

Getting the measure of outcomes in clinical practice

Glyn Lewis & Helen Killaspy

ARTICLE

SUMMARY

It has been argued that the routine use of patient-reported outcome measures (PROMs) should be encouraged in order to improve the quality of services and even to determine payment. Clinician-rated outcome measures (CROMs), patient-reported experience measures (PREMs) and process measures also should be considered in evaluating healthcare quality. We discuss difficulties that the routine use of outcome measures might pose for psychiatric services. When outcome and experience measures are used to evaluate services they are difficult to interpret because of differences in case mix and regression to the mean. We conclude that PROMs and CROMs could be useful for monitoring the progress of individuals and that clinical audit still has an important role to play in improving the quality of services.

LEARNING OBJECTIVES

- Understand the difference between process measurement and outcome measurement.
- Understand the limitation of using outcome measures to assess and promote quality of services.
- Understand the difficulties in assessing the psychometric properties and validity of outcome measures.

DECLARATION OF INTEREST

None.

Current healthcare policy

Recent years have seen a plethora of policy documents focusing on outcomes in healthcare (Department of Health 2008, 2010, 2011). Outcome frameworks for the National Health Service (NHS), social care and public health have followed (Department of Health 2012a, 2013a,b) and the most recent mental health policy, *No Health Without Mental Health* (Department of Health 2012b), identifies six high-level outcomes related to the aims of greater prevention, well-being, recovery and social inclusion. This shift in focus is away from the preoccupation of previous governments with targets such as reduced waiting times. It also moves away from the emphasis on clinical audit to maintain the quality of healthcare. The new policy is that health services should provide evidence for their effectiveness by measuring outcomes. This appears to be reasonable, but is not quite as simple as it may at first seem.

Processes and outcomes

One problem is that there is often a conflation of the term ‘outcome’ with measures and indicators of process. A typical dictionary definition of an outcome is a result or a visible effect. To give a clinical example relevant to psychiatry, it could be a reduction in symptoms or an improvement in social functioning assessed using a standardised tool. Processes are the inputs that drive or mediate these improvements, for example, the delivery of interventions that have demonstrated effectiveness. Processes can also be measured, such as the proportion of people meeting eligibility criteria who are offered cognitive-behavioural therapy for psychosis as recommended by national guidance on the treatment of schizophrenia

Glyn Lewis is Professor of Psychiatric Epidemiology at University College London. He works as a psychiatrist for Camden and Islington NHS Foundation Trust in a community team specialising in the treatment of depression and anxiety disorders. His research includes evaluating treatments for depression using randomised controlled trials. He is interested in how to apply research findings to clinical practice. **Helen Killaspy** is Professor of Rehabilitation Psychiatry at University College London. She works as an honorary consultant in rehabilitation psychiatry for Camden and Islington NHS Foundation Trust. Her research focuses on the quality and effectiveness of services and interventions for people with complex mental health problems. **Correspondence** Dr Glyn Lewis, Division of Psychiatry, University College London, 67–73 Riding House Street, London W1W 7EJ, UK. Email: glyn.lewis@ucl.ac.uk

Health services are there to improve clinical outcome for people with health problems. Measuring those outcomes would therefore seem a key part of ensuring that the services are doing their job. With the decline of medical paternalism and the supposed empowerment of patients, the patient would seem the correct person to ask about outcomes. So there has been increasing emphasis on the routine use of patient-reported outcome measures or PROMs in health services in order to improve the quality of service provision. This article will discuss some of the main issues concerning the choice and use of PROMs in routine health services, along with other measures concerned with the quality of healthcare (Box 1).

BOX 1 Common abbreviations

PROMs: patient-reported outcome measures
PREMs: patient-reported experience measures
CROMs: clinician-rated outcome measures
QoL: health-related quality of life

(National Collaborating Centre for Mental Health 2009). Of the six 'outcomes' included in *No Health Without Mental Health* (Department of Health 2012b), one is a measure of process (more people will have a positive experience of care), three are measures of outcome (more people will have good mental health; more people with mental health problems will recover; more people with mental health problems will have good physical health) and two can be thought of as measures of both process and outcome (fewer people will experience stigma and discrimination; fewer people will suffer avoidable harm).

