

International research

Making decisions about patient progress: the application of routine outcome measurement in stepped care psychological therapy services

Peter Bower BSc PhD

Senior Research Fellow, National Primary Care Research and Development Centre, University of Manchester, UK

Simon Gilbody BSc MBChB MSc MRCPsych Dphil DipLSHTM PGDip

Senior Lecturer in Mental Health Services Research, Department of Health Sciences, University of York, UK

Michael Barkham BEd MA MSc PhD

Professor of Clinical and Counselling Psychology, Psychological Therapy Research Centre, University of Leeds, UK

ABSTRACT

In a stepped care system, initial treatment is the most efficient intervention that is still likely to be effective. Patient outcome is then monitored, and patients are 'stepped up' to a more intensive treatment if the initial treatment is deemed to be unsuccessful. Stepped care has been proposed as a method of organising psychological therapy services that provides the optimal balance between effectiveness and access. However, the practical application of stepped care is complex, and a key challenge is making effective decisions about which patients are 'stepped up' to more complex and intensive services in the stepped care pathway. Routine outcome measures have been proposed as one method of improving the delivery of

psychological therapy services, and in this article, we set out a range of issues concerning the use of routine outcome measures in support of stepped care within mental health. These include measurement issues (such as reliability, validity and acceptability), and the key issues relating to their use in decision making, including the role of patient preferences and choice. We conclude with recommendations for researchers, practitioners, and measure developers to maximise the yield from routine outcome measurements in the implementation of stepped care.

Keywords: decision making, outcome measurement, stepped care

Introduction

There is a wealth of literature attesting to the efficacy of a range of psychological therapies within the context of controlled trials, the foundation of evidence-based practice.¹ In parallel, there is a growing body of evidence reporting the effectiveness of psychological therapies delivered in routine practice

settings, the foundation of practice-based evidence.² However, comparisons between the effectiveness of treatments in the context of controlled trials and routine practice yield inconsistent results. Some reviews have suggested that results from controlled trials in the psychological therapies closely mimic

routine practice, in that the patients are clinically representative.³ However, other studies have yielded smaller effects for routine practice.^{4,5}

One reason for the occurrence of lower rates of improvement in routine settings may be that the interventions in these settings are not delivered within a clearly defined protocol. One of the functions of treatment protocols within trials is to help reduce the role of error through more systematic delivery of treatment. The introduction of appropriate protocols within routine practice settings might similarly enhance effectiveness.

The role of protocols has been highlighted by current developments in psychological therapy delivery in the UK, which has seen a shift away from concerns about clinical effectiveness alone, towards wider issues of demand, supply and the efficient use of finite psychological therapy resources to meet population needs. Recent work has highlighted the potential for stepped care models of service delivery, using minimal interventions and guided self-help, as well as brief and longer-term psychological treatments.⁶

In a stepped care system, the firstline treatment choice is the most efficient intervention that is still

likely to be effective. Following this firstline treatment, patient progress is systematically monitored, and if the treatment is deemed to have been unsuccessful, patients are 'stepped up' to a more intensive treatment.⁷⁻⁹

Figure 1 shows a possible five-level stepped care system for depression proposed by the National Institute for Clinical Excellence (NICE) in the UK.¹⁰ As an example, a patient with mild to moderate symptoms, demonstrating low risk of self-harm might be initially dealt with through watchful waiting by the general practitioner (GP). If an objective assessment showed that symptoms had persisted for 2-3 weeks despite support from the GP, this patient might then be offered guided self-help from a graduate worker. After receipt of this treatment and another objective assessment, a decision would again be required as to whether the patient was 'stepped up' to brief cognitive-behavioural therapy (CBT) from a more experienced practitioner, or whether the benefit received from the self-help treatment was sufficient to end the treatment episode at that point.

