The QOF, NICE, and depression:

a clumsy mechanism that undermines clinical judgment

The National Institute for Health and Clinical Excellence (NICE) has commissioned guidelines for the treatment and management of depression in adults from the National Collaborating Centre for Mental Health. The full 700-page updated edition (guideline 90)¹ was published in 2010, replacing the 2004 guideline 23 which has been credited with providing the rationale behind the Quality and Outcomes Framework (QOF) depression indicators.² Guideline 90 now eschews population or opportunistic screening, instead promoting increased awareness and case finding of depression. A step-wise approach to treatment based on semi-structured severity assessment is also recommended with an emphasis on the importance of considering functional impairment.

Guideline 90 is an interesting if substantial read. It contains a collection of viewpoints which collectively provide an excellent example of how and why general practice can be an uncomfortable and confusing place. How should it respond to simultaneous accusations of over and under diagnosing depression, over and under treating, and of being both under and over structured in its approach to a 'condition' that, by its nature, is impossible to define clearly, and which at one end overlaps with the emotional ups and downs of normal human existence in an increasingly stressful society?

How then was a controversial part of a NICE guideline allowed to become enshrined into one-size-fits-all indicators for the uncontrolled pay-for-performance experiment that is the QOF?³

The evidence to support a population or opportunistic screening approach for recognising depression in general practice probably never existed. It has been extensively reviewed in the UK, and neither approach is recommended as they do not fulfil the criteria of the UK National Screening Committee (NSC).^{4,5} However, with seemingly poor evidential justification, the systematic application of 'screening questions' in the guise of 'case finding' for those with two important but seemingly arbitrarily-chosen chronic diseases (diabetes and heart disease) arrived through the QOF back door.

This approach has been given an unexplained and, it appears, uncomfortable

evidence waiver by the NSC,⁵ which runs for cover, deferring instead to the NICE guidance. That guidance can no longer be marshalled as support, as guideline 90 now aligns with the NSC view. Indeed, it is less than enthusiastic on the QOF case-finding approach, acknowledging the questions around the usefulness of ultra short questionnaires and the danger of encouraging an oversimplified and reductionist, biomedical approach.¹

UNINTENDED CONSEQUENCES

Many recent qualitative studies, including several published in the *BJGP*, help us to understand the effects and unintended consequences of introducing the QOF depression indicators to general practice under the QOF. Themes are emerging of deleterious effects on the consultation and the work of both GPs and practice nurses, and growing qualms about the severity scales used to inform decision making.⁶⁻⁹

Interviews with doctors and nurses have demonstrated a common theme of frustration at the intrusion of the QOF measurements into consultations. As soon as the possibility of depressive symptoms are raised, decisions have to be made: not so much about how to listen and to offer help, but rather about how to introduce the screening tools without disrupting the flow and direction of the consultation. Nurses also report difficulty in introducing screening and formulaic severity assessment questions into the limited time they have available to run chronic disease consultations. They express concern they may enter areas of discussion that they have neither the time nor the training to deal with.⁷

To be less intrusive, clinicians have found creative and adaptive ways to administer the severity scores (by their own admission on occasions solely to gain QOF points). Unfortunately, these generally involve using the instruments in ways for which they were never designed or validated. The additional question 10 of the Patient Health

Questionnaire-9 on functional impairment is not part of the QOF, and there are reports of practitioners therefore not recording it, despite its increased emphasis in the updated NICE guideline. If these workaround strategies are commonplace, subsequent evaluations of the benefit of implementing these indicators on patient outcomes will be compromised by further reduction of the credibility, reliability, and clinical utility of the severity score data.⁷

Further evidence appears in the paper by Cameron *et al* in this month's Journal demonstrating a poor correlation between all of the QOF-approved severity screening questionnaires with the Hamilton Rating Scale for Depression-17, which underpins the stepped-care model.¹⁰ This increases doubt about their validity as severity-based general practice decision support tools, which is worrying as this severity-based treatment is a central plank of the NICE recommendations.^{1,10} None of this will be a surprise to GPs who have long since observed that the scores are counterintuitive, intrusive, and unnecessary. They have (notwithstanding QOF point collection) continued throughout to rely more on their clinical judgement.⁷

AN OPPORTUNITY FOR REVIEW

A set of financially-incentivised indicators have been put in place promoting a 'screening by the back door' approach which lacks a sound evidence base^{1,4,11} and which mandates the use of severity scoring tools which are seen by those using them as an impediment, rather than an aid, to good care⁶⁻⁸ and which now seem, in the primary care setting at least, to be unfit for their stated purpose.^{7,10} The chain of adaptive behaviours, and unintended consequences that follow on, are all too familiar with QOF and are ultimately deprofessionalising.

