

Monitoring depression with the PHQ-9 in primary care: a qualitative study

Ching, Brian CF; Dewar-Haggart, Rachel; May, Carl R; Leydon, Geraldine; Kendrick, Tony; Trapasso, Emilia; Patel, Tasneem; Bird, Molly; Bridewell, Lauren; Bui, Lien; Corcoran, Emma; Hahn, Jane S; Tiwari, Riya; Logan, Mekeda X; Dowrick, Christopher; Lewis, Glyn; Gabbay, Mark; Geraghty, Adam WA

DOI: <https://doi.org/10.3399/BJGPO.2025.0159>

To access the most recent version of this article, please click the DOI URL in the line above.

Received 07 August 2025

Accepted 04 September 2025

© 2025 The Author(s). This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 License (<http://creativecommons.org/licenses/by/4.0/>). Published by BJGP Open. For editorial process and policies, see: <https://bjgpopen.org/authors/bjgp-open-editorial-process-and-policies>

When citing this article please include the DOI provided above.

Author Accepted Manuscript

This is an 'author accepted manuscript': a manuscript that has been accepted for publication in BJGP Open, but which has not yet undergone subediting, typesetting, or correction. Errors discovered and corrected during this process may materially alter the content of this manuscript, and the latest published version (the Version of Record) should be used in preference to any preceding versions

Monitoring depression with the PHQ-9 in primary care: a qualitative study

Brian CF Ching¹, Rachel Dewar-Haggart², Carl R May³, Geraldine Leydon², Tony Kendrick², Emilia Trapasso⁴, Tasneem Patel⁴, Molly Bird¹, Lauren Bridewell², Lien Bui², Emma Corcoran¹, Jane S Hahn¹, Riya Tiwari², Mekeda X Logan¹, Christopher Dowrick⁴, Glyn Lewis¹, Mark Gabbay⁴, & Adam WA Geraghty^{2*}

¹Division of Psychiatry, University College London, Maple House, 149 Tottenham Court Road, London W1T 7NF, UK.

²School of Primary Care, Population Sciences & Medication Education, University of Southampton, Aldermoor Health Centre, Aldermoor Close, Southampton, SO16 5ST, UK.

³Faculty of Public Health and Policy, London School of Hygiene and Tropical Medicine, 15-17 Tavistock Place, London WC1H 9SH, UK.

⁴Department of Primary Care and Mental Health, University of Liverpool, Waterhouse Block B 1-5 Brownlow St Liverpool L69 3GL, UK.

*Corresponding author: Adam WA Geraghty; a.w.geraghty@soton.ac.uk.

Brian CF Ching, BSc MSc, Research Assistant, <https://orcid.org/0000-0002-2179-9793>. Rachel Dewar-Haggart, BSc MSc PhD CPsychol, Research Fellow, <https://orcid.org/0000-0002-3757-1152>. Carl R May, PhD FAcSS, Professor of Health Systems Implementation, <https://orcid.org/0000-0002-0451-2690>. Geraldine Leydon, BA MA PhD, Professor of Medical Sociology and Communication, <https://orcid.org/0000-0001-5986-3300>. Tony Kendrick, BSc MD FRCGP FRCPsych FHEA, Professor of Primary Care, <https://orcid.org/0000-0003-1618-9381>. Emilia Trapasso, BSc MSc PhD, Postdoctoral Research Fellow in Public Engagement, <https://orcid.org/0000-0002-7539-7204>. Tasneem Patel, PhD, Research Assistant, <https://orcid.org/0000-0003-4496-7603>. Molly Bird, MSc, Research Assistant, <https://orcid.org/0000-0003-1676-1907>. Lauren Bridewell, BMedSci, Medical Student, <https://orcid.org/0000-0001-9569-7813>. Lien Bui, BA MSc PhD, Research Fellow, <https://orcid.org/0000-0003-3434-4066>. Emma Corcoran, MSci DClinPsy, Clinical Psychologist, <https://orcid.org/0000-0001-5811-4615>. Jane S. Hahn, BA MSc, Research Assistant, <https://orcid.org/0000-0002-4584-9441>. Riya Tiwari, BSc MSc, Research Assistant, <https://orcid.org/0000-0002-3002-3276>. Mekeda X Logan, BSc MSc, Research Assistant, <https://orcid.org/0000-0001-7899-7531>. Christopher Dowrick, BA MSc MD FRCGP, Emeritus Professor of Primary Medical Care, <https://orcid.org/0000-0002-4245-2203>. Glyn Lewis, PhD FRCPsych, Professor of Epidemiological Psychiatry, <https://orcid.org/0000-0001-5205-8245>. Mark Gabbay, MD FRCGP DFSRH Dip Psychotherapy, Professor of General Practice, <https://orcid.org/0000-0002-0126-8485>. Adam WA Geraghty, PhD CPsychol FHEA, Associate Professor of Psychology and Behavioural Medicine, <https://orcid.org/0000-0001-7984-8351>.

Abstract (250 words)

Background: Primary care guidelines recommend general practitioners consider using depression symptom questionnaires as patient-reported outcome measures (PROMs) to monitor depression in adults to inform treatment and evaluate management strategies. The PROMDEP randomised controlled trial assessed the use of the Patient Health Questionnaire's (PHQ-9) effectiveness and cost-effectiveness for monitoring depression. We gathered qualitative evidence on the views and experiences of participating patients and practitioners to inform interpretation of the findings.