Improving healthcare services

The quality of care

Assessment of the quality of care requires an understanding and measurement of the relevant processes and outcomes for any specific service. Therefore assessment of both is to be encouraged. Recent government policy has adopted Lord Darzi's definition of quality as incorporating the effectiveness and safety of treatment and care alongside a positive experience for people using services (Department of Health 2008). This latter point is particularly pertinent in the context of recent quality of care scandals (Department of Health 2013c; Francis 2013). In response to this, the latest Care Quality Commission (CQC) consultation on changes to the way they inspect, regulate and monitor care services suggests a framework for future assessment of the quality of services where four of the five constructs to be evaluated relate to processes (safe, caring, responsive, well led) and only one (effectiveness) is a measure of outcome (Care Quality Commission 2013).

Value-based healthcare

Conversely, there has been recent interest in the concept of 'value-based healthcare' (Porter 2010), which focuses on the relative relationship between the cost of care and clinical outcomes (where value = outcome/cost), with no specific measurement of the inputs (processes). In other words, it matters less what you do, as long as it provides good outcomes for the money spent. Here, although the monetary costs of care are obviously in focus, the value for money of a specific treatment or episode of care also takes into account the non-monetary values of those receiving healthcare. This has synergy with the concepts used in health economics, where the costs of care are weighed against the likelihood of improving quality of life over a certain period of time. These approaches potentially provide a framework for more

patient-focused decision-making in healthcare investment, though the population-based models on which they are based are difficult to extrapolate to individual cases.

Quality and outcome

Recent health policy assumes that better-quality services will produce better clinical outcomes. Since 2004, this assumption has financially incentivised the delivery of primary care for chronic medical conditions in England through the Quality and Outcomes Framework (Department of Health 2004). Although the same approach is now being encouraged for other healthcare systems, including mental health, the relationship between service quality and clinical outcomes has had little empirical evaluation. One large study that investigated the impact of the Quality and Outcomes Framework on diabetes care found no clear association with improved clinical outcomes over the 3 years before and after its introduction (Calvert 2009). However, a recent national survey of mental health rehabilitation services found a positive association between quality of care and patient outcome (Killaspy 2013).

Despite the relatively limited evidence, there are increasing demands across health services to deliver data on service activity and performance, with a number of external bodies (such as Monitor, the NHS Information Centre and the CQC) requiring regular 'outcome' reports. Local commissioning bodies also request data on care quality indicators (CQuINs) to justify continued investment in services. The impending introduction of a tariff-based mental healthcare system[†] (in England at least) will further embed the need for regular data collection to describe in quantitative terms what mental health services deliver and what impact this has on patients.

It is well recognised that offering financial incentives can lead to unintended consequences and 'gaming' in order to improve apparent outcomes. Cross-validation of data to check for inconsistencies can address this to some degree but it is an inevitable consequence of providing incentives.

What outcomes do we need to measure?

There are several categories of outcomes that could be measured. Clinical outcomes would include mortality or depressive and psychotic symptoms. Another major category is often termed health-related quality of life (QoL for short). Quality of life measures are designed to assess important non-symptom outcomes for the patient. In other areas of medicine, this often also includes psychiatric symptoms. For example, there is concern that

[†]National tariffs for healthcare is discussed by Jacobs, pp. 155–164, this issue. Ed.

the treatments for some cancers might extend life expectancy, but in doing so reduce social functioning and emotional well-being (Bowling 2005). Therefore many of the QoL measures used in medicine include symptoms of depression and anxiety. As psychiatrists, we tend to conceive of QoL as social functioning, the ability to maintain relationships, to work and to fulfil responsibilities to family and friends. For completeness, it is also worth noting that some important outcomes can affect people other than the patient, although we are not discussing these in this article. Examples include the burden on carers or the victims of crime.

