

# What outcomes to measure in routine mental health services, and how to assess them: a systematic review

Mike Slade

---

**Objective:** Routine outcome assessment in adult mental health services involves the on-going assessment of patient-level outcomes. Use of outcomes to inform treatment is widely recommended, but seldom implemented. The goals of this review were (i) to identify principles that have been proposed for implementing routine outcome assessment, (ii) to identify the full range of outcome domains that have been proposed for assessment, and (iii) to synthesize proposals for specific outcome domains into emergent categories.

**Method:** A systematic review of published and unpublished research was undertaken, using electronic databases, research registers, conference proceedings, expert informants and the World Wide Web. For goal (i) studies were included that proposed principles for implementing routine outcome assessment. For goal (ii) studies were included that identified at least two patient-level outcome domains for patients using adult mental health services and made some reference to a broader literature base.

**Results:** Six thousand four hundred publications matched initial search criteria. Seven distinct sets of principles for choosing patient-level outcomes were located, which showed a fair degree of consensus. Sixteen outcome domain proposals were identified, which were synthesized into seven emergent categories: wellbeing, cognition/emotion, behaviour, physical health, interpersonal, society and services.

**Conclusions:** The findings from this review were used to develop a four-step method for adult mental health services wishing to implement routine outcome assessment.

**Key words:** mental health, outcome assessment.

**Australian and New Zealand Journal of Psychiatry 2002; 36:743–753**

A standard definition of ‘outcome’ in mental health care is ‘the effect on a patient’s health status attributable to an intervention by a health professional or health service’ [1, p.3]. Despite being widely used, this definition has been challenged for several reasons – outcome could result from self- rather than professional help, the link between intervention and outcome is not straightforward, outcomes are not always positive, outcomes may be influenced by the absence rather than presence of

an intervention, and outcome may differ from different perspectives [2]. There is as yet no consensus about an agreed definition of outcome for individual patients. Further complexity arises when evaluating outcome in mental health services. Three levels of mental health service can be identified: specific treatments, combinations of treatments (such as a community mental health centre) and population-wide treatments (all programmes for a defined population, such as a managed care organization) [3]. The outcome data needed to inform each level are very different.

Perhaps because of this complexity, the systematic measurement of outcome in mental health services has traditionally been the preserve of researchers. In general, most efforts to assess outcome take place in ‘research

---

Mike Slade, Medical Research Council Clinician Scientist Fellow

Health Services Research Department (PO29), Institute of Psychiatry, De Crespigny Park, Denmark Hill, London, SE5 8AF, United Kingdom.  
Email: m.slade@iop.kcl.ac.uk

Received 19 March 2002; 2nd revision 23 July 2002; accepted 30 July 2002.

contexts where specifically funded and trained external raters parachute into routine clinical settings in order to guarantee the validity and reliability of study measures' [4]. However, in the last decade a new international emphasis on the ongoing measurement of outcome in routine mental health services – routine outcome assessment – has begun to emerge [5–8].

Internationally, a range of approaches have been taken to assess outcome routinely. In the US, the focus on outcomes as the measure of success for mental health services has been driven by cost-containment. Difficulties in implementing Diagnostic Related Groups [9] and the growth in proportion of healthcare costs spent on mental health services from 4% in the early 1980s to 25% in the early 1990s [8] led to an emphasis on the use of outcomes. This emphasis fits with other quality enhancement initiatives, the increasing availability of an information technology infrastructure, the importing of 'learning organization' values from business into health care, and the pressure from consumer organizations for issues of choice, quality and value to be considered [8].

In the UK, a similar impetus has come from the drive towards evidence-based medicine. This is exemplified by the development of the National Institute for Clinical Excellence to provide cost-effectiveness information about health interventions, and the development of national standards for mental health care [10]. Other influences include an emphasis on clinical governance and practice guidelines, a political emphasis on quality and on patient experience, the development of high-profile (if not widely used) outcome measures for routine clinical use [11] and a societal shift towards consumerism, with concomitantly increased expectations about mental health services.

Most approaches to collecting data within routine mental health services have therefore been intended to inform programmes and systems. However, there is an increasing recognition of the importance of treatment-level outcomes, which can inform the future care provided to individual patients [2]. Australia has the most coherently developed approach to treatment-level routine outcome assessment. A systematic review of patient outcomes was undertaken as part of the first national mental health strategy, resulting in proposals for specific assessments to use routinely [1]. These assessments were then independently field-tested, to evaluate their utility [12]. The resulting recommendations have informed Australian practice in routine outcome assessment.