There are many issues to consider in the development of effective stepped care systems,⁶ but one

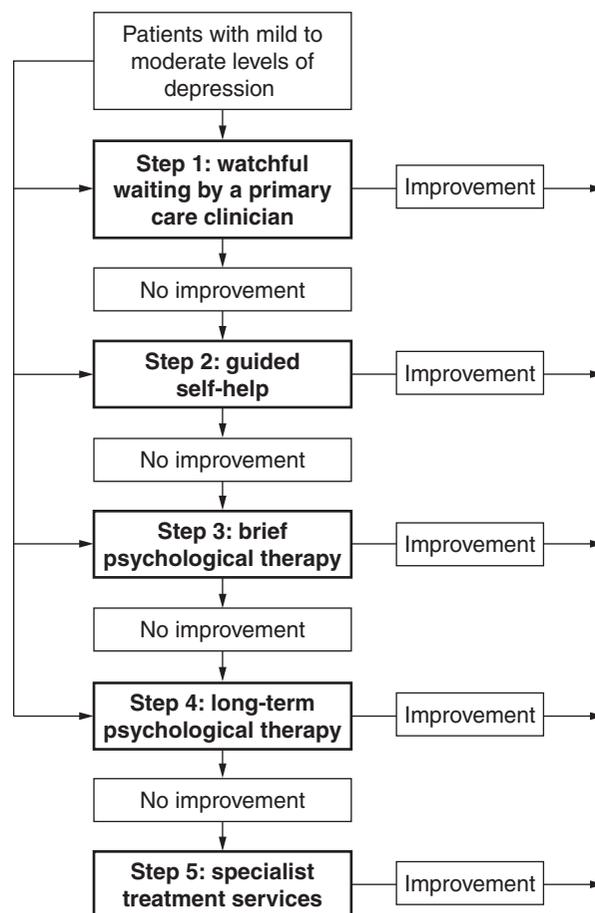


Figure 1 Stepped care system for depression

particularly pressing issue is how decisions are to be made about whether patients have had a 'good' outcome within a step, and whether they need 'stepping up' to a more intensive form of treatment. At present there is limited guidance as to how such decisions should be made.

Routine outcome measurement is one technology that may be important in stepped care. The Department of Health Practice Guidance on outcomes views routine outcome measurement technology as having significant potential for quality improvement in the NHS.¹¹ Currently there are a range of outcome measures that are used in routine practice settings. These include generic self-report measures such as the Clinical Outcomes in Routine Evaluation – Outcome Measure (CORE-OM) and How are you (HRU) which have been developed in the UK,^{2,12,13} and equivalent measures from the US, such as the Beck Depression Inventory (BDI) and the Patient Health Questionnaire (PHQ)-9 which target specific diagnostic presentations such as depression.^{14,15}

The current article considers how such measures used routinely in practice settings could help inform decision making within stepped care.

The aim of routine outcome measurement within stepped care

The aim of routine outcome measurement within stepped care is effective 'decision support', i.e. to support decision-making activities about 'stepping'. For any outcome measure to be useful, it needs to improve upon routine judgements, to identify those patients with a 'good' outcome, and ensure that only those patients who fail to demonstrate such an outcome are stepped up to further care.

The process of routine outcome measurement within stepped care is a two-stage one. The first stage is the measurement of outcome, while the second stage is the decision making about stepping, based on that measurement. Each of these will be dealt with in turn.

Measurement of outcome within stepped care

Clearly, there is an enormous literature on the assessment of the quality of outcome measures. However, the current discussion will highlight those issues that are particularly relevant to stepped care.

Reliability

As with any measure, it is necessary for routine outcome measures to show good evidence of reliability. Low levels of reliability mean that scores made on different occasions will vary significantly, which means that decisions about 'stepping up' would not be on a firm basis. Indeed, the levels of reliability of measures that are designed to make decisions about individuals may be higher than those conventionally required in research studies where group comparisons are made.¹⁶

Validity

Traditionally, any measure is expected to show evidence of 'validity', which is often referred to as a demonstration that a scale 'is measuring what it is supposed to measure'. Although superficially useful, in the current context this definition begs the question of how a 'good outcome' in mental health is defined. There are a number of issues of relevance.

