66*	0.73				
IPS	0.49*	0.07	0.89*	0.77*	0.61*	0.83			
RSS	0.35*	0.11	0.75*	0.55*	0.41*	0.69*	0.93		
TSS	0.37*	−0.04	0.74*	0.49*	0.52*	0.58*	0.53*	0.74	
LSS	0.38*	0.11	0.74*	0.50*	0.60*	0.56*	0.52*	0.59*	0.72

Coefficient alphas are presented in boldface along the diagonal

PPAQ-E Primary Care Behavioral Health Provider Adherence Questionnaire Essential Items Subscale, *PPAQ-P* Primary Care Behavioral Health Provider Adherence Questionnaire Prohibited Items Subscale, *LIM* Level of Integration Measure Total Score, *SIS* Level of Integration Measure Systems Integration Subscale, *BCS* Level of Integration Measure Beliefs and Commitment Subscale, *IPS* Level of Integration Measure Integrated Practices Subscale, *RSS* Level of Integration Measure Relationship Subscale, *TSS* Level of Integration Measure Training Subscale, *LSS* Level of Integration Measure Leadership Subscale

* $p < .001$

overlooked aspect of CCC implementation. Within a field that has struggled to define and operationalize its procedures, this scale is the first to allow for quantification of CCC fidelity at the BHP level.

The approach to reliability assessment of the PPAQ followed conventional procedures of evaluating individual item statistics in an effort to improve the internal consistency of the subscales. Overall, few items were removed from the PPAQ during item analysis procedures, pointing to the value of the Delphi process for determining which items to include in the PPAQ. The exclusion of the six PPAQ-C items eliminated sources of low reliability and reduced the overall length of the PPAQ. The poor reliability of the PPAQ-C might be explained by the heterogeneous nature of the items. These six items were classified as a set of behaviors that are neither essential to nor prohibited by the CCC model. Individually, the items may be relatively common among BHPs, but they do not appear to represent a unitary dimension of “compatible” behaviors. In electing to eliminate these items based on their poor reliability, some degree of content validity is likely lost. However, we felt that removing these items was balanced by improved interpretability of a measure consisting of only two discrete subscales.

In contrast to the PPAQ-C, the PPAQ-E and PPAQ-P were both found to be reliable. The alpha estimate for the PPAQ-P was considerably lower than the estimate for the PPAQ-E which also may be due to the diverse set of items it contained. CCC is frequently described as a patient-centered, highly flexible model that addresses a wide range of patient concerns [28–30]. Because of this patient-centered emphasis, identifying behaviors that should be prohibited in CCC presented a challenge because they are seldom discussed in the literature [31]. Nonetheless, some very time-consuming practices that might occur routinely in specialty mental health settings should be discouraged in CCC to maintain a population-based, brief treatment focus. Accordingly, the ten PPAQ-P items appear to focus on behaviors associated with providing (a) specialty assessments and

interventions, (b) sessions of longer than 30 min, or (c) frequent or open-ended sessions. It should also be noted that although the term prohibited may carry an unrealistically rigid connotation, we defined prohibited simply as behaviors that should be avoided when working in CCC [12, 13]. Ultimately, the frequency with which providers engage in behaviors that are classified as prohibited by the PPAQ should be balanced with the need to remain patient-centered when providing clinical care.

In terms of validity assessment, the pattern of highly significant low to moderate correlations of the PPAQ-E and LIM subscales suggested very good convergent validity. The nonsignificant correlations between the PPAQ-P and the LIM subscales similarly suggest good discriminant validity. Combined with the lack of correlation between the PPAQ-E and PPAQ-P, these results reinforced the conceptual utility of identifying both essential and prohibited items as indicators of protocol adherence. The results of the known-groups comparisons suggested partial support for the validity of the PPAQ. We were unable to adequately assess our hypothesis regarding differences between BHPs in CCC and care management roles because the vast majority of care managers we surveyed were also engaging in CCC roles. Although this finding may reflect limitations of our sampling strategy, alternatively it may reflect progress in the VA mandate to provide both CCC and care management in primary care as cohesive or unified services. In retrospect, it is not completely surprising that we did not find differences between BHPs in different PC-MHI roles. Despite our emphasis on identifying CCC-specific behaviors, the PPAQ is likely tapping into behaviors that generalize to the functions of PC-MHI providers more broadly, including those elements of care management that overlap with CCC, such as interaction with primary care teams, engaging in patient education, and delivering brief interventions. The care management-specific functions not assessed in the PPAQ focus on using telephone-

based care, an informatics platform, routine follow-up at highly structured intervals, and heavy emphasis on patient assessment. Whereas this finding does not diminish the value of PPAQ, it suggests that comparison between BHPs in CCC-only and those in both CCC and care management roles may have limited utility for specifying the domain of CCC.