Aim: To explore the views and experiences of patients and practitioners of using the PHQ-9 in the PROMDEP trial of monitoring depression in primary care.

Design and Setting: Nested qualitative study and process evaluation of the trial in primary care in England and Wales.

Method: Twenty-nine patients and 15 practitioners took part in semi-structured telephone or video interviews. Interview data were analysed using thematic analysis.

Results: Patients and practitioners both valued the PHQ-9 and reported limitations in its use for monitoring depression. This included its role in improving understanding of depression, impact on consultation and care, and integration within current primary care processes. In the context of the PROMDEP trial, our findings highlight potential reasons for the mixed trial findings, including how resistance in its use in practice may be due to barriers that make it hard for practitioners to integrate the PHQ-9 in consultations.

Conclusion: Monitoring of depression using PROMs need to be considered within the context of current primary care processes and resources. Further research is warranted to understand how the PHQ-9 can be successfully integrated into consultations.

Keywords (up to six, MeSH headings): primary care, depression, PHQ-9, patient-reported outcome measure, qualitative research

How this fits in (no more than four short sentences what was previously known or believed on the topic and what your research adds, particularly focusing on the relevance to clinicians):

- Qualitative studies have shown mixed patient and practitioner views of using PROMs for depression.
- The PROMDEP trial was a cluster randomised controlled trial that found the PHQ-9 for the monitoring of depression did not incur significant benefit for depression scores in the short term.
- Our findings suggest potential benefits of using the PHQ-9 as a PROM for improving understanding of depression and impact on consultations and care, but also limitations in usability and implementation.
- The mixed trial findings of the PROMDEP trial may be explained by complex mechanisms underlying the use of the PHQ-9 in consultation.

Introduction

Depression symptom questionnaires, such as the Patient Health Questionnaire (PHQ-9), have been recommended for use as patient reported outcome measures (PROMs) in primary care depression guidelines internationally.¹⁻³ However, despite its use across healthcare settings, there is mixed evidence about the effectiveness for the use of PROMs in monitoring depression in primary care.⁴⁻⁷ Following The National Institute for Health and Care Excellence (NICE) guidance in the UK between 2006-2013, the Quality and Outcomes Framework (QOF) financially incentivised GPs to use symptom questionnaires to assess and follow-up depression. Qualitative research found that patients and GPs found PROMs acceptable to use and valued their utility in confirming diagnosis, facilitating structured conversations, and supporting decision-making around appropriate treatments based on symptom severity.^{5,8} Conversely, some practitioners and patients felt PROMs intruded in consultations, hampered practitioner-patient dialogue, and were not motivated to implement their use in practice, and did not capture all depressive symptoms.⁹⁻¹¹

Relatively few studies have investigated specifically the use of PHQ-9 for the monitoring of depression in primary care. The PROMDEP trial aimed to investigate the effectiveness of using PHQ-9 to assess and monitor depression in patients in UK practices.¹²⁻¹³ We found no significant benefit in depression scores at 12-weeks, suggesting that using PHQ-9 to monitor depression in practices were not significantly beneficial for depression in primary care patients. However, quality of life was better in the intervention arm at 26-weeks, and a clinically significant difference in depression scores at 26-weeks could not be ruled out. Importantly, only 41% of patients in the intervention arm had a GP follow-up PHQ-9 recorded. There may be unmeasured complex mechanisms not captured in the quantitative findings that may explain the mixed trial findings. To understand this, we also conducted a nested qualitative study and process evaluation within the PROMDEP trial to contextualise the quantitative trial findings.

In this paper, we report the nested qualitative study and process evaluation that aimed to explore the views and experiences of patients and practitioners of using the PHQ-9 to monitor depression in primary care, to inform greater understanding of the PROMDEP trial's quantitative findings.

Method

The PROMDEP trial

The PROMDEP trial was a cluster randomised controlled trial across 141 practices in England and Wales between November 2018 and December 2021 (Trial registration: ISRCTN 17299295). Ethical approval was obtained from the West of Scotland NHS Research Ethics Service committee (ref. 18/WS/0144). Practices were randomly allocated to the intervention or control group. Inclusion criteria were patients aged 18 years or older, with a new episode of depressive disorder or symptoms. Patients recruited into the study in the practices allocated to the intervention group had an initial PHQ-9 administered by a researcher within 2-weeks of recruitment. Practitioners, which included GPs or nurse practitioners, were then asked to repeat the PHQ-9 in a follow-up consultation in clinic. Patients were provided with an infographic with written feedback on their PHQ-9 scores and potential treatment options to discuss with their GPs after completing the PHQ-9 (see Supplementary Figure 1). Patients recruited into the control group received treatment as usual as provided by primary care practitioners. Further details of the trial and the PHQ-9 intervention are published elsewhere.¹²⁻¹³

The recruitment process and delivery of the PHQ-9 was impacted by the COVID-19 pandemic and changes were implemented during the trial as a result. Recruitment was significantly slowed down

as practices were inundated and involvement in research activities were paused or not a priority for practitioners. As a result, recruitment took much longer than planned. Additionally, due to national lockdown and social restrictions, the initial PHQ-9 was not administered in-person on paper by researchers and completed digitally either over the telephone or via video conferencing. Primary care consultations also changed, with most practices offering remote consultations.

