Caroline Mitchell, Rachel Dwyer, Teresa Hagan and Nigel Mathers

Impact of the QOF and the NICE guideline in the diagnosis and management of depression:

a qualitative study

Abstract

Background
The National Institute for Health and clinical Excellence (NICE) depression guideline (2004) and the updated Quality and Outcomes Framework (QOF) (2006) in general practice have introduced the concepts of screening severity assessment, for example using the Patient Health Questionnaire 9 (PHQ-9), and 'stepped care' for depression.

Aim
To explore primary care practitioner perspectives on the clinical utility of the NICE guideline and the impact of the QOF on diagnosis and management of depression in routine practice.

Design and setting
Qualitative study using focus groups from four multidisciplinary practice teams with diverse populations in south Yorkshire.

Method
Four focus groups were conducted, using a topic guide and audiotaping. There were 38 participants: GPs, nurses, doctors in training, mental health workers, and a manager. Data analysis was iterative and thematic.

Results
The NICE guideline, with its embedded principles of holism and evidence-based practice, was viewed positively but its impact was compromised by resource and practitioner barriers to implementation. The perceived imposition of the screening questions and severity assessments (PHQ-9) with no responsive training had required practitioners to work hard to minimise negative impacts on their work, for example: constantly adapting consultations to tick boxes; avoiding triggering open displays of distress without the time to offer appropriate care; positively managing how their patients were labelled. Further confusion was experienced around the evolving content of psychological interventions for depression.

Conclusion
Organisational barriers to the implementation of the NICE guideline and the limited scope of the QOF highlight the need for policy makers to work more effectively with the complex realities of general practice in order to systematically improve the quality and delivery of 'managed' care for depression.

Keywords
depression; primary health care; qualitative.

INTRODUCTION

Not only does depression affect the quality of life and functioning of individual patients, but the high prevalence and projected increasing disease burden have significant societal and economic implications. For example, depression is in the top three causes of receipt of long-term sickness benefits in the UK,[1] and the World Health Organization estimates that by 2020, major depression will be second only to ischaemic heart disease as the leading cause of disability.[2] At least 80% of depressed patients are managed exclusively in primary care.[3]

Two recent initiatives, the National Institute for Health and Clinical Excellence (NICE) guidelines for the management of depression,[3] and the updated Quality and Outcomes Framework (QOF) in general practice[4] have introduced the concept of 'stepped care' for depression; financial incentives for general practices to administer two depression-screening questions to patients who have diabetes and/or ischaemic heart disease; and standardised measurement of symptom illness severity for patients with new-onset depression, for example the Patient Health Questionnaire 9 (PHQ-9).[5]

In addition, the Improving Access to Psychological Therapies (IAPT) programme (2008), is delivering an unprecedented investment of £170 million in additional primary care mental health workers from 2008–2011.[6]

The introduction of standardised depression-screening and diagnostic questionnaires into routine practice is controversial. A recent Cochrane systematic review found that the use of screening instruments has little impact on the recognition, management, or outcome of depression in primary care and in the general hospital.[7] Previous studies have qualitatively explored GP and patient perspectives on the QOF requirement of using questionnaires for assessing the severity of depression, analysed practice databases to investigate the relationship between management of depression and severity scores and, by survey, examined GP self-reported adherence to the NICE guideline.[8–10] GPs have a strong preference for clinical judgement over scores on depression-severity measures, and 'gaming' has also been described in the decision to code depression. This paper considers the clinical utility and impact of both the NICE and QOF quality initiatives for depression, in routine care, from the perspective of the overall multidisciplinary primary care team.

METHOD

A maximum variation sampling approach was used, based on socioeconomic population characteristics and ethnic

C Mitchell, MD, FRCGP, DRCOG, senior clinical lecturer/GP, Academic Unit of Primary Medical Care, University of Sheffield. **R Dwyer**, BSC, research assistant; **N Mathers**, MD, PhD, FRCGP, DCH, head of unit/GP, Academic Unit of Primary Medical Care, University of Sheffield and Samuel Fox House, Northern General Hospital, Sheffield. **T Hagan**, BA, PhD, MSc, consultant clinical psychologist, Sheffield Health and Social Care NHS Foundation Trust, Community Mental Health Team, Sheffield.

Address for correspondence
Caroline Mitchell, Academic Unit of Primary Medical Care, University of Sheffield, Samuel Fox House, Northern General Hospital, Herries Road, Sheffield, S5 7AU.

E-mail: c.mitchell@sheffield.ac.uk

Submitted: 4 October 2010; **Editor's response:** 9 November 2010; **final acceptance:** 11 January 2011.