The first issue is the source of the data, which may include the patient, significant others, practitioners, and service managers. Each perspective has advantages and disadvantages. For example, clearly there is no better measure of the subjective impact of an intervention than patient self-report. However, patients' judgements of the benefits of treatment may be made without knowledge of normative considerations such as the expected benefit from treatment or the response of other patients, and therefore may not be sufficient, given the focus on population health and the efficient use of resources, to meet wider need in stepped care.

The second issue is the content of the measure. The traditional content of routine outcome measures has been symptoms.¹⁷ To the degree that the appropriate level of stepped care may be determined to a large extent by issues of symptom severity and complexity, this may be appropriate.

It may be that more than a unitary symptom score is required in order to capture the potential complexity of many cases. For example, a single score on a measure can be arrived at via numerous combinations of single-item responses. Although some variation is part of the diversity of clinical presentation, systematic variations in how the same score is achieved may reflect more significant underlying factors, which may have implications for assignment to the level of stepped care. For example, a similar score may be achieved via relatively high scores on a cluster of items, or by low scores on a broad spread of items. The former might indicate a relatively circumscribed focus, while the latter might be indicative of a more pervasive pathology requiring a greater level of professional input.

Moving beyond symptoms, there is also increasing interest in a wider range of domains of outcome, including issues such as social functioning, satisfaction, and new user-defined outcomes such as 'recovery'.¹¹ There are also other models of the change process within psychological therapy which might provide an alternative perspective, such as the assimilation model, the phase model, and the stages of change model.¹⁸⁻²⁰ Each is premised on an assumption that there are discrete stages or phases in the progression from distress to health, and might provide another way of measuring outcomes.

Many of the measures available are standardised, in the sense that patients are asked to complete a set series of questions, with responses combined in a scoring algorithm that has been defined by the measure developers. However, given current interest in making services more responsive to patients, there is interest in the possibility of weighting the scoring of measures to reflect the importance placed on them by patients rather than measure developers. This is the approach generally taken by utility measures, which use weights derived from sources such as general population surveys to calculate a utility score based on certain patterns of symptoms.²¹ Techniques such as discrete choice experiments may be useful ways of examining what aspects of outcome are most important to patients.²²

Although such methods may make scales more responsive to the general views of patients, an even more radical approach is the use of individualised patient-defined measures of outcome. This involves individual patients selecting the particular issues or domains that are most important to them, or defining personal treatment goals to be attained.^{23,24} Rather than imposing an external standard, such measures are highly likely to reflect the preferences and needs of the individual patient. However, this comes at the cost of increased complexity in their administration.

Acceptability

Traditional criteria for the assessment of measurement technologies tend to focus on the psychometric properties of individual instruments.²⁵ However, for an instrument to be clinically useful, it must be simple and quick to administer, and should be acceptable to the patient and practitioner alike.²⁶ Stepped care is a method of improving the efficiency of services, and when a measurement technology becomes expensive to administer, the system becomes cumbersome and inefficient. This is especially important at the lower levels of stepped care, where rapid decisions must be made for large volumes of patients. Inefficient measurement at this stage will quickly block flows through a stepped care system.

The 'outcomes movement' has grappled with how to reflect complexity, while maintaining the idea that tools and instruments must be brief and informative in every day clinical decision making.^{26,27} This idea was neatly summed up by Kane, who described an optimal measure as being: 'the clinical equivalent of the Swiss Army knife – something small and easily taken into the field, with enough blades and attachments to fit any number of circumstances that may arise'.²⁸ An example of this is that any 'systems approach' to the organisation and delivery of care (such as stepped care) must bridge primary and secondary care settings. Therefore, any routine outcome measurement programme must incorporate measurement that is suitable for use in both primary and secondary care, and which is able to reflect varying severity and complexity of problems across these two settings.