Mental health services implementing routine outcome assessment will want to base their efforts on principles developed through the experience of other services. One specific decision will be the outcome domains (conceptually distinct components of outcome, such as quality of

life, symptomatology or satisfaction with care) to assess. This article is intended to assist adult mental health services in implementing routine outcome assessment, by using a systematic review of the available literature to inform a proposed method for implementation. The review goals were (i) to identify the principles that have been proposed for implementing routine outcome assessment, (ii) to identify the full range of outcome domains that have been proposed for assessment, and (iii) to synthesize proposals for specific outcome domains into emergent categories.

## Method

### Study selection

The main sources for published information were the electronic databases shown in Table 1. However, electronic searching will not identify all relevant research, partly through missing relevant indexed journal papers, and partly through not accessing technical reports, discussion papers and other forms of 'grey literature' [13]. Efforts were made to access these studies using three methods. First, researchers active in the field were consulted, and findings presented at the four European Network for Mental Health Service Evaluation conferences were reviewed. Second, the World Wide Web was searched using Copernic 2000, an internet search engine which collates the findings from other search engines. Third, the UK National Research Register and the Research Findings Electronic Register were searched. No language restrictions were employed in any search, and non-English articles were included where an abstract in English was available. Pre-publication and 'in press' manuscripts were included and the literature review was completed by the author.

It was not possible to identify a search strategy that differentiated between publications relating to mental health research and to routine mental health services. Both were therefore included. Some of the outcome domains identified were described as models of 'health status', 'wellbeing' or 'quality of life', but no search strategy was identified that allowed searching on any of these key words with sufficient specificity. Similarly, no satisfactory synonym for routine (as in 'routine mental health services') could be found, so this aspect was incorporated where possible when reviewing abstracts (although often the distinction between research and routine clinical uses was not made). To maximize sensitivity, the search strategy was deliberately over-inclusive.

Several electronic databases were searched and the *MEDLINE* search engine allowed the most comprehensive search strategy. For the *MEDLINE* search, all studies relating to mental health or psychiatry (identified from title, abstract or medical subject heading (MeSH) heading) with either the word 'outcome' in their title or abstract or a MeSH heading of 'Outcome and process assessment (mental health)' were identified. Treatment trials and animal-only studies were excluded. The search was then adapted for other electronic databases, which were less sophisticated. For instance, the IBSS engine only allowed one search term, so 'outcome' was used. Duplicates of all identified articles were removed using Reference Manager Professional Edition, Version 9.5 (ISI ResearchSoft, Berkeley, California).

Table 1. Electronic databases used for the literature review, and number of publications from each database matching initial search criteria

Name of database	Brief description	Search engine	Web site	Search dates	Number found
<b>Primary sources</b>					
Medline	Human medicine and biomedical research	Ovid version 7.8 accessed via Biomed	http://www.biomed.niss.ac.uk	1993 – June 2000	1973
Cumulative Index of Nursing and Allied Health Literature (CINAHL)	Nursing literature	Ovid version 7.8 accessed via Biomed	http://www.biomed.niss.ac.uk	1993 – April 2000	255
PsycINFO	Psychology research	Ovid version 7.8 accessed via PsycINFO	psycinfo.umds.ac.uk	1993 – April 2000	1941
International Bibliography of the Social Sciences (IBSS)	Social Sciences database	Ovid version 7.8 accessed via IBIDS	http://www.bids.ac.uk	1951 – June 2000	539
National Research Register 2000, Issue 3 (completed projects only)	UK National Health Service	Update software	http://www.update-software.com	Unspecified	615
Research Findings electronic Register (ReFeR)	UK National Health Service	Unspecified	tap.ccta.gov.uk/doh/refr_web.nsf/basicsearch	1995 – August 2000	147
World Wide Web	'grey' literature	Copernic 2000	http://www.copernic.com	August 2000	450
<b>Secondary sources</b>					
The Cochrane Methodology Register	Articles on the science of research synthesis	Update software	http://www.update-software.com	Unspecified	161
NHS Economic Evaluation Database	Economic evaluation of health interventions	NHS CRD	http://www.york.ac.uk/Institute/crd	1994 – June 2000	248
Health Technology Assessment database	Information on health technology assessments	NHS CRD	http://www.york.ac.uk/Institute/crd	1993 – January 2000	28

Since a high-quality review was published in 1994 [1], the electronic search was restricted to publications in or since 1993. The review was undertaken between September and October 2000.

## Data extraction

The inclusion criterion for principles was that (potentially) measurable principles were proposed for implementing routine outcome assessment. The main reasons for exclusion were that proposals were too narrow (e.g. relating to minimizing staff resistance to outcome measurement, relating to measuring outcome of psychotherapy, relating to outcome data solely for service funders, and (most commonly) relating to desirable psychometric properties of assessments) or too general (e.g. relating to measuring outcome in all medical settings).