One further area that is important in health service evaluation is the patient's experience of and satisfaction with healthcare. Although these will usually be reported only by the patient (the measures that assess them are sometimes referred to as patient-reported experience measures, or PREMs), other outcomes can be reported by the patient or clinician (using PROMs and CROMs). However, methodological difficulties apply to all of these measures, as shown in the next section.

Research and clinical outcomes

Reliability

Most current outcome measures were developed primarily for research studies. Accurate measurement lies at the basis of all scientific activity and so there has been an understandable preoccupation in psychiatry with studying the reliability and validity of the measures that we use (Carmines 1979; Streiner 1989). Reliability is best thought of as the repeatability of an assessment. If the same test is used again on the same person within a short enough time period (where no change in their rating would be expected) then the agreement between the two measures is an estimate of the test-retest reliability. Similarly, agreement between two raters assessing the same patient is known as interrater reliability. The more reliable the test, the more closely should the two results agree. However, it could be reliably providing the wrong answer.

One important principle is that the reliability of a test is specific to the population within which it is tested. Reliability is usually calculated as the proportion of variance that can be attributed to the true scores. The variance will depend on the spread of scores in the population being studied, so the reliability will also depend on the characteristics of that population. A test might therefore perform less well in a clinical population than in the published results from other settings.

BOX 2 Types of validity

Criterion validity	The measure agrees with a gold standard
Concurrent validity	The measure agrees with another scale that measures the same construct
Predictive validity	The measure predicts something of importance, such as a good outcome
Face validity	The items in the measure appear to address the construct of interest

Validity

Validity concerns whether the test is measuring what it intends to measure (the construct). Many textbooks list different forms of validity, such as criterion, concurrent, predictive and face validity (Streiner 1989) (Box 2). Criterion validity is the agreement between the measure and a gold standard or error-free measure of the construct. Unfortunately, gold standards are completely absent in psychiatry, as they are in most areas of medicine. Often, clinician-rated assessments have been used as the gold standard, but clinicians still disagree with each other and this will always be a limitation in psychiatric studies. As a result, validity is very difficult to establish for psychiatric measures. Face validity, concurrent validity and predictive validity are also used to justify tests when there is no gold standard.

Validity is often summarised as the sensitivity and specificity of a test in relation to a gold standard. A reliable test may or may not be valid, but an unreliable test cannot be valid. There will always be some uncertainty about the validity of measures in psychiatry, in part because we are not certain about the nature and pathophysiology of the psychiatric disorders we are trying to measure.

The validity of PROMs

It has been argued that the validity of PROMs assessing symptoms of anxiety and depression is likely to be good since these are primarily subjective states and the patient is, by definition, the best person to report on them (Lewis 1989). The validity of a PROM, though, also depends on the insight of the patient. For psychotic phenomena a PROM might be less valid than a clinician measure in which some cross-examination is allowed. Measures of self-reported psychotic symptoms, such as the psychosis screening questionnaires (Bebbington 1995; Horwood 2008), lead to much higher estimates of symptoms than measures that require some degree of

cross-examination (Horwood 2008). This could be because psychotic phenomena might be difficult to explain in a self-reported format and because lack of insight might affect self-reported information. For these reasons, some investigators prefer to use clinician- or researcher-rated scales to assess psychotic symptoms rather than relying on self-reported assessment.

Using research measures in clinical practice

There often appears to be a divide between the measures used in clinical practice and those used in research. However, psychiatric research is meant to inform clinical practice and so ideally the measures used in research should be the same as those used in clinical practice. In this way results from research can easily be applied to clinical situations and *vice versa*.

The Improving Access to Psychological Therapies (IAPT) initiative in the UK is an example where routine outcome measurement has been included as a core element. The IAPT website states that 'Routine outcomes measurement is central to improving service quality – and accountability' (www.iapt.nhs.uk/data). The NHS is expecting IAPT services to increase the proportion of patients who recover after treatment (National IAPT Programme Team 2011). IAPT services use the Patient Health Questionnaire for depression (PHQ-9) (Gilbody 2007) and the Generalised Anxiety Disorder Assessment (GAD-7) (Spitzer 2006) as their main outcome measures, and research studies in the UK are increasingly using the same measures (Richards 2013). This should enable services to compare their outcomes with research results.