How this fits in

Previous research has explored the perspectives of GPs on the use of the Patient Health Questionnaire (PHQ-9); GPs have a strong preference for clinical judgement over scores on depression severity measures, and 'gaming' has also been described in the decision to code depression. The holistic National Institute for Health and Clinical Excellence depression guideline was viewed positively by practitioners, but its impact was compromised by limited resources and application of the stepped-care model. Idiosyncratic and adaptive behaviours reported in the 'routine' use of the PHQ-9, raise doubts about its credibility, reliability, and clinical utility in general practice, as incentivised by the Quality and Outcomes Framework. The introduction of the two depression screening questions without appropriate training had a negative impact on the work of practice nurses, for example, adaptation of chronic disease-management consultations was needed to avoid triggering open displays of distress without the time to offer appropriate care.

diversity (by reference to census data). Multidisciplinary focus groups in four diverse practices were purposively identified, following a first postal invitation to participate sent to 26 practices in south Yorkshire (five responded positively). The four practices (list sizes between 4750 and 8200 patients) comprised:

- one inner-city practice with an ethnically diverse population (the team frequently required translators for consultations);
- two urban practices with average levels of socioeconomic deprivation; and
- one mixed urban/rural practice.

Practice and neighbourhood characteristics are summarised in Tables 1 and 2. All the participating practices had attached counsellors and/or additional primary care mental health workers providing a range of psychological therapies, ('counselling' and cognitive-behavioural-based therapy sessions including guided self-help).

Four focus groups (a minimum of eight and a maximum of 10 participants) were conducted, one in each practice. In total there were 38 individual participants, mostly GPs and practice nurses; however, mental health workers, doctors in training, community nurses, and a manager also attended (Table 3). The topic guide was developed after literature review, discussed within a multidisciplinary group of primary and secondary care clinicians, piloted, and given to participants before the focus group (Appendix 1). Informed consent was obtained for digital recording and verbatim transcription.

The focus groups were led by a trained facilitator, who was identified as an academic GP. Another researcher took contemporaneous field notes (for example observing levels of participation and engagement). The facilitator used the topic guide to ask the groups to describe and explore the impact of the introduction of the NICE guideline and the QOF clinical indicators for depression, and also the factors influencing the implementation of both initiatives. An open questioning style was used, minimising the number of prompts, and allowing the opportunity for participants to discuss key themes.[11,12]

Analysis

Data analysis was independently undertaken by the multidisciplinary research team, and was iterative, thematic,

Table 1. Practice information

Practice focus group	F2/GP registrar training	QOF 2007/2008 depression statistics[a]	
		% of patients at practice with new diagnosis of depression 2007/2008	% of patients at practice with history of depression (prevalence rate)
1	No	0.1	8.5
2	Yes	0.9	10.8
3	Yes	2.1	19.1
4	Yes	2.8	9.6

[a]*Data sources: NHS Information Centre, http://www.ic.nhs.uk; QOF database, http://www.gpcontract.co.uk/.*

Table 2. Neighbourhood information[a]

	Neighbourhood[b]			
Demographic	1	2	3	4
Ethnic group, %				
White	54.0	97.9	96.9	97.3
Mixed	1.9	1.1	0.9	1.0
Asian/Asian British	38.5	0.4	1.0	0.8
Black/Black British	3.0	0.4	0.8	0.7
Chinese or other	2.6	0.3	0.4	0.2
Employment status, %				
Employed	34.9	62.1	63.9	57.2
Unemployed	7.4	4.4	2.4	3.9
Permanently sick/disabled	11.2	6.7	3.9	7.8
Other	30.1	12.1	9.8	11.3
Deprivation quintile[c]	Most deprived	Above average	Average	Above average

[a]Data source: NHS Sheffield, Neighbourhood and services district atlas 2006/2007, http://www.sheffield.nhs.uk/healthdata/atlas_nhoods.php. [b]Based on Census 2001, Office for National Statistics. [c]Based on the proportion of households claiming income support in 2006, (data source: Housing Benefits System, Sheffield City Council 2006).

and self-conscious. QSR NVivo8 software was used to organise transcribed data and field notes. Following each focus group, transcripts were read, and emergent content units were identified, coded, and grouped into themes and later compared across the groups. Each content unit was linked to a referenced item of original data. Analysis meetings were held that included discussion of the observational field notes, systematic data verification, challenging of interpretive analysis within the coding framework, and potential researcher bias. Consideration of reflexivity and potential researcher bias were important, as the facilitator is a local academic GP, and the research team included a researcher new to the area of primary care and a consultant psychologist. A second academic GP provided independent identification of a number of new themes, which were fed back to the other members of the research team for further discussion and also provided verification of the preliminary identified themes.

RESULTS

In summary, the key emergent themes concerned:

- the mechanistic and intrusive impact of administering the PHQ-9 questionnaire during consultations;
- idiosyncratic use of the PHQ-9 depression questionnaire and score interpretation;
- GP strategic labelling of depression;
- nurse-perceived barriers to depression screening within chronic disease-management reviews; and
- resource and organisational barriers to primary care implementation of stepped care.