Data collection

For the nested qualitative study and process evaluation embedded within the PROMDEP trial, we conducted interviews with participants. We recruited practitioners and patients in both trial arms using maximum variation sampling to capture perspectives from participants with a diverse range of characteristics, including gender, ethnicity, education level, and location of practice. Patients were interviewed between completing their 12-week and 26-week follow-up research assessments, and practitioners were invited during their participation in the trial. Patients gave consent to be contacted to take part in an interview as part of the consent process for the main part of the trial. Researchers approached potential participants by telephone and emailed the participant information sheet. Written, verbal or electronic informed consent was obtained at the start of the interviews.

Interviews were completed by researchers (BCFC, RDH, ET, TP, MB, LBr, LBu, EC, JSH, RT, and MXL) and conducted face-to-face, by telephone, or online video call using Microsoft Teams. A semi-structured interview guide (Supplementary Table 1) informed by normalisation process theory¹⁶ was developed by the research team and patient and public involvement (PPI) contributors was used. Interviews took place in August 2019-July 2022. Interviews lasted 13-117 minutes. Interviews were audio recorded and transcribed verbatim with potentially identifying information removed from transcripts to ensure participant confidentiality and anonymity.

Analysis

Analysis was conducted in two stages. First, we conducted thematic analysis.¹⁵ We sought immersion in the data by reading and re-reading all transcripts and reflecting on interviews and discussing potential themes in data sessions. BCFC and RDH independently coded a set of transcripts and collaboratively developed an initial coding frame. This framework was then used to code subsequent transcripts and iteratively extended and revised as new codes were developed. Coding was inductive and derived from the data. BCFC developed the final set of themes from the coded data, refined with feedback from the research team and PPI representatives. Second, we undertook an attribution analysis¹⁶ of the results of the thematic analysis. CRM identified key attributions about the workability and integration of the intervention within the trial. These were then mapped onto the action constructs of normalisation process theory.¹⁷

We adhered to frameworks for conducting and writing up high quality qualitative research.¹⁸ To maximise our result's validity, multiple researchers were involved in the data collection, coding, and analysis. At different stages of the analysis, we presented preliminary findings to the research team and discussed the face validity of themes.

Our research team for this study is made up of a diverse group of academic and clinical researchers at different career stages, including research assistants, associates, and professors. Study team members have expertise across multiple clinical specialties, including general practice and psychiatry, and have contributed to the development of the NICE guidelines for the treatment and management of depression in adults. We bring a range of research interests and experiences into the study, including qualitative methods, development and evaluation of complex interventions

(including PROMs), implementation science, and primary care service development. We represent a range of gender, culture, and racial identities (including Black, Asian, and White British). This range of identities and perspectives enhanced our knowledge and insight on experiences of depression and its monitoring in primary care, but we still maintained curious and reflected on our preconceived notions, allowing diverse ideas to be developed inductively from the data. These important reflections informed how we interpreted and analysed the qualitative data.

Results

Twenty-nine patients and 15 practitioners agreed to be interviewed for the study. Tables 1 and 2 show participant characteristics for patients and practitioners. Five inductive themes were developed after analysis of all interview data which are discussed below (see Table 3). We included quotes to illustrate and evidence our analysis. Quotes have been allocated participant identification codes to maintain confidentiality and anonymity (GP for practitioners and PT for patients).

Table 1. Participant characteristics: patient interviews

Characteristics N=29		Intervention N=18	Control N=11
Gender			
Female	10	8	
Male	8	3	
Age Mean years (SD)	36.7 (12.7)	46.8 (19.7)	
Ethnicity			
White	15	10	
Chinese	1	0	
Indian	1	0	
African and Irish	1	0	
Persian	0	1	
Marital Status			
Married/Cohabiting	7	7	
Single	11	3	
Divorced	0	1	
Education Level			
None	2	0	
CSE/NVQ Level 1	1	0	
GCSE/O Level/NVQ Level 2	3	1	
HNC/HND/City & Guilds/Teaching qualification/NVQ Level 4	2	1	
Degree/higher degree/NVQ Level 5	10	7	
Vocational qualification	0	2	
Employment Status			
Full time work/Self-employed	11	5	
Part time work	2	5	
Homemaker	1	0	
Retired	1	1	
Student	3	0	
IMD* (Median)	6.5	9	

*Index of Multiple Deprivation score for practice area

Table 2. Participant characteristics: practitioner interviews

Characteristics N=15	Intervention N=11	Control N=4
Gender		
Female	4	3
Male	7	1
Location		
Urban	8	3
Rural	3	1
Practice size		
Large	7	3
Small	4	1
IMD* (Median)	7	7.5

*Index of Multiple Deprivation score for practice area

Table 3. Themes and sub-themes

1. Improved understanding of depression	
	1.1 Recognising symptoms
	1.2 Monitoring over time
	1.3 Motivation and hope
2. Usability	
	2.1 Being pigeonholed
	2.2 Accessibility
3. Impact on the consultation	
	3.1 Driver of discussion
	3.2 Patient-GP relationship
	3.3 Person-centred care
4. Impact on care	
	4.1 Evidence to inform treatment and management
	4.2 Objectivity versus subjectivity
5. Organisational barriers and facilitators to implementation	
	5.1 Time restraints
	5.2 Technological integration
	5.3 Frameworks and guidelines

Thematic analysis

Improved understanding of depression

Recognising symptoms

Practitioners and patients believed the PHQ-9 was a helpful tool in providing information on the range of symptoms and severity categories of depression and how particular symptoms affected them.