The only positive in this story is that the relevant QOF committee is currently reviewing the depression indicators in the face of vociferous opposition. There is also accumulating evidence that GPs continue to

"To be less intrusive, clinicians have found creative and adaptive ways to administer the severity scores (by their own admission on occasions solely to gain QOF points)."

"Wisdom, clinical experience, and judgment must remain the central foundations on which often transient evidence is laid"

rely on their clinical judgement to decide who, when, and how to manage their patients with symptoms of depression. Wisdom, clinical experience, and judgement must remain the central foundations on which often transient evidence is laid.^{6,9} Shortcuts and checklists must be properly piloted prior to introduction^{6,8} and their use must never be mandated through coercive financial incentives.

MOVING ON FROM QOF

That experienced clinicians yet again feel duty bound to work round the perverse incentives inherent in so many QOF indicators surely continues to prompt the question: 'Is this process-driven quality framework of net benefit?'.¹² If the answer is 'yes', then surely by now the QOF architects and supporters should have sufficient confidence that the intrinsic worth of the indicators are compelling enough for the activity to be self-sustaining without the need for financial incentive. If the answer is 'no', then why persist with it? Either way, why not adopt a hands off approach? Unlike QOF, a rigorously evaluated pilot should precede widespread adoption.

Why not redirect decisions on the distribution of the incentive pay-for-performance funding to the combined wisdom of local or regional consortia of providers and end users of the system? Is there the leadership, the courage, and sufficient trust to allow such a powerful coalition to decide how this funding should be used to promote meaningful quality outcomes. Perhaps one day there will be.

A greater recognition, understanding, and acceptance is needed of the complex range of skills employed by general practice teams in helping patients across the spectrum: from normal fluctuations in mood through to serious depressive illness. We need to develop sustainable, empathic, and implementable strategies to help all patients. The QOF depression indicators have become a distraction and a barrier rather than an aid to quality care. As in 2007, this pay-for-performance system remains Quite Obviously Flawed.¹²

Les Toop,

Professor, University of Otago, Christchurch, New Zealand.

Provenance

Commissioned; not externally peer reviewed.

DOI: 10.3399/bjgp11X582994

ADDRESS FOR CORRESPONDENCE

Les Toop

University of Otago, Christchurch, 1 Seamount Terrace, Mount Pleasant, Christchurch, 8081, New Zealand.

E-mail: les.toop@otago.ac.nz

REFERENCES

1. National Collaborating Centre for Mental Health. *Depression: the treatment and management of depression in adults (updated edition). National Clinical Practice Guideline 90*. London: The British Psychological Society and The Royal College of Psychiatrists, 2010.
2. BMA, NHS Employers. *Quality and Outcomes Framework guidance for GMS contract 2009/10*. London: British Medical Association, 2009.
3. Glasziou P. Punished by rewards [rapid response]. *BMJ* 2007. http://www.bmj.com/content/335/7629/1075.full/reply#bmj_el_181700 [accessed 1 Jun 2011].
4. Gilbody S, House A, Sheldon T. Screening and case finding instruments for depression. *Cochrane Database Syst Rev* 2005; **(4)**: CD002792.
5. Allaby M. *Revised report. Screening for depression: a report for the UK National Screening Committee*. 2010. http://www.screening.nhs.uk/policydb_download.php?doc=73 [accessed 1 Jun 2011].
6. 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. doi: 10.1136/bmj.b663.
7. Mitchell C, Dwyer R, Hagan T, Mathers N. Impact of the QOF and the NICE guideline in the diagnosis and management of depression: a qualitative study. *Br J Gen Pract* 2011; doi: 10.3399/bjgp11X572472
8. Leydon G, Dowrick C, et al. Questionnaire severity measures for depression: a threat to the doctor-patient relationship? *Br J Gen Pract* 2011; **61**: 117-123.
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 doi: 10.1136/bmj.b750.
10. Cameron IM, Cardy A, Crawford JR, et al. Measuring depression severity in general practice: discriminatory performance of the PHQ-9, HADS-D, and BDI-II. *Br J Gen Pract* 2011; doi: 10.3399/bjgp11X583209.
11. Gillies J, Mercer S, Watt GCM, et al. Questionnaire severity measures for depression. *Br J Gen Pract* 2011; **61(586)**: 324-325.
12. Mangin D, Toop L. The Quality and Outcomes Framework: what have you done to yourselves? *Br J Gen Pract* 2007; **57(539)**: 435-437.