Interpreting outcome measures

Case mix

It is well recognised that outcome measurement in a clinical service is difficult to interpret. This applies to any measure of outcome or patient experience. It may be meaningful for an individual, but as a way of evaluating a whole service it is influenced mostly by the characteristics of the patients entering that service. This is often called 'case mix' (Box 3) and there have been efforts to adjust for case mix over the years (Orchard 1994) in order to use routine outcome data to evaluate services. In economically deprived areas, for example, the patients entering IAPT are likely to have more severe conditions, and the outcome for people of lower socioeconomic status who have depression is likely to be worse (Weich 1998; Lorant 2003). This will make it harder for IAPT services in such areas to meet centrally imposed

targets than for services based in more affluent regions. Patients with more severe illness will also have a poorer prognosis. When outcome measures are routinely used it is important to adjust for the different patients seen by different services. If this is not done, services might be discouraged from taking on the more difficult patients and comparisons might be misleading.

Regression to the mean

The other phenomenon that can interfere with routine outcome measurement is regression to the mean (Barnett 2005) (Box 3). This is a statistical phenomenon that can make natural variation in repeated data look like real change. It is particularly likely when someone is selected because they have especially high scores. In effect this happens all the time in clinical practice as patients consult when they are at their worst. As a result, they are likely to appear 'better' merely because the subsequent measurements will usually be closer to the average. This is often interpreted clinically as 'spontaneous recovery' or even as evidence that the treatments have been effective, although of course both of these can happen as well. Spontaneous recovery refers to a real change in the clinical state of the patient that is not a result of any clinical intervention.

Regression to the mean is an inevitable consequence of measurement error, and the outcome measures used in psychiatry are not that reliable. Regression to the mean is sometimes described as 'the physician's friend' and it encourages services to think they are being effective when in reality they may be having little impact.

The usual way of addressing regression to the mean and spontaneous recovery is by having a comparable group not receiving the intervention – in other words, a randomised controlled trial (RCT). However, conducting RCTs is not possible as a routine part of clinical services.

BOX 3 Problems with interpreting outcome data for a service

Case mix – the composition of the patients in the service that affects the outcome. For example, patients with more severe illness will also have a poorer prognosis, so a service that treats people with worse illness will have worse outcome measures.

Regression to the mean – a statistical phenomenon that can make natural variation in repeated data look like real change. Patients will appear 'better' over time merely because the subsequent measurements will usually be closer to the average.

Recovery as an outcome

Over recent years there has been a growing literature concerned with ‘recovery’ from mental health problems, largely with the perspective of people with psychosis (Jacobson 2001). This approach is based on the primacy of the patient experience and the patient perspective. This literature has highlighted areas such as ‘hope’ and ‘empowerment’ as aspects of recovery that are valued by patients but not adequately addressed by current outcome measures. What this indicates is that a narrow focus on psychiatric symptoms may be missing aspects that are valued by patients. In analogy to the use of quality of life measures in parallel with symptom measures, one can envisage a time when the measurement of recovery from the patient perspective will also be an important element of outcome measurement. It would seem appropriate that such measures should be completed by the patient.

Choosing outcome and process measures

As part of the tariff-based approach, the Department of Health is very likely to mandate regular collection and reporting of data from mental health services using a small set of standardised outcome measures. These will include a CROM, a PROM and a PREM that will be used across all mental health services. These measures will need to be universally relevant and will assess broad constructs such as symptoms, well-being and patient satisfaction with care. Beyond these, it may be appropriate to add one or two additional measures that are specific to an individual specialty or service (Trauer 2010). The remainder of this section (summarised in Box 4) describes how to decide on and set up such measures.