'Having your depressive symptoms split up into nine sections I think is helpful. It's not a big accumulation of just feeling horrible, it's splitting it up and making sense of it. I think it works for lots of people who want to understand their symptoms better.' – GPO1002-01

'As I was reading [the items] I was thinking to myself – yes, that is me, that is me and I kind of understood [depression] bit more.' - PT02024-03

Some patients felt that seeing each item and the infographic was validating, highlighted that they needed support, and provided solace that they were not alone.

'It reassured me that I wasn't going mad, that it is an illness, and it does affect a lot of other people as well.' - PT01023-02

Monitoring over time

The PHQ-9 was reported to be a useful tool for practitioners to monitor their patients' depressive symptoms temporally, identifying improvements and deteriorations. Regularly completing the PHQ-9 highlighted changes in depressive symptoms, which was reassuring and rewarding to patients, allowing patients to 'map out' areas for improvement.

'It's mainly mapping to see which bits needed to be improved and which bits could be not focused on for a little while, what to work on or how to make yourself feel better.' - PT03037-03

In the control group, patients described how they were monitored over time without questionnaires. They described how practitioners used the conversation in the consultation to check in, using both what the participants said, and how they appeared to look for worsening or improvement.

*'P: She said I looked like I was improving, like mood - wise.
Int: Yes, it sounds like she's kind of gauged that from the mood an in the consultation perhaps?
P: I think a combination of both, really, yes, I guess maybe how I was like saying these things on top of what I was saying.'* - PT 0305103

Motivation and hope

Using the PHQ-9 and realising the severity of their depression motivated some patients in actively seeking support. Over time, seeing improvements motivated patients to continue their efforts and emphasised that their perseverance was worthwhile, providing hope.

'When I see it go down, you're relieved. You're like, oh, it's working; it's worth getting help.' - PT01002-04

Usability

Being pigeonholed

Some practitioners and patients experienced the PHQ-9 as limited in capturing people's experiences and symptoms; suggesting response options on the PHQ-9 pigeonholed patients into arbitrary categories. This was perceived by patients as inaccurate and reductive of their complex experiences.

'The frequency of symptoms remained the same, but the intensity decreased. I felt like I couldn't reflect it on the questionnaire.' – PT01096-05

Accessibility

Some practitioners suggested, that beyond the trial, the waiting room as a prime location to complete the PHQ-9. Staff in reception could support those who need help filling them in and this would ensure patients complete the PHQ-9 as asked. Especially for patients who 'can't motivate themselves to get out of bed, they may not get the motivation to do the form' (GP01026-01).

Impact on the consultation

Driver of discussion

As each item highlighted a specific symptom, practitioners felt that the PHQ-9 was helpful in identifying specific symptoms patients were experiencing and its impact. The PHQ-9 reminded them to conduct necessary risk assessments. This enabled safety netting and necessary referral to mental health services and prompted practitioners to have sensitive conversations with patients.

'You are forced to ask the question about the self-harm... In a consultation where you have other things to discuss with them, it can be easy to forget.' – GP01002-01

The PHQ-9 was described helpful in guiding discussions around patients' difficulties, needs, and care. It facilitated productive conversations as all parties had information of the problem as assessed by the PHQ-9.

'It sometimes is really difficult to explain how you're feeling. Being able to put it down on paper like that both yourself and your GP can see how you have been feeling and what areas are a bit worse than others.' – PT01002-04

For some patients, the PHQ-9 facilitated conversations about depression which was difficult to speak about due to reasons such as stigma, feeling like they could not articulate their experiences, or the perception that the GP practice was not a place for mental health.

Patient-GP relationship

The PHQ-9 facilitated practitioners' engagement and support to patients through validation, praise, monitoring, and maintenance of progress.

'If the PHQ 9 score was decreasing, then I would congratulate the patient... I say, that's great progress.' – GP02013-01

The structure provided by the PHQ-9 ensured patients could develop a trusting and personal relationship with their GP and be reassured they were supported. Continuity of care was voiced by patients as paramount here, as they could build rapport with their doctor and not need to repeatedly retell their problems.

'It was good knowing that there's somebody there to monitor that progress. I think that's reassuring.' – PT03044-04

Person-centred care

Practitioners described how the PHQ-9 provided a richer understanding of how depression impacted their patients. The PHQ-9 also facilitated conversations between practitioner and patient about their care based on what patients want.

'I think it's a more holistic assessment than you would do if you're left to your own devices and you ask just three random questions, which might not actually show you the whole extent of how things affect you... I negotiate with the patients what they think might be acceptable to them.' - GP01002-01

Impact on care

Evidence to inform treatment and management

Most practitioners and patients found the PHQ-9 solidified decisions about potential treatments and informed the urgency of suggestions, where worse scores would indicate more intensive treatment and insistence.