The mechanistic and intrusive impact of administering the PHQ-9 during consultations

The use of depression-assessment instruments such as the PHQ-9 within the

Table 3. Focus group participants

Primary care team participants	Numbers of professionals (*n* = 38)
GPs	17
Specialist training registrars in general practice	2
Foundation (F2) training GP	2
Practice nurses	11
Community nurses	2
Primary care mental health workers	3
Manager	1

Box 1. Differing modes of administration of the PHQ-9 questionnaire reported

- Self-completion, including question 10.[a]
- Professional completion with the patient, reading out questions using paper forms.
- Professional completion with the patient, reading out the questions using a computer screen (question 10 not on the screen version).
- Using a translator and adapting the wording.
- 'Short cuts': missing out PHQ-9 question 10 as this response isn't needed for the QOF score.
- Recalling questions from memory (to make them fit better with a traditional consultation style) and calculating the score afterwards.
- Telephone administration of the PHQ-9.[a]

[a]*Validated mode of administration of the PHQ-9.*

'10-minute' primary care consultation was a new concept, usually seen by GPs as counterintuitive, intrusive, and unnecessary. One GP explained:

'*... they sort of, you know, break down in tears and tell you how depressed they're feeling and* [they had a] *crap month because lost their job, you do this and you ask them the questions and then "oh now I've got this questionnaire to fill out". I just think it's so inappropriate sometimes ... it's a personal transaction you've had and you know it's not going to change your management because it might be someone you know quite well but you have to, you know, to get the QOF thing you have to do it.*' (GP3 FG2)

And another GP commented:
'*QOF tick-box exercise as far as I'm concerned.*' (GP4 FG4)

Idiosyncratic use of the PHQ-9 depression questionnaire and score interpretation

Adapting the PHQ-9 to 'fit' the consultation style and accommodate diversity. The PHQ-9 was very difficult to administer in an ethnically diverse, non-English-speaking practice population and clearly neither valid nor reliable in this context:

'*It's very very difficult to do a PHQ-9 with an interpreter because actually the words are not right ... there isn't actually a word for depression ...* [the translator] *talks about pressure, not depression, and pressure actually means something different but they talk about it as stress and pressure, not depression. Or we can talk about sadness. Umm, there is, there is, a word for that but that's not quite the same as what you mean, is it?* (GP1 FG1)

The GPs described several adaptive behaviours in their use of the PHQ-9, to make the process less time consuming and more in tune with their usual consulting style. There were wide variations, inter- and intra-GP and inter- and intra-practice, in the mode of administration of the PHQ-9. This theme generated considerable discussion between participants and provided an opportunity to share all the different ways they administered the PHQ-9. An example of one of these discussions is:

GP2: '*That's an interesting one isn't it?*
[group laughing]
GP1: '*I sit with it at my desk with my pen in my hand and I ask them questions and I tick because if I get them to do it, it takes too long.*'
GP3: '*I sometimes get them to go and sit in the waiting room.*'
GP1: '*Yes, I've done that before.*'
GP3: '*And call them back in after my next patient.*'
Female: '*I've done that.*'
GP4: '*I always give it to them to take away ...*' (FG4)

In all, seven different modes of administration of the PHQ-9 were described and were seen as appropriate adaptive behaviours by these GPs (Box 1).

Treat the patient not the score. The potential for missed targets was described as the main impetus to use the PHQ-9, rather than confidence in formal severity assessment as a means of informing management decisions; this is described by one GP:

'*It* [diagnosis] *isn't always entirely linked into a PHQ-9 score because your gut feeling about how depressed someone is and their PHQ-9 score often don't marry up do they? ... so you use your gut feeling much more than you do the strict adherence to the PHQ-9's sort of criteria.*' (GP1 FG4)

GPs described how the administration of the PHQ-9 narrowed the focus of the consultation and reduced the time to deal with the wider context of the illness:

'*... the interesting thing for me is that since the introduction of PHQ-9 I find in terms of material I'm treating the score, not the patient. Because, you know, it's such a sort of barrier in the consultation.*' (GP1 FG3)

GP strategic labelling of depression

The GPs described a range of influences on the actual decision to code and the choice of codes for depressive illness, for example, the QOF; complexity of presenting problems; stigma; and 'watchful waiting' (active monitoring). The QOF provided a disincentive to code 'depression' if a PHQ-9 was not completed by the patient. The use of alternative labels such as 'low mood', 'stress', or other diagnostic codes was

reported, in order to avoid labelling 'mild' symptoms as depression:

'... diagnoses of what would be "QOF-able" depression has probably dropped ... we realised if we kept labelling people as depressed when they perhaps weren't, then we weren't going to see them again and lose the points ... so we had to adapt our coding and I don't know if it's just me but I mean I used to use one of the things "single measure depression mild" that was sort of my preferred code, and that was the wrong code to use because it put everybody on the mental health register.' (GP1 FG2)

Nurse-perceived barriers to depression screening within chronic disease-management reviews — confidence, templates and time constraints

Incorporating depression-screening questions into chronic disease-management consultations was new to the nurses, who felt the questions were imposed, and created additional work, with no responsive training. One practice nurse (PN) said:

'I think we had little education about it really, they've just said this is QOF, this is what you've got to ask and they're the questions. We didn't really have any training.' (PN1 FG2)

Perhaps consequently, nurses found the screening questions intrusive and expressed discomfort in asking patients about low mood, for example:

'It's very difficult because we're supposed to see a large number of people just for one thing and ... you sort of do think twice about asking those questions if you see they need to be asked ... but also I think if you do ask them, then it's very difficult if someone's telling you about some problem, it's very difficult to just fob them off and say "oh well, you can have an appointment". I feel you have to listen, you have to listen there and then and we only ... have 10 minutes.' (PN FG1)

This perceived burden led to the screening questions not being asked in full or being skipped. The practice nurses described concerns that if the questions were asked too early in the 'QOF' chronic-disease-template-driven list of tasks, the patient might become distressed, which would impact on the rest of the consultation and leave insufficient time to complete the review. In an ethnically diverse population, where a telephone translation service was required, the problem was worse, as acknowledged by a practice nurse:

'Yeah but I never get anything else done! Yeah. But I do do, obviously it is at the top of my mind, and I do do it with the people that I know are going to be able to quite quickly brush over it ..., I know that's not good, but that's the pressures of practice nursing, what we've got at the moment with the allotted time that we've got ...' (PN1 FG1)

Community nurses expressed an alternative perspective on screening for depression in people with 'long-term conditions', and less concern about the integration of the screening questions into routine care. For example, a district nurse described more comfort in talking to patients about their mood, less pressure on time, and peer support:

'... we go out and do the house-bound reviews, we do ask it but yet again, I think in a way we've perhaps got a little bit more time than what [the practice nurse] *has 'cos we're not set to set minutes or whatever and ... ummm especially as the majority of people with long-term conditions do have a depressive illness ... we've also got access to case managers and community matrons that step it up a little you know ...'* (district nurse FG1)

Resource and organisational barriers to primary care implementation of stepped care

Poor access to complex non-pharmacological interventions for depression. The NICE stepped-care model, with a renewed focus on non-drug interventions as an adjunct or alternative to antidepressants, was seen as a credible and holistic approach to the management of depression. GPs were keen to avoid 'over-medicalising' and over-prescribing of antidepressants:

'... the big difference to the way we manage is having [the mental health worker] *here more often, because none of us like to*

prescribe antidepressants and yet if you haven't got access to the other psychological therapies then it's hard not to ...' (GP2 FG3)

However, there was a perceived failure of the NHS to provide adequate services to support adherence to the guideline. Practitioners expressed the desire to adapt their practice as much as possible, but poor or absent access to non-drug treatments, for example, the specified range of psychological therapies, tailored exercise programmes, and lengthy waiting times, were the main barriers to the implementation of stepped care. One GP commented:

'... it's interesting when you look at the sort of treatments that they [NICE] *recommend, you know, how many of them are still readily available in primary care? It's, it's very few so it really comes back to you're going to offer patients things which you can actually access.'* (GP3 FG3)

'Bounced back': barriers to 'stepped care' at the interface with secondary care. The NICE guideline presents a stepped-care model of 'seamless' patient care across services. However, practitioners were concerned that there was no 'clear-cut' boundary about where secondary and primary care provision should lie. Specialist secondary care was seen as primarily a crisis-intervention service for actively suicidal patients. GPs and mental health workers described very limited access to specialist input for patients with more complex, treatment-resistant, or recurrent depression (five references to referrals being 'bounced back' across all four focus groups). Discussions took place between individual members of focus groups:

Mental health worker (MHW): *'... if you feel that you need more help then I can refer on. But quite often if you refer on then sometimes they get bounced back.'*
Others: *'mmmm.'*
GP2: *'Always.'*
MHW: *'So it's it's difficult.'* (FG1)

Lack of clarity about how the evolving content of psychological treatments applies to 'stepped care'. Good access to psychological therapies was seen as key to improving outcomes for patients. The IAPT programme has increased the number and range of primary care mental health workers, but there was confusion about the names of these workers and what psychological therapy would actually be provided for patients. For example:

'... we've got a graduate mental health worker so that's CBT [cognitive behavioural therapy] *here. I mean you hear positive things about her from the patients. I don't think we use her enough. She's sort of built up a bit more of a waiting list, because the non-directive counselling's got an enormous waiting list. The CBT person, she's quite constrained. She's not really supposed to see people with a PHQ over 14 anyway.'* (GP3 FG2)

'... a lot of counsellors haven't been specifically trained in CBT therapies but are doing CBT anyway.' (mental health worker FG1)

GP3: *'... CBT is a psychological treatment of choice but — we talk about so many different ...'*
MHW: *'But I mean to get the proper thing I would be thinking of going to* [secondary care outpatients]*, do you know what I mean?'* (FG3)

DISCUSSION

Summary

The holistic NICE depression guideline was viewed positively by practitioners, but its impact was compromised by limited resources and application at practice level. The QOF focus on standardised screening and assessment questions was experienced as unlikely to actually improve the quality of care. The administration of the PHQ-9 interfered with the flow and holistic focus of patient-centred consultation models favoured by GPs, nurses, and mental health workers. Perhaps consequently, the idiosyncratic and adaptive behaviours reported by primary care professionals in their use of the PHQ-9 raises doubts about its credibility, reliability, and clinical utility in routine practice, as incentivised by the QOF.