Relationships between the criteria

There are clearly tensions between the reliability, validity and acceptability criteria. The most valid and reliable measures might in fact be the most impractical, unwieldy and unacceptable. There is also some suggestion that techniques for ensuring the reliability of measures can have a negative impact on their validity and sensitivity to change.²⁹ Making appropriate tradeoffs between these criteria is a key challenge for the development of useful measures.

Decision making in stepped care

Given that a measurement has been made of patient outcome, how is that to be used in decision making?

In traditional evidence-based medicine, decisions about treatment are based on knowledge from controlled trials of the effectiveness of treatment. That is, once a treatment has been 'empirically validated', it is recommended within guidelines and provided to patients on the basis of that previous empirical validation. In contrast, stepped care requires that each individual patient's outcome is measured, and placed against some criteria of psychological functioning. It is most likely that this will involve normative criteria derived from external sources. These criteria may relate to (a) the treatment; (b) the patient population; or (c) the specific disorder. Examples of each of these criteria are described below.

In terms of a treatment-based criterion, a particular psychological therapy might be generally associated with a reduction of a certain number of points on a symptom-based measure in a particular population, and patients showing a similar reduction would be judged to have demonstrated the appropriate outcome for that treatment.

A criterion based on patient populations would work through the empirical identification of population norms. One of the most frequently used normative measures is the reliable and clinically significant change (RCSC) criterion.³⁰ The primary use of the RCSC criterion has been to provide a firmer empirical basis upon which judgements of improvement can be made, based on changes that are needed to move a patient from membership of one population (e.g. clinical) to another population (e.g. general population). In principle, this provides a method to determine the presence or absence of criteria that would result in 'stepping up'.

The final alternative is the use of disorder-specific criteria. For example, in the case of depression, research has indicated that residual symptoms are associated with later relapse. Because of this, treatment until full remission has been recommended, and might be the basis for decision making.⁹

One limitation of all these normative methods is that they define outcome as relative to the performance of others, and may be fairly insensitive to the particular characteristics of the patient. A more patient-centred approach might be the use of individualised 'expected outcomes' based on patient pre-treatment variables.³¹ This relates to work on patient profiling, which seeks to estimate a patient's expected outcome during treatment based on baseline variables, and then compare this estimate with the actual outcome observed, in order to determine the success or otherwise of the particular treatment.

Thresholds in stepped care decision making

In stepped care systems there exists an inherent trade-off between false positives (i.e. inappropriate 'stepping up' when it is not required) and false negatives (failure to 'step up' when appropriate). By adopting very conservative cut-off points for 'stepping up', false positives can be avoided, but many patients with unmet needs will languish at lower levels for too long. Conversely, by adopting more liberal criteria, many people will be stepped up too quickly, rendering the system unwieldy, unworkable and inefficient.

Here, much can be learned from evidence-based criteria that have been developed in assessing whether or not a screening technology should be adopted. Here it is recognised that there is no such thing as a perfect test, that explicit trade-offs between sensitivity and specificity must be made, and that there are mathematical and practical considerations in setting decision rules and adopting a screening technology.³² The application of these criteria in stepped care represents an unexplored issue.

Clearly, the development of thresholds within stepped care may be even more complex than a

simple binary approach. Using the example of depression, a number of differing populations are of relevance on the continuum of severity, such as the non-distressed, the general population, and then mild, moderate, and severely distressed, and people with more chronic, severe and enduring problems. It will be necessary for measures to define with greater precision the various cut-off points between key subpopulations. This will then be able to guide decision making on movement between the different steps of the system.

The role of clinical judgement in decision making in stepped care

As noted earlier, the role of routine outcome measurement within stepped care is 'decision support'. However, the exact relationship between routine outcome measures and clinician judgement in stepped care decision making is unclear. There are those that claim that routine outcome measures are only ever an adjunct to clinical judgement, and would resist the use of 'relatively untrained personnel making judgements solely on the basis of questionnaires'.⁸ However, others have criticised the supposed advantages of clinical judgement, pointing out that questionnaires may increase reliability, and are clearly more transparent.³³ It is also necessary that the complexity of the decision-making process reflects the focus of stepped care on efficiency and throughput. Highly trained personnel able to give expert clinical judgements may simply not be available at lower levels of the stepped care pathway.