'I perhaps might be more persuasive towards someone who's dark red saying I think we do need to start you on treatment... If someone's yellow, I might sit more on the fence and just say – this is an option, how do you feel? Would you prefer to have treatment or not have treatment?' - GP01056-01

The PHQ-9 score was also indicative of whether the treatment plan was working. If scores did not change or increased, this made practitioners and patients consider stepping up treatment.

Objectivity versus subjectivity

Many practitioners saw the PHQ-9 as an objective measure in providing clarifying quantitative information about patients' difficulties and needs where symptoms may not present the same or as clearly in all patients. The PHQ-9 was deemed useful for patients where symptoms may not be overtly visible during consultation.

'[The PHQ-9] is something we should think about using, especially with something as grey as depression... People could be tearful but not have really bad depression and there could be other people who look quite together with it but actually who were severely depressed. It's a helpful, objective tool.' - GP03044-01

Other practitioners, especially those with more experience, preferred using their clinical judgement in making decisions about patients' care. They did not like the rigidity of the categories and their suggested treatments as it felt like 'tick-box medicine' (GP01002-01). These practitioners expressed resistance to the continued use of the PHQ-9. However, they suggested it could be used as a guidance for younger practitioners with less experience.

'Doctors as a rule don't like tick-box medicine unless they are starting out as new doctors and haven't found their feet yet... As you're becoming more senior, doctors tend to do less of those things, because they feel they are being patronised and being told what to do... They'd rather do what they think is right rather than somebody else telling them to do so.' - GP01002-01

Organisational barriers and facilitators to implementation

Time restraints

Time limitations for each consultation was identified as a huge barrier to practitioners' use of the PHQ-9. Within a ten-minute consultation, GPs felt unable to spend time administering or discussing the PHQ-9 on top of general discussions around the patients' depression.

'They'll be given a ten-minute appointment which is the usual in GP land... Which is a bit of a shame when assessing depression because it clearly takes longer than that, let alone then having to do a PHQ-9.' - GP01026-01

Technological integration

Beyond the trial processes, for practitioners where the PHQ-9 was embedded within their practices' digital systems, benefits were perceived, including increased access to the PHQ-9, efficiency of appointment distributions by need, and maintained privacy for patients. This may have preserved resources and promoted better care.

'We can just text them the PHQ-9 link and they can fill it out. We can ask them to do the score once a month and they can text it or email it to the surgery. It gets automatically downloaded on the patient notes.' - GP02013-01

Frameworks and guidelines

Many practitioners felt that if the PHQ-9 was to be recommended in national guidelines they would feel more encouraged to implementing it in practice. To facilitate this, practitioners would need access to a clear evidence-base for its effectiveness for depression monitoring.

'If it was part of NICE guidelines, that would be different. We have to adhere to them... I think you find most people won't have a better system to assess patients, certainly not an evidence-based system. That probably would be the single best way to get it into the general GP population and use it in practice.' - GP01002-01

Attribution analysis

Thematic analysis provided a robust account of factors that promote or inhibit incorporation of the PHQ-9 in clinical practice. We then moved to an attribution analysis (see Table 4 for key attributions mapped onto action constructions of normalisation process theory). In the PROMDEP trial, the PHQ-9 conferred interactional advantages on patients because it provided them with a tool that enabled them to make sense of their depression, even though it captured only a limited range of their experienced symptoms. It also conferred an advantage on inexperienced practitioners because it provided an objective measure of symptom severity.

However, within the trial, there was evidence that the PHQ-9 was time-consuming and difficult to integrate into the consultation. Practitioners' accounts suggested that it could be pushed out of the in-person consultation and configured as an element of e-consult or could be administered by receptionists or nurses.

Table 4. Attribution analysis mapped onto action constructions of Normalisation Process Theory

NPT action constructs	Usability of the PHQ-9	Impact on the consultation	Improved understanding of depression	Impact on care	Other factors that promote or inhibit implementation	Conclusion
<i>Interactional workability – how does the PHQ-9 affect interactions between patients and care processes?</i>	(+) assists patient in (subjective) recognition of symptoms	(+) validates patient's need for care	(+) allows patients to understand severity of symptoms and actively seek support	(+) allows patients to 'map out' areas for improvement	(-) assumes that patients are motivated to effectively participate in their care	PHQ-9 has an interactional workability advantage for patients ++++Normalisation potential
<i>Relational integration – how does the PHQ-9 relate to existing knowledge and relationships?</i>	(+) supports inexperienced practitioners in (objective) diagnostic process	(+) assures patient of continuity of support from practitioners.	(-) results may be perceived by patients as inaccurate and reductionist	(-) captures limited range of patient experiences	(+) allows practitioners to effectively assess urgency of presentation	PHQ-9 has a relational integration advantage for less experienced practitioners, but not for patients or more experienced practitioners +++ Normalisation potential
<i>Skill-set workability – how is the character of patient and practitioners work affected by the PHQ-9?</i>	(+) provides a sense-making tool for patient and practitioners	(-) can shift patient work to the waiting room, e-consult, or online	(+) supports practitioners' clinical judgement	(+) permits monitoring of changing symptoms over time	(+) allows risk assessments around self-harm and permits objective referrals to mental health services	PHQ-9 has a skill-set workability advantage for practitioners ++++ Normalisation potential
<i>Contextual integration – how does the PHQ-9 relate to the setting in which it is operationalised?</i>	(+) changes in scores show whether practitioners' treatment plan is working	(-) adds time and complexity to the consultation	(+) score links to evidence-based treatment suggestions	(-) needs to be embedded in digital systems, adds to efficiency of access	(-) not included in national guidelines	PHQ-9, in the trial, has clinical value but is poorly contextually integrated in practice ++ Normalisation potential

Note: (+) and (-) indicates positive mechanisms/facilitators and barriers to implementation of the PHQ-9 in the monitoring of depression in primary care.