The perceived imposition of the screening questions with no responsive training had required nurses to work hard to minimise the negative impacts on their work, for

example, constantly adapting consultations to tick boxes and avoiding triggering open displays of distress without the time to offer appropriate care. Further confusion was also experienced around the evolving roles of the mental health workforce and the content of psychological interventions provided in primary care.

Strengths and limitations

Recently published research has explored the perspectives of GPs on the use of the PHQ-9 incentivised by the QOF.[8] The present study, with the emphasis on multidisciplinarity and the whole primary care team, including depression screening by nurses, adds new knowledge about the implementation of the guideline in routine primary care.

This study sought to achieve conceptual transferability rather than statistical generalisability in a purposive sample, using the 'maximum variety' approach of a theoretical sampling framework for sociodemographic patient population characteristics and the disciplines of participating professionals (Tables 1–3).[13–15] The socioeconomic and ethnic diversity of the populations served by the four teams allowed a greater understanding of the challenges of implementing the depression guideline and the QOF in different primary care settings. Specifically, this study reports the difficulties of screening for depression and administering the questionnaires in ethnically diverse populations, using translation support.

The perspectives of a large sample of primary care health professionals is relevant, as achievement of the QOF target requires a 'team approach'; the NICE guideline emphasises that multiprofessional care for depression and collaborative models of care can improve outcomes.[16]

Limitations include the single primary care trust recruitment of participating practices; the organisation of primary care mental health services may differ in other healthcare settings. All practices already had mental health workers (counsellors, graduate mental health workers, and low-/high-intensity mental health workers), but the type of mental health worker and content of the therapies offered was in a process of change. While specific prompts were used in the topic guide to encourage contributions by nurses, mental health workers, and junior doctors in groups where GPs were a vocal majority, observational notes recorded relatively fewer contributions by nurses in one group.

A focus-group methodology was felt to be the most dependable method to explore the implementation of the guideline and the impact of the QOF on the diagnosis and management of depression on the primary care team as a whole, as well as individuals within it, in terms of group norms and individual behaviours. A potential limitation is that there may be a difference between what practitioners report and actually do, while peer dynamics within the group may lead to apparent consensus that does not truly represent the views of participants.

The authors acknowledge the importance of issues of reflexivity, potential researcher bias, and the Hawthorne effect.[17] In a qualitative study assessing the impact of professional identity and status on the research process, the GP qualitative interviewer has been identified as both 'expert and judge'.[18] In the present study, a local GP facilitated the focus groups, which may have influenced the group dynamic. While bias inherent in the individual characteristics of the trained group facilitator cannot be eliminated in qualitative research, there are potential advantages in self-conscious peer facilitation of a health professional group, since there is potential for a more direct understanding of the participants' viewpoints, as well as skilled facilitation of a multidisciplinary group to encourage contributions by less vocal members of the group.[19] The research process also included critical appraisal of the literature, verbatim transcription, researcher observation, recording of the group process, and multidisciplinary analysis (as described in the methodology section) .

Comparison with existing literature

In previous studies, individual practitioner barriers to the implementation of guidelines, for example credibility and dissemination strategies, have been identified as important factors in uptake.[20] In the present study, the NICE depression guideline had less influence on professional practice, because of variable access to many of the complex non-drug interventions described as part of stepped care, a factor that may also explain the lack of impact of the QOF incentives to

score severity on the management of depression in practice.

The importance of 'holism', 'clinical judgement', and GP scepticism about the validity of the PHQ-9 has previously been described in a study of patients' and GPs' views on depression-severity questionnaires.[5] Dowrick describes controversy in 'border' disputes in psychiatric diagnostic classification, the overlap of depressive symptoms with those experienced by patients with physical disorders that are prevalent in undifferentiated primary care consultations, and the importance of both attention to psychosocial factors and a positive practitioner–patient therapeutic relationship.[21] In the present study, GPs and mental health workers considered established patient-centred, non-directive primary care consultation models to be more useful than the narrow, directive focus of the PHQ-9, in the assessment of patients' concerns and expectations, the social context of the illness, and the formulation of a management plan

The consequent adaptive behaviour in administration of the PHQ-9 questionnaire between practices and practitioners may compromise the reliability of the instrument in routine care and further limit the value of the 'score' for management decisions.[22–25] A systematic review of routinely administered questionnaires for depression and anxiety has previously reported a tendency of clinicians to 'ignore raw scores on psychometric questionnaires when they have to add them up and interpret them themselves'.[26]

A practitioner's decision to code an episode of depression is complex. In the present study there was evidence of 'gaming', as previously reported, in the timing of coding, or decision to code, in order to avoid missed QOF targets.[5,27] However, there were other factors influencing the coding decision, such as perceived stigma and 'over-medicalisation' of psychological distress. 'Mild' depression (NICE step 2) would not usually be coded as 'depression'by the GPs in this study, as antidepressants would not usually be prescribed.