Factors to consider

A number of factors must be borne in mind when deciding on the data you plan to gather. First, clarify whether you wish to assess processes, outcomes, experience or aspects of all three. Choose an indicator or outcome that is clinically meaningful (that has good face validity). Consider whether the data you will need to report on this indicator or outcome are already available, or potentially easily attainable. If you plan to use a standardised measure, choose one with good reliability and validity that is appropriate not only to the outcome you wish to assess, but also to the setting you plan to use it in. Consider how user friendly it is for those you will be expecting to complete it in terms of its length, comprehensibility and rating scheme. If it is a staff-rated measure, will staff need training

BOX 4 Choosing an outcome measure or indicator

- Know what you are assessing: processes, outcomes, experience or aspects of all three
- Choose a clinically meaningful outcome or indicator for which data can be easily obtained
- Choose a measure that is valid, reliable, population-appropriate and user friendly
- For activities or processes, know the indicator's numerators and denominators
- Pilot any new measure to iron out the problems

to learn how to complete it? Is it subject to any copyright restrictions and, if so, is there any cost associated with using it? If you are thinking about introducing a new measure, pilot it first to identify any problems with its feasibility. This applies even if the measure has well-established psychometrics, as it will clarify how long it takes to complete, whether those completing it find it easy enough to use (both of which will affect response rates when the measure is rolled out to a bigger population) and whether it really taps into the construct you wish to report on.

Activity and process indicators

If you want to collect activity or process data, be clear about the figures that will constitute your indicator's numerator and denominator. For example, if you want to report on whether your patients are having regular care reviews, you first need to consider which staff have to be at a meeting for it to be classified as a care review. Is attendance at care reviews recorded somewhere in an easily accessible place/record? Are patients always expected to attend? What is the frequency of care reviews you wish to set as your standard? In fact, a number of separate indicators may be needed to assess what seems a fairly straightforward process such as this. Having defined what constitutes a care review meeting and the frequency, one indicator could be the proportion of the team's patients for whom a care review meeting was held within the past 6 months (numerator = number of team's patients for whom a care review meeting was held attended by consultant psychiatrist and care coordinator in the past 6 months; denominator = team's total case-load). Another might be the proportion of these meetings that the patient attended (numerator = total number of team's care review meetings in past 6 months where patient attended; denominator = total number of team's care review meetings held in past 6 months).

MCQ answers

1 e 2 c 3 e 4 b 5 d

Collecting, collating and reporting the data

The data need then need to be collected and collated. In an ideal world, data collation would be carried out by a computerised data management system that has been well designed to identify and extract the specific numerators and denominators you need and collate these into an easy-to-understand report. Unfortunately, the real world tends to disappoint. For the example above, unless there is a specific 'tick box' for staff to code that a patient has had a care review meeting and another to indicate whether the patient attended (and the staff are conscientious about ticking the relevant boxes), the data management system (or person) would have to screen entries in the patients' case notes to identify the numerators and denominators required. This is clearly not feasible on a regular basis. It is therefore wise to carefully consider the resource implications involved in reporting on your chosen indicators and outcomes and to discuss these with the relevant personnel, including the team staff and data managers.

The data reports need to be presented in a format that everyone can understand. Simple charts work well visually, but can be misleading when only proportions and percentages are presented rather than raw data.

A further point to note is that, although there are numerous standardised measures available for assessing a wide range of specific psychiatric symptoms (Royal College of Psychiatrists 2011), many of these have been developed for research studies that assess change at group rather than individual level. If you are able to choose measures that can feed into an individual's clinical review and care planning processes as well as being useful at the team or service level, all the better (Royal College of Psychiatrists 2011). However, you still have to establish a process for collecting and reviewing an individual's data at care review meetings.

Improving services through outcome or process measurement

The ultimate aim of encouraging the use of outcome measurement in health services is to improve quality. Quality applies to all aspects of healthcare, including those that might influence patient experience as well as processes and outcomes. Proponents who argue for the routine use of outcome measures say that this will improve quality. For example, it is thought that the collection of routine mortality data for heart surgery has improved standards in that area (Bridgewater 2013). However, randomisation is the best way to evaluate a healthcare intervention

(Altman 1999) and we are not aware of any examples where routine outcome measurement has been properly evaluated in that way.