The role of the patient in decision making in stepped care

Previous sections have highlighted some ways in which the views of patients can be indirectly encompassed within decision making in stepped care (for example, through the use of individualised patient-defined measures of outcome). However, the focus of the paper has been on the traditional model of decision making which sees the patient's role as secondary to that of the clinician and service. This reflects the fact that stepped care seeks to derive the greatest population benefit from available treatment resources, rather than meeting the needs and preferences of the individual patient.

However, recent government policy has highlighted the increasing importance of 'the patient voice' and 'patient choice' in service delivery, and models of decision making are increasingly based on ideas of 'partnership' and 'shared decision making'.^{34,35} The operation of patient choice within stepped care is potentially problematic, because of the clash between the population and individual

perspectives. It may be that patient choice is focused on choice *within* steps rather than between them, so that patients can choose between different brief interventions at step 2 (such as guided self-help from a graduate worker, computerised CBT, or group psycho-education). However, if patient choice is to operate in decisions *between* steps, then the operation of routine outcome measures within stepped care decision making will need to be conducted in a way that seeks to balance the importance of normative measures with the preferences of patients concerning further treatment. The practicality of achieving such a balance remains a crucial issue for the implementation of stepped care services.

Stepped care and stratified care

Although the NICE guidelines are based on stepped care principles, there is still an assumption within the guidelines that the initial assessment of patients will allow some to be allocated to different steps of the hierarchy immediately, rather than working their way through the steps sequentially (see Figure 1). For example, patients with high levels of initial severity or previous experience of treatment failure at lower steps may immediately access higher steps. Equally, patient preferences could potentially be used to determine which step they initially access. This may be described as 'stratified care'. The optimal mix of 'stepped' and 'stratified' care within psychological therapy services is currently unclear. At present the science of how to predict individual treatment response based on baseline criteria is in its infancy,³⁶ and it is precisely this lack of specificity that makes stepped care a potentially effective alternative.

However, developments in prediction about patient response to treatment based on baseline characteristics suggest the future possibility of both 'stepped' and 'stratified' care, with baseline measures being used initially to place patients at the right point of the stepped care hierarchy, and outcome measures used to control their movement through the system. The need to collect both these forms of data is an additional argument for the adoption of a systems approach.²

Dealing with relapsing and remitting conditions

Most of the preceding discussion has concerned decision making about 'stepping up'. Mental health problems such as depression have a relapsing and remitting course, and may be best considered as chronic disorders.³⁷ Stepped care can be seen as a method of organising and delivering care for chronic conditions, and common disorders like depression

may benefit from organisational changes that have been shown to be effective in other conditions, such as diabetes and hypertension.³⁸ Dealing with a chronic condition means that 'stepping down' is potentially as important as 'stepping up', and any routine outcome measurement technology may need to be sensitive to this issue.

Conclusions

If stepped care is to deliver its promise, then significant consideration needs to be given to the issue of decision making, and the role of routine outcome measures. The current paper has described a number of issues and challenges. Key issues for the future are now briefly discussed.

At a strategic level, policies concerning the implementation of routine outcome measurement need to be clearly and explicitly linked to the stepped care philosophy, so that policy development relating to outcome measures reflects the particular needs of stepped care systems.

Similarly, the development and validation of measures needs to be sensitive to the particular issues raised by stepped care. For example, it is necessary to establish the empirical relationship between measures and the different levels of stepped care, and to make the data accessible in order to allow standardisation across services.

There is also a substantial development agenda relating to the need to link routine outcome measurement technologies to tracking systems, in order to ensure that measures are implemented in a context where the data can be used effectively and efficiently to assist in the monitoring of patient progress and in clinical decision making.