Implications for practice and research

In the present study, the QOF rather than the NICE guideline was the catalyst for depression screening and severity assessment. It has been suggested, in other healthcare settings, that quality monitoring should be multidimensional in order to address the variation in process, quality, and cost-effectiveness of primary care for depression, and thus maximise the potential of 'managed' care.[28]

GPs had already adapted the mode of administration of the PHQ-9 to fit established consultation styles at the time of this study and did not explicitly link the score to a stepped-care model. Two authors of the PHQ-9 described the potential for its use as a quantifiable method of monitoring depression, which could be completed 'in the clinic or by telephone (for example, nurse administration or interactive voice recording)' to provide 'an efficient means to assess the number and severity of the nine DSM-IV [Diagnostic and Statistical Manual of Mental Disorders][29] symptoms', akin to a laboratory test for blood glucose.[30] The sample of primary care professionals in the present study was resistant to this 'mechanistic' approach. The most recent update of the QOF (2009/2010) includes a new clinical indicator for depression that adds a repeat depression-severity measure at 5–12 weeks to the initial assessment.[31] It also recommends a revised scoring system for the PHQ-9, based on several studies suggesting that the PHQ-9 overestimates severity compared to another recommended severity-assessment tool, the Hospital Anxiety and Depression Score (HADS).[32–37] This ambiguity in a national policy document may compound the concerns expressed in this study about the credibility of the score, and, in addition, potentially confuse clinicians.

International depression-screening guidelines emphasise that the benefits of screening are unlikely to be realised unless clinical practices have effective systems for accurate diagnosis, follow-up, and access to evidence-based treatments.[38–41] A lack of confidence and education in the area of primary care mental health and depression screening by the practice nurses was identified, and may limit the impact of the QOF national depression-screening programme on chronic disease management and patient outcomes.

Participants in this study expressed confusion about the burgeoning and jargonistic nomenclature for mental health

Funding

The study was funded by the Sheffield Health and Social Research Consortium. The researchers carried out this study independently from the funders. University of Sheffield R/115705-11-1. Grant from Sheffield NHS Health and Social Care Research Consortium/ Woodhouse Health Centre.

Ethical approval

The study had research ethics approval from the South Yorkshire NHS Research Ethics Committee (ref 08/4130/46) and research governance was received from the Sheffield Health and Social Research Consortium.

Competing interests

The authors have declared no competing interests.

Discuss this article

Contribute and read comments about this article on the Discussion Forum: **http://www.rcgp.org.uk/bjgp-discuss**

professionals and described in policy documents; for example, IAPT workers, graduate/primary care mental health workers, CBT techniques, counselling, and guided self-help. There was also a lack of clarity among GPs about what psychological therapies were actually provided after referral to mental health workers. A lack of knowledge about the effectiveness and content of primary care psychological therapies and the gatekeeper burden of limited access have previously been described as factors that may influence the individual GP's decision to refer.[42,43]

There has been a growing international consensus towards a shift from 'ad hoc' acute-episode management of depression to a 'systematic and structured' proactive approach to evidence-based primary care, using chronic disease-management templates of established value for other diseases such as asthma.[44,45] Unfortunately, enhanced treatment of acute depressive episodes in accordance with the 'Wagner criteria' of chronic disease management (by which the UK QOF falls short in ambition),[46] produced no better outcomes at 12 months than usual care.[47]

Further research is needed to evaluate the impact of the use of depression-severity measures and the stepped-care model on the actual outcomes of treatment. The quality of the IAPT and QOF databases may be compromised, unless there is standardisation in the mode of administration of the patient-reported outcome measures.[5,22,23] Implementation of the guideline and QOF in an ethnically diverse population (Table 2) is extraordinarily difficult. If severity-assessment tools are key to quality improvement in the management of depression, development and validation of current patient-reported outcome measures for use in UK minority ethnic groups is appropriate.

There is evidence from the present study that the introduction of the QOF quality targets has failed to positively influence patient access and practitioner behaviour, in terms of referral to the diverse range of psychosocial interventions envisaged by the stepped-care model. Concerns have been reported about the applicability of complex interventions to practice contexts, particularly the challenge of introducing 'complex' interventions due to organisational barriers,[48,49] inadequate description of all the components of the intervention[50] and limited evidence on costs, and the intensity and length of treatments compared to drug treatments.[51] Five years after the national dissemination of the UK NICE guideline, organisational and practitioner barriers to its implementation in primary care, and the limited scope of the QOF, have compromised the delivery of the policy goals of evidence-based, primary care provision based on an individual patient's depression as well as their personal and social circumstances.[3]

Pawson and Tilley remind us that 'a particular programme will only 'work' if the contextual conditions into which it is inserted are conducive to its implementation as it is put into practice.[52] This study has helped to identify factors that can operate to impede or help the effectiveness of any proposed changes in practice aimed at improving care for people with depression and alert us to the need for policy makers to work with the complex realities of primary care in dialogue with practitioners and healthcare commissioners.