Our remaining hypotheses were largely supported among the other group comparisons with the exception of time in role. Greater adherence to CCC essential behaviors was associated with a CBT orientation and greater perceived resources to enact CCC. In contrast, working at a CBOC or having a social work educational background was associated with poorer fidelity due to more frequent engagement in prohibited behaviors. The reason why these variables impact fidelity differently across essential and prohibited behaviors is not clear, but further supports the value of identifying two dimensions of BHP behavior. Future research in this area is warranted, but additional efforts should be taken to understand the specific implementation challenges within CBOCs, including staffing these clinics with BHPs who are experienced in CBT, and are receptive to additional training to adapt CBT for CCC.

There are several limitations to our study that warrant mention. Our sampling frame included BHPs with a high probability of being in PC-MHI, but we ultimately relied on individuals to self-report the nature of their role. We were therefore unable to assess if some participants erroneously classified themselves as a CCC provider, care manager, or both. Furthermore, our sampling frame was based on lists of BHPs who either participated in previous CCC-related research or trainings rather than from the full population of BHPs within the VA system. Thus, those BHPs who participated in CCC-related trainings may differ in their responses to this instrument due to increased knowledge or familiarity with CCC practice than those BHPs who did not have such trainings. Additionally, although the PPAQ was designed for use in any CCC-type setting, our reliance on VA BHPs alone may further limit the generalizability of our findings to non-VA providers. Validation of the PPAQ in non-VA samples is warranted.

Another issue related to use of self-report relates to the issue of social desirability bias. For example, participants may have been reluctant to endorse low levels of knowledge and skills related to CCC if they felt it cast them in an unfavorable light. Therefore, the limited range of responses to this item made it impossible to assess its association with PPAQ scores. Similarly, participants may have been inclined to overreport their level of fidelity on PPAQ items if they believed that their self-report was somehow linked to individual performance appraisal. Although the ultimate impact of social desirability is difficult to determine, the web-based survey format and multiple assurances of anonymity may have reduced the influence of this type of bias [32].

Further limitations included that our methodology did not allow us to investigate multiple forms of

construct validity. We considered a range of options by which to assess validity of the PPAQ, including confirmatory factor analysis to test if our assertions about essential, compatible, and prohibited items could be upheld. However, with the large number of items in the PPAQ, the ratio of observations per parameter required for factor analysis demanded an unfeasibly large sample size. We further excluded attempts at assessing predictive validity through longitudinal data because of our cross-sectional survey design and use of anonymous data collection. Additionally, there are no well-accepted “gold standard” measures of CCC adherence that could be measured alongside the PPAQ that could support assessment of concurrent validity. Future validation of the PPAQ through the above approaches will no doubt be valuable, yet our assessment using known-groups and correlational analyses was nonetheless very encouraging. Additional research will also need to assess for test-retest reliability and sensitivity to change over time. The acceptability of this instrument on behalf of BHPs as a fundamental tool for assessing CCC implementation should also be investigated to identify any barriers to utilization.

STUDY IMPLICATIONS

Applying the PPAQ to future effectiveness research can begin to address the lack of literature regarding if and how provider protocol adherence is associated with clinical outcomes. It has been our experience that site-specific modification of CCC to meet the needs of local system requirements is common, even in the relatively well-resourced VA system. While these modifications meet the more immediate goal of rapid implementation of CCC, it has been argued previously that evidence-based models should not be substantially altered without sufficient additional research to guide selection of elements to retain, discard, or amend [15]. In the worst case scenario, rather than so-called model “evolution,” spontaneous modifications could lead to *devolution* (i.e., degeneration) away from critical components responsible for clinical effectiveness. Therefore, research that employs the PPAQ can provide the basis for measuring and assessing decomposition of effect that can better guide a more systematic and empirically based approach to local adaption of CCC [12]. For example, it may be determined that essential behaviors are more strongly associated with beneficial clinical outcomes than prohibited behaviors. If this were the case, then training and facilitation efforts could focus on supporting BHPs to engage more frequently in behaviors consistent with CCC with less focus on avoiding certain prohibited behaviors.