Discussion

Our nested qualitative study and process evaluation highlighted that patients and practitioners reported various benefits of using the PHQ-9 to monitor depression in primary care. Particularly, the PHQ-9 provided information on range and severity of depression symptoms, temporal changes in mood, informed treatment plans, facilitated dialogue, ensured risk assessments were carried out, and promoted a trusting practitioner-patient relationship and person-centred care.

However, the PHQ-9 was also perceived to have limitations. Some patients expressed the PHQ-9 as limited in capturing their nuanced experiences and meaningful changes. Practitioners also described the PHQ-9's severity categories and treatment suggestions as 'tick-box medicine'. There was also resistance in some practitioners to its use in practice.

In the context of the PROMDEP trial, these findings shed light on potential mechanisms that may explain the quantitative findings. The trial reported no significant effect on depression scores at 12-weeks. Use of the PHQ-9 in clinics may not have translated to changes that positively impacted depression outcomes due to factors related to usability and organisational factors. Practitioners highlighted issues with the integration of the PHQ-9 into current primary care systems and consultations. This is relevant considering a proportion of patients in the intervention arm did not have an initial or a follow-up PHQ-9 recorded, suggesting some patients in the intervention arm were not provided with the PHQ-9 itself. Inconsistent fidelity across GPs who were provided with compensation and training to administer the PHQ-9 suggests that the use of PHQ-9 as a depression monitoring tool involves a myriad of complex mechanisms that need to be addressed in primary care. This includes practitioner perceptions of the PHQ-9, guidelines and frameworks, and how it fits with current processes and technology in practices.

Strengths and limitations

Our nested qualitative study and process evaluation purposefully sought a diverse group of patients and practitioners. We explored a wide range of patient and practitioner experiences and views which maximise the potential transferability of our findings. We also harnessed the diversity of our study team in the data collection and analysis, which provided unique perspectives and input from different backgrounds (e.g., gender and racial identities) and expertise (e.g., clinical and academic roles). Rigorous qualitative methodology was followed to ensure validity and trustworthiness of our findings. Our study provided in-depth detail of potential mechanisms that may underly the findings of the main PROMDEP trial. We demonstrate that the use of PHQ-9 in monitoring depression in primary care is a complex intervention that requires integration into current processes for successful uptake by practitioners.

Potential biases may have been present in our recruitment. Patients and practitioners with more positive opinions about the PHQ-9 may have volunteered to participate as only interested participants were recruited into this study. Alternative experiences and views about the PHQ-9 may have been left out. Additionally, despite sampling efforts, certain characteristics dominated our sample (i.e., White and employed). Experiences of the PHQ-9 and monitoring of depression in primary care may vary across minoritised identities,²⁰ which refers to groups that have intersecting identities that may be socially minoritised and disadvantaged, such as women, people of colour, and people in deprived communities.²¹ Other characteristics that may have been important in the experiences of using the PHQ-9, such as practitioner job role, were not collected. This needs to be considered in interpretation of our findings.

Comparison with existing literature

Our findings complement previous studies that suggest potential utility of using PROMs in monitoring depression in primary care. Previous research found that PROMs helped patients identify and understand their depressive symptoms, provided structure to the consultation, and supported practitioners in developing and confirming treatment decisions and monitoring changes in depressive symptoms.^{5,8} Our study extends this by evidencing its utility in also instilling hope of recovery in patients, and supporting patients and practitioners in monitoring changes in depressive symptoms, developing a trusting and communicative relationship, and fostering person-centred care. However, this needs to be interpreted within the context of low fidelity within the PROMDEP trial, suggesting that this benefit may only be translated in practices when practitioners feel able to integrate this within their consultations. When implementing new methods or components of primary care consultations, adding to clinical workloads of practitioners is a barrier in its use.²² Malpass et al. found that some patients experienced the PHQ-9 as limited in accurately capturing their depressive symptoms due to the limited scope of the items.¹⁰ Additionally, research suggest that practitioners prefer using their clinical judgement ('phronesis') over PROMs in consultations as they value the human element of the practitioner-patient dialogue.⁷ Our study additionally highlighted how those views may vary depending on years of experience, where older practitioners may perceive the need to use the PHQ-9 as patronising and a burden. Our findings suggest that patients want to be able to identify the impact or magnitude of symptoms which the PHQ-9 may be limited in monitoring, and practitioners need to be inherently motivated to use a tool for its regular use.

Implications for research and practice

The quantitative findings of the RCT were that the intervention arm practitioners recorded follow-up PHQ-9 results in the medical records of only 41% of patients, suggesting that questionnaires were not used routinely to inform treatment decisions by all practitioners, even in practices which had volunteered to use them in the trial.¹²⁻¹³ It is possible the increased demands on practices during the COVID-19 pandemic made the completion of the study a lower priority during 2020-21.