REFERENCES

1. J Moncrieff, Pomerleau J. Trends in sickness benefits in Great Britain and the contribution of mental disorders. *J Public Health* 2000; **22(1):** 59–67.
2. Murray C, Lopez A. Global mortality, disability, and the contribution of risk factors: Global Burden of Disease Study. *Lancet* 1997; **349(9063):** 1436–1442.
3. National Institute for Health and Clinical Excellence. *Depression: management of depression in primary and secondary care. National clinical practice guideline 23.* London: National Institute for Health and Clinical Excellence, 2004.
4. British Medical Association, NHS Employers. *Revisions to the GMS contract 2006/07.* London: British Medical Association, 2006.
5. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001; **16(9):** 606–613.
6. Byng R, Gask L. Improving access to psychological therapies: implications for mental health care in general practice. *Br J Gen Pract* 2009; **59(566):** 640–641
7. Gilbody S, House A, Sheldon T. Screening and case finding instruments for depression. *Cochrane Database Syst Rev* 2005; **(4):** CD002792.
8. Dowrick C, Leydon GM, McBride A, *et al.* Patients' and doctors' views on depression severity questionnaires incentivised in UK quality and outcomes framework: qualitative study. *BMJ* 2009; **338:** b663.
9. Kendrick T, Dowrick C, McBride A, *et al.* Management of depression in UK general practice in relation to scores on depression severity questionnaires: analysis of medical record data. *BMJ* 2009; **338:** b750.
10. Toner R, Snape C, Acton S, Blenkiron P. Do general practitioners adhere to NICE guidelines for depression? Systematic questionnaire survey. *Prim Health Care Res Dev* 2010; **11:** 123–131
11. Kitzinger J. Qualitative research. Introducing focus groups. *BMJ* 1995; **311(7000):** 299–302.
12. Kamberelis G, Dimitriadis G. Focus groups: strategic articulations of pedagogy, politics and inquiry. In: Denzin NK, Lincoln YS (eds). *The Sage handbook of qualitative research. 3rd edn.* London: Sage Publications Inc, 2008: 887–907.
13. Pope C, Mays N. Reaching the parts other methods cannot reach: an introduction to qualitative methods in health and health services research. *BMJ* 1995; **311(6996):** 42–45.
14. Mays N, Pope C. Qualitative research in health care: assessing quality in qualitative research. *BMJ* 2000; **320(7226):** 50–52.
15. Lincoln YS, Guba EG. *Naturalistic enquiry.* Newbury Park, CA: Sage, 1985.
16. Dowrick C. Reasons to be cheerful? Reflections on GPs' responses to depression. *Br J Gen Pract* 2009; **59(566):** 636–637.
17. Holden JD. Hawthorne effects and research into professional practice. *J Eval Clin Pract* 2001; **7(1):** 65–70.
18. Chew-Graham CA, May CR, Perry MS. Qualitative research and problems of judgement: lessons from interviewing fellow professionals. *Fam Pract* 2002; **19(3):** 285–289.
19. Hewitt J. Ethical components of researcher researched relationships in qualitative interviewing. *Qual Health Res* 2007; **17(8):** 1149–1159.
20. Rashidian A, Eccles MP, Russell I. Falling on stony ground? A qualitative study of implementation of clinical guidelines' prescribing recommendations in primary care. *Health Policy* 2008; **85(2):** 148–161.
21. Dowrick C. *Beyond depression: a new approach to understanding and management. 2nd edn.* Oxford: Oxford University Press, 2009.
22. Bowling A. *Measuring health: a review of quality of life measurement scales. 3rd edn.* Oxford: Oxford University Press, 2005.
23. McDowell I. *Measuring health: a guide to rating scales and questionnaires. 3rd edn.* Oxford: Oxford University Press, 2006.
24. Spitzer RL, Kroenke K, Williams JBW, *et al.* Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. *JAMA* 1999; **282(18):** 1737–1744.
25. Spitzer RL, Williams JBW, Kroenke K, *et al.* Validity and utility of the PRIME-MD Patient Health Questionnaire in assessment of 3000 obstetric-gynecologic patients: The PRIME-MD Patient Health Questionnaire Obstetrics-Gynecology Study. *Am J Obstet Gynecol* 2000; **183(3):** 759–769.
26. Gilbody S, House AO, Sheldon T. Routinely administered questionnaires for depression and anxiety: systematic review. *BMJ* 2001; **322(7283):** 406–409.
27. Davies HTO, Kinloch C. Interpreting health outcomes. *J Eval Clin Pract* 1997; **3(3):** 187–199.
28. Well KB, Schoenbaum M, Unutzer J, *et al.* Quality of care for primary care patients with depression in managed care. *Arch Fam Med* 1999; **8(6):** 529–536
29. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders (DSM-IV). 4th edn.* Arlington, USA: American Psychiatric Publishing, Inc, 1994.
30. Kroenke K, Spitzer RL. The PHQ-9: a new depression diagnostic and severity measure. *Psychiatr Ann* 2002; **32:** 509–521.
31. British Medical Association, NHS Employers. *Quality and Outcomes Framework guidance for GMS contract 2009/10.* London: British Medical Association, 2009.
32. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand* 1983; **67(6):** 361–370.
33. Lowe B, Spitzer RL, Grafe K, *et al.* Comparative validity of three screening questionnaires for DSM-IV depressive disorders and physicians' diagnoses. *J Affect Disord* 2004; **78(2):** 131–140.
34. Gilbody S, Richards D, Barkham M. Diagnosing depression in primary care using self-completed instruments: UK validation of PHQ-9 and CORE-OM. *Br J Gen Pract* 2007; **57(541):** 650–652.
35. Cameron IM, Crawford JR, Lawton K, Reid IC. Psychometric comparison of PHQ-9 and HADS for measuring depression severity in primary care. *Br J Gen Pract* 2008; **58(546):** 32–36.
36. Hansson M, Chotai J, Nordstom A, Bodlund O. Comparison of two self-rating scales to detect depression: HADS and PHQ-9. *Br J Gen Pract* 2009; **59(566):** 650–654.
37. Lowe B, Unutzer J, Callahan CM, *et al.* Monitoring depression treatment outcomes with the patient health questionnaire-9. *Med Care* 2004; **42(12):** 1194–1201.
38. Whooley MA, Avins AL, Miranda J, Browner WS. Case-finding instruments for depression; two questions are as good as many. *J Gen Intern Med* 1997; **12(7):** 439–445.
39. Barth J, Schumacher M, Herrmann-Lingen C. Depression as a risk factor for mortality in patients with coronary heart disease: a meta-analysis. *Psychosom Med* 2004; **66(6):** 802–813.
40. Anderson RJ, Freedland KE, Clouse RA, Lustman PJ. The prevalence of comorbid depression in adults with diabetes. *Diabetes Care* 2001; **24(6):** 1069–1078.
41. The US Preventive services Task Force. Screening for depression in adults. *Ann Intern Med* 2002; **136:** 760–764.
42. Stavrou S, Cape J, Barker C. Decisions about referrals for psychological therapies: a matched-patient qualitative study. *Br J Gen Pract* 2009; **59(566):** e289–298.
43. Mathers N, Mitchell C. Are the gates to be thrown open? *Br J Gen Pract* 2010; **60(574):** 317–318.
44. Scott J. Depression should be managed like a chronic disease. *BMJ* 2006; **332(7548):** 985–986.
45. Andrews G. Should depression be managed as a chronic disease? *BMJ* 2001; **322(7283):** 419–421.
46. Wagner EH, Austin BT, Von Korff M. Organizing care for patients with chronic illness. *Millbank Q* 1996; **74(4):** 511–543.
47. Lin EHB, Simon GE, Katon WJ, *et al.* Can enhanced acute-phase treatment of depression improve long-term outcomes? A report of randomized trials in primary care. *Am J Psychiatry* 1999; **156(4):** 643–645.
48. Greenhalgh T. Evidence based medicine: new storylines needed. *BMJ* 2009; **339:** b3296.
48. Thompson C, Kinmonth AL, Stevens L, *et al.* Effects of a clinical-practice guideline and practice-based education on detection and outcome of depression in primary care: Hampshire Depression Project randomised controlled trial. *Lancet* 2000; **355(9199):** 185–191.
49. Thomson R. Evidence based implementation of complex interventions. *BMJ* 2009; **339:** b3124.
50. Glasziou P, Meats E, Heneghan C, Shepperd S. What is missing from descriptions of treatment in trials and reviews? *BMJ* 2008; **336(7659):** 1472–1474.
51. Arroll B, Moir F. Time for a rethink of treatment for patients with depression in primary care. *Br J Gen Pract* 2010: **60(578):** 641–642.
52. Pawson R, Tilley N. *Realistic evaluation.* London: Sage Publications Ltd, 1997.