The acceptability of different measurement technologies to patients and professionals needs to be examined. This may be linked to a wider debate about patient and practitioner expectations concerning outcomes in mental health care, and the role of patient choice and preference in treatment delivery within a stepped care system.

Finally, there is a significant research agenda relating to the actual operation of decision making in stepped care, and the role of routine outcome measurement. Initially, routine quantitative data collection within functioning stepped care systems would be useful to examine patient flows through the system in relation to measures of outcome, and to examine variation in patient 'stepping' in relation to those outcomes and clinician characteristics. Given the complexity associated with clinical decision making, qualitative studies of the process of

decision making in stepped care might also be useful. Such studies may usefully consider both the patient and practitioner perspective on decisions on 'stepping up'.

ACKNOWLEDGEMENTS

The National Primary Care Research and Development Centre is funded by the Department of Health, but the views expressed in this paper are those of the authors alone.

REFERENCES

- Roth A and Fonagy P. *What Works for Whom? A critical review of psychotherapy research*. Guildford: London, 1996.
- Mellor-Clark J, Barkham M, Connell J *et al*. Practice based evidence and the need for a standardised evaluation system: informing the design of the CORE system. *European Journal of Psychotherapy, Counselling and Health* 1999;3:357-74.
- Shadish W, Navarro A, Crits-Christoph P *et al*. Evidence that therapy works in clinically representative populations. *Journal of Consulting and Clinical Psychology* 1997;65:355-65.
- Cavanagh K, Shapiro D, Van Den Berg S *et al*. The effectiveness of computerised cognitive behavioural therapy in routine care. *British Journal of Clinical Psychology* (in press).
- Proudfoot J, Ryden C, Everitt B *et al*. Clinical efficacy of computerised cognitive-behavioural therapy for anxiety and depression in primary care: randomised controlled trial. *British Journal of Psychiatry* 2004;185:46-54.
- Bower P and Gilbody S. Stepped care in psychological therapies: access, effectiveness and efficiency. *British Journal of Psychiatry* 2005;186:11-17.
- Lovell K and Richards D. Multiple Access Points and Levels of Entry (MAPLE): ensuring choice, accessibility and equity for CBT services. *Behavioral and Cognitive Psychotherapy* 2000;28:379-91.
- Sobell M and Sobell L. Stepped care as a heuristic approach to the treatment of alcohol problems. *Journal of Consulting and Clinical Psychology* 2000; 68:573-9.
- Scogin F, Hanson A and Welsh D. Self-administered treatment in stepped-care models of depression treatment. *Journal of Clinical Psychology* 2003;59: 341-9.
- National Collaborating Centre for Mental Health. *Management of Depression in Primary and Secondary Care* (National Clinical Practice Guideline 23). London: National Institute for Clinical Excellence, 2005. www.nice.org.uk/page.aspx?o=235213 (accessed 11 May 2006).
- Department of Health. *Outcomes Measures Implementation: best practice guidance*. Leeds: Department of Health, 2005.
- Barkham M, Leach C, Lucock M *et al*. Service profiling and outcomes benchmarking using the CORE-OM: toward practice-based evidence in psychological therapies. *Journal of Consulting and Clinical Psychology* 2001;69:184-96.
- Clifford P, Katsavdakakis K, Lyle J *et al*. How are you? Further development of generic quality of life outcome measure. *Journal of Mental Health* 2002;11: 389-404.
- Beck A, Steer R and Brown G. *Manual for Beck Depression Inventory-II (BDI-II)*. San Antonio, TX: Psychological Corporation, 1996.
- Lowe B, Unutzer J, Callahan C *et al*. Monitoring depression treatment outcomes with the patient health questionnaire-9. *Medical Care* 2004;42: 1194-201.
- Nunnally J. *Psychometric Theory*. New York: McGraw-Hill, 1967.
- Gilbody S, House A and Sheldon T. *Outcomes Measurement in Psychiatry: a critical review of patient based outcomes measurement in research and practice*. CRD Report number 24. York: University of York, 2003.
- Stiles W, Elliott R, Llewelyn S *et al*. Assimilation of problematic experiences by clients in psychotherapy. *Psychotherapy* 1991;27:411-20.
- Howard K, Lueger R, Martinovich Z *et al*. The cost-effectiveness of psychotherapy: dose response and phase models. In: Miller N and Magruder K (eds). *Cost-effectiveness of Psychotherapy: a guide for practitioners, researchers and policy makers*. New York: Oxford University Press, 1999, pp. 143-52.
- Prochaska J, DiClemente C and Norcross J. In search of how people change. *American Psychologist* 1992; 47:1102-4.
- Euroqol Copyright Group. Euroqol - a new facility for the measurement of health related quality of life. *Health Policy* 1990;16:199-208.
- Ryan M and Farrar S. Using conjoint analysis to elicit preferences for health care. *British Medical Journal* 2000;320:1530-3.
- Ruta D, Garratt A, Leng M *et al*. A new approach to the measurement of quality of life: the patient-generated index. *Medical Care* 1994;11:1109-26.
- Newton M. Evaluating the outcome of counselling in primary care using a goal attainment scale. *Counselling Psychology Quarterly* 2002;15:85-9.
- McDowell I and Newell C. *Measuring Health: a guide to rating scales and questionnaires*. Oxford: Oxford University Press, 1996.
- McDowell I and Jenkinson C. Development standards for health measures. *Journal of Health Services Research and Policy* 1996;1:238-46.
- Relman A. Assessment and accountability: the third revolution in medical care. *New England Journal of Medicine* 1988;319:1220-2.
- Kane R. Functional assessment for geriatric patients - or the clinical Swiss Army Knife. *Medical Care* 1987;(suppl):S95-S99.
- Fitzpatrick R, Davey C, Buxton M *et al*. Evaluating patient-based outcome measures for use in clinical trials. *Health Technology Assessment* 1998;2(14).