There is also an opposing argument that outcome measures are not necessary. We have already discussed the difficulties of interpreting outcome measures for a service. Although outcome measurement is an important part of monitoring the progress of an individual patient, it might be better for the service to ensure that the process of care is well carried out rather than be concerned with potentially misleading aggregate outcomes.

An alternative (and older) approach is to rely on process measures and clinical audit (Benjamin 2008). This approach continues to be used by the Healthcare Quality Improvement Partnership (www.hqip.org.uk), which conducts regular audits such as the National Clinical Audit and Patient Outcomes Programme, mandated in the NHS standard contract. Randomised controlled trials can provide good unbiased evidence concerning the effectiveness of treatments. These results are incorporated in a standard, for example, 'all people with diagnosis A should receive treatment X'. Audit monitors the process of care against that standard, thereby ensuring that all the appropriate patients receive an effective treatment. Many factors affect outcome in addition to medical care. Audit therefore concentrates just on providing the effective treatments. There is high-quality randomised evidence that audit and feedback can be an effective means of improving both processes and outcomes (Ivers 2012).

Conclusions

Outcome measures, whether rated by clinicians or patients, are good at monitoring the progress of individual patients. They are less good at monitoring the quality of services, as patient outcomes will also depend on a variety of factors that cannot be influenced by the health service. Despite these potential limitations, it seems likely that the government, and other funders of the NHS, will increasingly use routine outcome measurement to monitor health service performance. Outcome measures are the new panacea for quality, but it is important to remember the role that clinical audit also plays in improving processes and ensuring that patients receive the appropriate care and treatment.

References

- Altman DG, Bland JM (1999) Statistics notes. Treatment allocation in controlled trials: why randomise? *BMJ*, 318: 1209.
- Barnett AG, Van Der Pols JC, Dobson AJ (2005) Regression to the mean: what it is and how to deal with it. *International Journal of Epidemiology*, 34: 215–20.