The inclusion criteria for outcome domains was that the proposal identified a range of (i.e. more than one) treatment-level outcome domains for patients using adult mental health service, and made some reference to a broader literature base beyond personal experience or expertise. Proposals relating to other areas of medicine were only included if the proposal was sufficiently generic to have relevance to mental health services, as rated by the reviewer. Exclusion criteria included a focus on choosing outcome measures (rather than domains), proposals relating to patient groups other than adults of working age or outcome domains for programmes (e.g. service uptake) or systems (e.g. interagency working) and reports of individual treatment trials. Emergent categories of outcome domains were identified, comprising the smallest conceptually distinct set of categories which could encompass all proposed outcome domains.

The titles of all publications identified in the initial electronic search were read, to identify those with possible relevance. The abstracts from these identified publications were then reviewed, and where they appeared to meet inclusion criteria the full publication was obtained and read, following which a decision was made as to its inclusion. The reference lists from all obtained articles were also hand-searched for relevant earlier publications. Where more than one publication referred to the same piece of work, only the earliest was included, even where the apparently later one indicated that it was the first publication (e.g. [14,15]). Where the date of 'publication' for grey literature was not clear, the date of the latest citation was used (e.g. [16]).

## Results

The databases searched are shown in Table 1, with the number of publications matching search criteria shown in the final column.

As well as the 6357 publications identified electronically, approximately 50 were identified from nonelectronic sources. Approximately 150 full papers were obtained. Seven papers presenting distinct sets of principles for implementing routine outcome assessment were located [5,12,14,17–20], which together identified 18 principles. Table 2 orders these principles by degree of consensus (with one principle identified in 6 of the 7 studies, two identified in 5 studies, etc.).

The authors of six of the studies worked in North American institutions and one [12] in Australia. Five studies were conducted under the auspices of national bodies – the National Institute for Mental Health [17,18] and the National Alliance for the Mentally Ill [14], university departments [20] and Government departments [12], and two by individuals [4,19]. Studies [14] and [17] are based on the findings of task

Table 2. Principles identified in 7 studies for implementing routine outcome assessment in mental health services

Principle	17	5	18	Reference	14	20	12
Standardized measures should be used	✓				✓		✓
Relevance to informing practice should be emphasized	✓	✓	✓	✓	✓	✓	✓
Multiple perspectives should be used	✓		✓	✓	✓	✓	✓
Standardized methods should be used	✓		✓		✓		✓
Data collection should be cheap and simple	✓		✓		✓		✓
Measures should be relevant to the patient group	✓	✓	✓	✓	✓	✓	✓
Treatment received should be characterized	✓	✓		✓			✓
Feedback should be quick, easy and meaningful	✓	✓					✓
Aggregated data should be comparable with benchmarks	✓	✓					✓
Meaning of measures should be comprehensible	✓			✓			✓
Data should be collected longitudinally			✓		✓	✓	
Casemix (e.g. diagnosis) should be assessed				✓			
Measures should show means/processes of change	✓				✓		✓
Measures should fit with psychopathology theories	✓				✓		✓
Outcomes chosen should be multidimensional							
Costs should be included							
Data on treatment leavers should be collected†							
Individual utility differences should be considered			✓				

†The outcome domains chosen are specific to the individual, since different patients may attribute different degrees of importance to particular outcome domains.

forces, studies [18] and [20] on expert panels, study [12] on a literature review (strongly based on [17] and [21]), and studies [9] and [19] on personal expertise.

Sixteen specific and distinct proposals for outcome domains to measure were identified. These are shown chronologically in Figure 1, organized using seven categories that emerged from the review process. The original terms for proposed outcome domains are shown, with vertical lines indicating the span of the outcome domain across the emergent categories.

All literature reviews used as justification for the proposed domains were selective – no systematic review was identified. Study [34] stood out as being underpinned by a sound (though not systematic) literature review. It included aspects such as sentinel events (undesirable outcomes of a magnitude to always warrant a detailed investigation of a clinician's actions) and technical proficiency (of the clinician), which featured prominently in the general medical but not the mental health literature. (Technical proficiency is regarded as a process rather than an outcome domain in most mental health-focused literature).

The affiliations of the authors of studies [13,19–23,25–28,31] were North American, of study [24] was Italian, of studies [29,30] were English, and of studies [32,33] were English and Italian. Studies [19,20] described domains of health-related quality of life, studies [21,28,30] described domains of health status, and the remainder described treatment outcome domains. Study [19] related to a cardiovascular patient group, studies [20,21,31] to general medicine, [28] to psychotherapy, [30] to clinical trials, and the remainder to patients of mental health services.