Using the questionnaire could be made easier than it was in the trial. The extra time taken in consultations could be significantly reduced by automatic remote administration of the questionnaire beforehand, through sending it by text message or email, or including its administration in online consultations (e-consults). Some practice computer systems already incorporate a PHQ-9 automatically in e-consult questions for patients who indicate they may have mental health problems. Alternatively, telephone administration of the PHQ-9 has been shown to be reliable²⁰ and could be done by other practitioners or even administrative staff whose time is less expensive than that of GPs.

Future research is warranted to better understand how uptake of PHQ-9 and related PROMs can be improved in primary care. Rigorous and thoughtful implementation research needs to be conducted and consistently incorporate patient and practitioner perspectives, as well as other important stakeholders, such as service planners and policymakers. Considerations need to be made within current primary care systems and infrastructure.

Monitoring of depression using PROMs need to be considered further within the context of the current drive towards digitalising primary care processes and increasingly stretched resources. Successful integration of the PHQ-9 needs to be facilitated in ways that increase practitioners' decision-making capacity and promote positive perceptions about PROMs including enablement of the seamless use of PROMs in time-limited consultations.

Crucially, our quantitative and qualitative evidence shows questionnaire monitoring is useful for increasing patients' understanding of the wide range of depressive symptoms, and that patients like to track their own symptom changes over time. Practitioners should recognise those advantages, even if the scores do not directly inform their treatment decisions.

Funding: This project was funded by the National Institute for Health and Social Care Research (NIHR) Health Technology Assessment programme (ref: HTA 17/42/02). The views expressed in this publication are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

Ethical approval: This study was approved by the NHS Research Ethics Committee West of Scotland REC 5, on 21st September 2018 (reference: 18/WS/0144).

Competing interests: Tony Kendrick, Christopher Dowrick, Glyn Lewis, Geraldine Leydon, Adam Geraghty, Carl May, and Mark Gabbay have received grant funding to their employer universities from the National Institute for Health and Social Care Research (NIHR) to carry out this study and other research. In addition Glyn Lewis has received grant funding from the MRC and Wellcome. Mark Gabbay has received consultancy fees from Spectrum Learning and Development as a board member for substance misuse training courses. Tony Kendrick was a member of the NHS England Quality Outcomes Framework (QOF) Advisory Committee 2009-2014, NICE Quality Indicators Advisory Committee 2015-2018, and NICE Depression Guideline Update Committee 2015-2022, and has been a member of the NHS England Improving Access to Psychological Therapies Expert Advisory Committee since 2020. Christopher Dowrick chaired the WONCA Working Party on Mental Health 2016-2021. Glyn Lewis is a member of the NIHR EME Funding Committee. Tony Kendrick, Glyn Lewis, and Michael Moore have been members of trial steering committees for other NIHR funded studies. All the other authors declared no relevant interests.

Acknowledgements: We thank the patients, practitioners, and practices who participated in this research. Tony Kendrick and Adam Geraghty have been supported by the NIHR School for Primary Care research (SPCR). Mark Gabbay is supported by the NIHR Applied Research Collaboration North West Coast (ARC NWC). Glyn Lewis is supported by the University College London Hospital Biomedical Research Centre (UCLH BRC). Thanks to the Trial Steering committee (TSC) members who were: Michael Barkham (academic psychologist and Chair); Susan Collinson (PPI representative); Laura Gray (statistician); Stavros Petrou (health economist); and Linda Gask (academic psychiatrist). Thanks also to the trial Independent Data Monitoring Committee (IDMC) members who were Richard Byng (academic GP and Chair), Jill Mollison (statistician), and Jaime Delgadillo (academic psychologist).

Data availability: Our research ethics permission is for use of the data by the research team only. Applications to use the data need to be in the form of a peer-reviewed protocol, and will be considered by the three lead investigators, Tony Kendrick, Glyn Lewis and Mark Gabbay, in the first instance, before a decision on data sharing will be taken by the rest of the co-authors. Qualitative data will not be shared, due to the relatively high risk of breaches of confidentiality arising from the nature of qualitative interviews, which are difficult to make completely anonymous.

References:

¹Dutch Society of General Practitioners. [NHG Guideline on Depression]. 2022. <https://richtlijnen.nhg.org/standaarden/depressie#volledige-tekst-3-beoordelen-van-de-ernst-van-de-depressieve-stoornis> (accessed 20 Sep 2023).

²US Federal Health Resources and Services Administration. [Uniform Data System Clinical Quality Measures]. 2020. <https://bphc.hrsa.gov/sites/default/files/bphc/data-reporting/2020-clinical-measures-handout.pdf> (accessed 20 Sep 2023).

³Kaiser Permanente Health Maintenance Organisation. [Mental Health Monitoring tool]. n.d. <https://wa-provider.kaiserpermanente.org/static/pdf/provider/patient-ed/screenings/bhi-monitoring-tool.pdf> (accessed 20 Sep 2023).

⁴Mathias SD, Fifer SK, Mazonson PD, Lubeck DP, Buesching DP, Patrick DL. Necessary but not sufficient: the effect of screening and feedback on outcomes of primary care patients with untreated anxiety. *J Gen Intern Med* 1994;9:606-15. doi: 10.1007/BF02600303.