In support of the implementation of CCC in diverse primary care settings, the PPAQ may additionally be useful for training BHPs. Those who are new to the CCC role often consider the transition to working in primary care challenging.

Novice BHPs may find self-assessment with the PPAQ to be particularly useful during orientation or for monitoring one's acclimation to primary care due to the PPAQ's emphasis on typical clinical behaviors. Administrators may also find it useful during initial provider selection. This type of clear-cut, plainly stated guidance may be especially valuable at a time when learning "the ground rules" of CCC is vital. Although the PPAQ alone does not provide the type of nuanced instruction necessary for comprehensive training, it has potential to prove its value by offering an easily employed, fine-grained assessment of BHP behavior in systems with limited resources. For those systems with greater availability of resources, the PPAQ may be used during pre- and post-implementation or program development strategies (e.g., facilitation) to assess the impact of comprehensive mentoring and training strategies. The PPAQ may also be useful for program evaluation when used in systems with multiple BHPs, particularly when assessing variability in implementation status across clinics or geographic sites. Although its sensitivity to change will need to be assessed in future work, the PPAQ may also be useful to evaluate intended or unintended program modifications over time by identifying aspects of BHP behavior that contribute to program drift. Little has been written about the sustainability of CCC, but like many complex interventions, provider behavior plays a key role in program maintenance. The PPAQ may provide an opportunity to first identify behaviors contributing to program or provider-level maintenance so that targeted quality improvement efforts can be initiated. For example, an examination of changes in the PPAQ-E total score and individual items would indicate if a generalized decline in fidelity is present or if particular behaviors, such as utilization of brief screening measures for symptom assessment, are not being maintained. As demonstrated in the current study, the PPAQ is easily administered electronically and may also prove useful when incorporated into online training modules.

In summary, CCC aims to improve the identification, assessment, and treatment of mental and behavioral health concerns by co-locating and integrating BHPs in primary care. The evidence base for CCC is growing, but research in this area has rarely considered how provider behaviors impact clinical outcomes. To that end, the PPAQ has emerged as the first reliable self-report instrument of BHP fidelity that has demonstrated strong evidence of criterion validity. Quantifying fidelity as provider adherence to essential and prohibited behaviors helps to disentangle a complex approach to collaborative care that has beneficial implications for provider training, program monitoring, and clinical research. While the PPAQ has the potential to identify service delivery gaps, it will be important to maintain a stance in which fidelity assessment is used to support BHPs' efforts to improve their capacity to deliver evidence-based, patient-centered care.

Acknowledgments: This study was supported with resources and the use of facilities at the VA Center for Integrated Healthcare and the VA Western New York Healthcare System. The information provided in this study does not represent the views of the Department of Veterans Affairs or the US Government.

Conflict of interest: There are no known conflicts of interest for reasons financial or otherwise, no known competing interests, and no companies or products are being featured in this research.