- Bebbington PE, Nayani T (1995) The psychosis screening questionnaire. *International Journal of Methods in Psychiatric Research*, **5**: 11–20.
- Benjamin A (2008) Audit: how to do it in practice. *BMJ*, **336**: 1241–5.
- Bowling A (2005) *Measuring Health: A Review of Quality of Life Measurement Scales*. Open University Press.
- Bridgewater B, Hickey GL, Cooper G, et al (2013) Publishing cardiac surgery mortality rates: lessons for other specialties. *BMJ*, **346**: f1139.
- Calvert M, Shankar A, McManus RJ, et al (2009) Effect of the quality and outcomes framework on diabetes care in the United Kingdom: retrospective cohort study. *BMJ*, **338**: b1870.
- Care Quality Commission (2013) *A New Start: Responses to Our Consultation on Changes to the Way CQC Regulates, Inspects and Monitors Care Services*. October 2013. CQC.
- Carmines EG, Zeller RA (1979) *Reliability and Validity*. Sage.
- Department of Health (2004) *Quality and Outcomes Framework*. Department of Health.
- Department of Health (2008) *High Quality Care for All: NHS Next Stage Review*. Department of Health.
- Department of Health (2010) *Equity and Excellence: Liberating the NHS*. Department of Health.
- Department of Health (2011) *Liberating the NHS: Transparency in Outcomes*. Department of Health.
- Department of Health (2012a) *NHS Outcomes Framework 2013–2014*. Department of Health.
- Department of Health (2012b) *No Health Without Mental Health*. Department of Health.
- Department of Health (2013a) *Adult Social Care Outcomes Framework 2013–2014*. Department of Health.
- Department of Health (2013b) *Public Health Outcomes Framework for England 2013–2016*. Department of Health.
- Department of Health (2013c) *Transforming Care: A National Response to Winterbourne Hospital*. Department of Health.
- Francis R (2013) *Report of the Mid Staffordshire NHS Foundation Trust*. TSO (The Stationery Office).
- Gilbody S, Richards D, Brealey S, et al (2007) Screening for depression in medical settings with the Patient Health Questionnaire (PHQ): a diagnostic meta-analysis. *Journal of General Internal Medicine*, **22**: 1596–602.
- Horwood J, Salvi G, Thomas K, et al (2008) IQ and non-clinical psychotic symptoms in 12-year-olds: results from the ALSPAC birth cohort. *British Journal of Psychiatry*, **193**: 185–91.
- Ivers N, Jamtvedt G, Flottorp S, et al (2012) Audit and feedback: effects on professional practice and healthcare outcomes. *Cochrane Database of Systematic Reviews*, issue 6, CD000259.
- Jacobson N, Greenley D (2001) What is recovery? A conceptual model and explication. *Psychiatric Services*, **52**: 482–5.
- Killaspy H, Marston L, Omar RZ, et al (2013) Service quality and clinical outcomes: an example from mental health rehabilitation services in England. *British Journal of Psychiatry*, **202**: 28–34.
- Lewis G, Williams P (1989) Clinical judgement and the standardized interview in psychiatry. *Psychological Medicine*, **19**: 971–9.
- Lorant V, Deliege D, Eaton W, et al (2003) Socioeconomic inequalities in depression: a meta-analysis. *American Journal of Epidemiology*, **157**: 98–112.
- National Collaborating Centre for Mental Health (2009) *Schizophrenia: Core Interventions in the Treatment and Management of Schizophrenia in Adults in Primary and Secondary Care* (NICE Clinical Guideline 82). National Institute for Health and Clinical Excellence.
- National IAPT Programme Team (2011) *The IAPT Data Handbook: Guidance on Recording and Monitoring Outcomes to Support Local Evidence-Based Practice* (Version 2.0.1). TSO (The Stationery Office).
- Orchard C (1994) Comparing healthcare outcomes. *BMJ*, **308**: 1493–6.
- Porter ME (2010) What is value in health care? *New England Journal of Medicine*, **363**: 2477–81.
- Richards D, Hill J, Gask L, et al (2013) CADET: clinical effectiveness of collaborative care for depression in primary care. A cluster randomised controlled trial. *BMJ*, **347**: f4913.
- Royal College of Psychiatrists (2011) *Outcome Measures Recommended for Use in Adult Psychiatry* (Occasional Paper OP78). Royal College of Psychiatrists.
- Spitzer RL, Kroenke K, Williams JB, et al (2006) A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of Internal Medicine*, **166**: 1092–7.
- Streiner DL, Norman GR (1989) *Health Measurement Scales*. Oxford Medical Publications.
- Trauer T (ed) (2010) *Outcome Measurement in Mental Health*. Cambridge University Press.
- Weich S, Lewis G (1998) Poverty, unemployment and the common mental disorders: population based cohort study. *BMJ*, **317**: 115–9.

MCQs

Select the single best option for each question stem

1 The following are measures of outcome:

- a the proportion of people with depression offered cognitive-behavioural therapy
- b patient satisfaction with care
- c time from referral to assessment
- d length of admission
- e gaining employment.

2 The following outcomes are relevant to psychiatric services:

- a quality of services
- b costs of care
- c side-effects of medication
- d number of missed appointments
- e patient satisfaction with care.

3 Which of the following is not a type of validity?

- a The items of the scale appeared to measure the construct
- b The measure agreed with a scale previously used to measure the same construct
- c The measure was associated with outcome
- d There was agreement with a better measure
- e Two raters gave the same answer.

4 The following have been demonstrated to improve quality:

- a outcome measurement
- b clinical audit
- c care quality indicators (CQINs)
- d financial incentives linked to outcomes
- e quality outcome frameworks.

5 The following do not need to be considered when using a PROM:

- a time to complete
- b agreement with construct
- c test-retest reliability
- d interrater reliability
- e usability.