⁵Kendrick T, Stuart B, Leydon GM, et al. Patient-reported outcome measures for monitoring primary care patients with depression: PROMDEP feasibility randomised trial. *BMJ Open* 2017;7(3):1-11. [e015266]. doi: 10.1136/bmjopen-2016-015266.

⁶Wikberg C, Westman J, Petersson EL, et al. Use of a self-rating scale to monitor depression severity in recurrent GP consultations in primary care – does it really make a difference? A randomised controlled study. *BMC Fam Practice* 2017;18:6. doi: 10.1186/s12875-016-0578-9.

⁷Yeung AS, Jing Y, Brenneman SK, et al. Clinical Outcomes in Measurement-based Treatment (Comet): a trial of depression monitoring and feedback to primary care physicians. *Depress Anxiety* 2012;29:865-73. doi: 10.1002/da.21983.

⁸Dowrick C, Leydon GM, McBride A, Howe A, Burgess H, Clarke P, Maisey S, Kendrick T. Patients' and doctors' views on depression severity questionnaires incentivised in UK quality and outcomes framework: qualitative study. *BMJ*. 2009 Mar 19;338. doi: 10.1136/bmj.b663.

⁹Pettersson A, Björkelund C, Petersson EL. To score or not to score: a qualitative study on GPs views on the use of instruments for depression. *Fam Pract* 2014;31(2):215–21. doi: 10.1093/fampra/cmt082.

¹⁰Malpass A, Dowrick C, Gilbody S, Robinson J, Wiles N, Duffy L, Lewis G. Usefulness of PHQ-9 in primary care to determine meaningful symptoms of low mood: a qualitative study. *Br J Gen Pract*. 2016 Feb 1;66(643):e78-84. doi: 10.3399/bjgp16X683473.

¹¹Spitzer RL, Kroenke K, Williams JB, Patient Health Questionnaire Primary Care Study Group, Patient Health Questionnaire Primary Care Study Group. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. *Jama*. 1999 Nov 10;282(18):1737-44. doi: 10.1001/jama.282.18.1737

¹²Kendrick T, Dowrick C, Lewis G, Moore M, Leydon G, Geraghty AW, Griffiths G, Zhu S, Yao G, May C, Gabbay M. Depression follow-up monitoring with the PHQ-9: open cluster-randomised controlled trial. *Br J Gen Pract*. 2024 Feb 23. doi: 10.3399/BJGP.2023.0539.

¹³Kendrick T, Dowrick C, Lewis G, Moore M, Leydon GM, Geraghty AW, Griffiths G, Zhu S, Yao GL, May C, Gabbay M. Patient-reported outcome measures for monitoring primary care patients with

depression: the PROMDEP cluster RCT and economic evaluation. *Health Technol Assess (Winchester, England)*. 2024 Mar;28(17):1. doi: 10.3310/PLRQ4216.

¹⁴Murray E, Treweek S, Pope C, Macfarlane A, Ballini L, Dowrick C, Finch T, Kennedy A, Mair F, O'Donnell C, Ong BN. Normalisation process theory: a framework for developing, evaluating and implementing complex interventions. *BMC Med*. 2010 Dec; 8:1-1. doi: 10.1186/1741-7015-8-63.

¹⁵Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol*. 2006 Jan 1;3(2):77-101. doi: 10.1191/1478088706qp063oa.

¹⁶Boddy CR. Causality in qualitative market and social research. *Qualitative Market Research: An International Journal*. 2019 Jun 10;22(3):405-13. doi: 10.1108/QMR-02-2018-0027.

¹⁷May CR, Mair F, Finch T, MacFarlane A, Dowrick C, Treweek S, Rapley T, Ballini L, Ong BN, Rogers A, Murray E. An interdisciplinary theory of implementation, embedding and integration: the development of normalization process theory. *Imp Sci*. 2009 May 21;4(29). doi: 10.1186/1748-5908-4-29.

¹⁸O'Brien, B. C., Harris, I. B., Beckman, T. J., Reed, D. A., & Cook, D. A. (2014). Standards for reporting qualitative research: a synthesis of recommendations. *Acad Med*, 89(9), 1245-1251. doi: 10.1097/ACM.0000000000000388.

¹⁹Lamb J, Bower P, Rogers A, Dowrick C, Gask L. Access to mental health in primary care: a qualitative meta-synthesis of evidence from the experience of people from 'hard to reach' groups. *Health*. 2012 Jan;16(1):76-104. doi: 10.1177/1363459311403945.

²⁰Pinto-Meza A, Serrano-Blanco A, Penarrubia MT, Blanco E, Haro JM. Assessing depression in primary care with the PHQ-9: can it be carried out over the telephone? *J Gen Med*. 2005 Aug; 20(8): 738-42. doi: 10.1111/j.1525-1497.2005.0144.x.

²¹Perkins K, Wiley S. Minorities. In *Encyclopedia of critical psychology* 2014 (pp. 1192-1195). Springer, New York, NY.

²²Baines R, Tredinnick-Rowe J, Jones R, Chatterjee A. Barriers and enablers in implementing electronic consultations in primary care: scoping review. *J Med Internet Res*. 2020 Nov 12;22(11):e19375. doi: 10.2196/19375.