1. Brown Levey SM, Miller BF, deGruy FV III. Behavioral health integration: an essential element of population-based healthcare redesign. *TBM*. 2012;2:364-371.
2. Miller B, Mendenhall T, Malik A. Integrated primary care: an inclusive three-world view through process metrics and empirical discrimination. *J Clin Psychol Med Settings*. 2009;16:21-30.
3. VHA: Uniform mental health services in the VA medical centers and clinics. Availability at: http://www1.va.gov/vhapublications/ViewPublication.asp?pub_ID=1762. Accessed 3 Dec 2012.
4. Zeiss AM, Karlin BE. Integrating mental health and primary care services in the Department of Veterans Affairs health care system. *J Clin Psychol Med Settings*. 2008;15:73-78.
5. Post EP, Metzger M, Dumas P, Lehmen L. Integrating mental health into primary care within the Veterans Health Administration. *Fam Syst Health*. 2010;28:83-90.
6. Kearney LK, Post EP, Zeiss A, Goldstein MG, Dundon M. The role of mental and behavioral health in the application of the patient-centered medical home in the Department of Veterans Affairs. *Transl Behav Med*. 2011;1:624-628.
7. United States Air Force Medical Operations Agency. Primary behavioral health care services: Practice manual. Available at: <http://www.integratedprimarycare.com/files/primary%20care%20practice%20manual.pdf>. Accessibility verified November 1, 2012.
8. Robinson PJ, Strosahl KD. Behavioral health consultation and primary care: lessons learned. *J Clin Psychol Med Settings*. 2009;16(1):58-71.
9. Cigrang JA, Dobmeyer AC, Becknell ME, Roa-Navarrete RA, Yerian SR. Evaluation of a collaborative mental health program in primary care: effects on patient distress and health care utilization. *Prim Care Commun Psychiatry*. 2009;11(3):121-127.
10. Bryan CJ, Morrow C, Appolonio KK. Impact of behavioral health consultant interventions on patient symptom and functioning in an integrated family medicine clinic. *J Clin Psychol*. 2009;65(3):281-293.
11. Ray-Sannerud BN, Morrow CE, Kanzler KE, et al. Longitudinal outcomes after brief behavioral health intervention in an integrated primary care clinic. *Fam Syst Health*. 2012;30(1):60-71.
12. Mowbray CT, Holter MC, Teague GB, Bybee D. Fidelity criteria: development, measurement, and validation. *Am J Eval*. 2003;24:315-340.
13. Perepletchikova F, Kazdin AE. Treatment integrity and therapeutic change: issues and research recommendations. *Clin Psychol Sci Pract*. 2005;12:365-383.
14. Beehler GP, Funderburk JS, Possemato K, Vair C. Developing a measure of provider adherence to improve the implementation of behavioral health services in primary care: a Delphi study. *Implement Sci*. 2013;8:19.
15. Carroll C, Patterson M, Wood S, Booth B, Rick J, Balain S. A conceptual framework for implementation fidelity. *Implement Sci*. 2007;2:40.
16. Sireci SG. Gathering and analyzing content validity data. *Educ Assess*. 1998;5:299-321.
17. Netemeyer RG, Bearden WO, Sharma S. *Scaling procedures: issues and applications*. London: Sage; 2003.
18. Hedrick SC, Chaney EF, Felker B, et al. Effectiveness of collaborative care depression treatment in Veterans' Affairs primary care. *J Gen Intern Med*. 2003;18:9-16.
19. Oslin DW, Ross J, Sayers S, Murphy J, Kane V, Katz IR. Screening, assessment, and management of depression in VA primary care clinics. The Behavioral Health Laboratory. *J Gen Intern Med*. 2006;21:46-50.8.
20. Kirchner JE, Owen RR, Dockter N, et al. Equity in Veterans' mental health care: Veterans Affairs medical center clinics versus community-based outpatient clinics. *Am J Med Qual*. 2008;23(2):128-135.
21. Beehler GP, Wray LO. Behavioral Health Providers' perspectives of delivering behavioral health services in primary care: a qualitative analysis. *BMC Heal Serv Res*. 2012;12:337.
22. Hunter CL, Goodie JL, Oordt MS, Dobmeyer AC. *Integrated Behavioral Health in Primary Care: Step-by-Step Guidance for*

- Assessment and Intervention*. Washington DC: American Psychological Association; 2009.
23. Fauth J, Tremblay G. *The Integrated Care Evaluation Project*. Antioch: Endowment for Health; 2011.
 24. Shih T, Fan X. Comparing response rates in e-mail and paper surveys: a meta-analysis. *Educ Res Rev*. 2009;4(1):26-40.
 25. Spector PE. *Summated rating scale construction: an introduction*. Newbury Park: Sage; 2009.
 26. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951;16:197-334.
 27. Kline P. *Handbook of psychological testing*. 2nd ed. London: Routledge; 1999.
 28. Nash JM, McKay KM, Vogel ME, Masters KS. Functional roles and foundational characteristics of psychologists in Integrated Primary Care. *J Clin Psychol Med Settings*. 2012;19:93-104.
 29. Pomerantz A, Shiner B, Watts BV, et al. The White River model of co-located, collaborative care: a platform for mental health and behavioral health care in the medical home. *Fam Syst Health*. 2010;28(2):114-129.
 30. Haley WE, McDaniel SH, Bray JH, et al. Psychological practice in primary care settings: practical tips for clinicians. *Prof Psychol Res Pract*. 1998;29(3):237-244.
 31. Pomerantz AS, Corson JA, Detzer MJ. The challenge of integrated care for mental health: leaving the 50 minute hour and other sacred things. *J Clin Psychol Med Settings*. 2009;16:40-46.
 32. Nederhof AJ. Methods of coping with social desirability bias: a review. *Eur J Soc Psychol*. 1985;15:263-280.