- 30 Jacobson N and Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology* 1991;59:12–19.
- 31 Lutz W. Efficacy, effectiveness, and expected treatment response in psychotherapy. *Journal of Clinical Psychology* 2003;59:745–50.
- 32 National Screening Committee. *The UK National Screening Committee's Criteria for Appraising the Viability, Effectiveness and Appropriateness of a Screening Programme*. London: HMSO, 2003. www.nsc.nhs.uk/pdfs/criteria.pdf (accessed 11 May 2006).
- 33 Dawes R. You can't systematize human judgement: dyslexia. In: Dowie J and Elstein A (eds). *Professional Judgment: a reader in clinical decision making*. New York: Cambridge University Press, 1988, pp.150–62.
- 34 Department of Health. *Our Health, our Care, our Say: a new direction for community services*. London: The Stationery Office, 2006.
- 35 Elwyn G and Charles C. Shared decision making: the principles and the competences. In: Edwards A and Elwyn G (eds). *Evidence-based Patient Choice*. Oxford: Oxford University Press, 2001, pp. 118–43.
- 36 Lutz W, Leach C, Barkham M *et al*. Predicting change for individual psychotherapy clients on the basis of their nearest neighbors. *Journal of Consulting and Clinical Psychology* 2005;73:904–13.
- 37 Andrews G. Should depression be managed as a chronic disease? *British Medical Journal* 2001;322: 419–21.
- 38 Wagner E, Austin B and Von Korff M. Organizing care for patients with chronic illness. *Milbank Quarterly* 1996;74:511–43.

CONFLICTS OF INTEREST

None.

ADDRESS FOR CORRESPONDENCE

Peter Bower, Senior Research Fellow, National Primary Care Research and Development Centre, Williamson Building, University of Manchester M13 9PL, UK. Tel: +44 (0)1612 757638; fax: +44 (0)1612 757600; email: peter.bower@manchester.ac.uk

Received 27 July 2005

Accepted 2 May 2006