Appendix 1. Practitioner focus group topic guide

Research project title: Quality of care for depression in general practice: patient and professional perspectives (ZJ75)

Below is a list of questions that the group might discuss. If there are any questions that you do not want to talk about then you do not have to.

Key question for the group: National Institute for Health and Clinical Excellence (NICE) guidelines for the management of depression and the Quality and Outcomes Framework (QOF) in general practice — can we do it, does it work?

- Have you any experience of the NICE guidelines changing practice for the treatment of depression? (Is there anything you've done, seen, read, or heard about?)
- Have you any experience of the QOF in general practice changing practice for the treatment of depression? (Is there anything you've done, seen, read, or heard about?)
- What has been your experience of incorporating depression questionnaires (for example, PHQ-9) into consultations?
- What has been your experience of incorporating the two depression screening questions into consultations and diabetes and heart disease checks?
- Is the NICE recommended range of treatments available (for example, counselling, primary care mental health worker, cognitive behavioural therapy, exercise)?
- Have you experienced any sort of waiting times for referral processes or the different treatments?
- Who are the patients you refer to secondary care psychology and what prompts you to make that referral?
- Are there any other facilitators or barriers to achieving the requirements of the NICE guidelines or QOF?
- Does it make a difference to patient outcomes when NICE guidelines are adhered to in primary care?
- Have any new ideas been generated during your participation in this group?