Most proposals defined the meaning of the outcome domain. For example, Ware [25] defined 'mental health' as both behavioural dysfunction and the frequency and intensity of symptoms of psychological distress and feelings of psychological wellbeing; 'physical health' as limitations in performance, ability to perform daily self-care, or undertake a range of physical activities; 'social functioning' as both social contacts and social ties or resources; and 'role functioning' as performance of role activities such as employment, school and housework. Hargreaves and Shumway [24] stated that humanistic goals are to maximize the patient's and the family member's sense of wellbeing and personal fulfillment; clinical goals are to improve or cure an illness or disorder, reducing or eliminating its signs and symptoms; rehabilitative goals are to restore or improve social and vocational functioning; and public safety goals are to prevent injury whether from assaultive or self-destructive behaviours that arise out of illness, or from 'destructive' (i.e. iatrogenic) side-effects of the services themselves. Finally, Campbell [16] described wellbeing as linked to the protection of a person's basic human freedoms, safety and privacy; personhood as a recognition of common humanity and a tolerance for individual differences; self-help as including both self-help groups and provision of specific services by consumers; recovery as the maximization of a consumer's life and the minimization of their illness with appropriate, relevant and continuous flexible service and supports collaboratively developed and chosen; empowerment as involving the help receiver having direct control over the help and there being reciprocity between help givers and receivers; iatrogenic effects and negative outcomes being undesired consequences from or side-effects of receiving certain public mental health services or treatments; and satisfaction and dissatisfaction (both being important to measure) relating to the consumer's view of services received and the results of the treatment.

Study	Wellbeing	Cognition/ emotion	Behaviour	Physical health	Interpersonal	Society	Services
Wenger [22]	Perceptions of health status & wellbeing	Intellectual Emotional Symptoms	Daily routine	Symptoms – other illnesses	Social	Economic	
Bergner [23]	Health perceptions General life satisfaction	Symptoms Emotional status Cognition	Functional status	Sleep and rest Energy and vitality	Role activities Social functioning		
Hargreaves [24]	Humanistic	Clinical			Rehabilitative	Public safety	
Ware [25]	Mental health			Physical health	Social functioning Role functioning		
Rosenblatt [26]	Life satisfaction and fulfillment	Clinical			Functional	Welfare and safety	
Ruggeri [27]	Quality of life Needs for care	Psychopathology			Social functioning and support	Burden of relatives	Satisfaction
Sederer [28]		Symptom			Functional		Satisfaction with treatment
Cook [29]	Quality of life				Vocational Educational Residential		Hospitalization Consumer satisfaction
McGlynn [30]	Quality of life	Clinical			Functional		Adverse events Satisfaction with medical care
Schloster [31]	Life satisfaction Life direction	Emotional Mental		Physical		Social	
Campbell [16]	Wellbeing Personhood	Recovery					Self-help Empowerment Iatrogenic effects and negative outcomes Satisfaction and Dissatisfaction
Clifford [32]		Psychological	Activities of daily living	Physical wellbeing	Interpersonal relationships	Social circumstances	Response to care
Fitzpatrick [33]	Psychological wellbeing Personal constructs			Physical function	Social well-being Role activities		Satisfaction with care
Jennings [34]	Health status/ health-related quality of life Patient knowledge	Diagnosis Psychological function Symptom management	Behavioural Activities of daily living	Comfort/dis-comfort Physical function Mobility Disability	Social function		Patient satisfaction Appropriateness of treatment Sentinel events Technical proficiency
Thornicroft [35]	Quality of life	Disabilities Needs				Impact of caring	Satisfaction with services
Thornicroft, 2000 [36]	Quality of life	Symptom severity	Needs		Social disabilities	Carer burden	Quality of care Satisfaction
			Global functioning				

Figure 1. Outcome domain proposals from 16 studies for use in mental health services, grouped into 7 emergent categories.

## Discussion

This study reviewed principles and outcome domains for routine outcome assessment in adult mental health services. Some agreement exists about the key principles for routine outcome assessment. There is consensus on the need for scientific rigour in routinely collected data to be achieved through the research strategies of standardized and relevant measures and methods, assessment from multiple perspectives, and that putting outcome information into context requires information on treatment received. There is also consensus about the need to ensure that the collection of data requires minimal effort and provides relevant information. There is less consensus about the specifics, since routine outcome assessment can be done for different purposes. Where the goal is to provide outcomes to inform the comprehensive provision of mental health care for a defined population, there is a need for benchmarking and cost information. Where the goal is to inform the planning, development and evaluation of a specific service (such as a community mental health team), the focus is more on casemix and treatment leavers. Finally, at the treatment level it is important to collect data longitudinally, and to consider the preferences (utility) of individual patients.

Two themes emerged from the review of outcome domains. First, early proposals did not include assessment of the experience of receiving services, which only came to prominence in the mid-1990s. Second, two distinct perspectives can be identified. Publications within the medical literature used a more psychiatric language, emphasized the staff perspective, and had a focus on the amelioration of disability. The only publication that was located from outside the medical literature [16] used a more phenomenological language emphasizing the patient's experience of care, and focused on increasing the patient's wellbeing and avoiding harm from and dependence on mental health services. The search strategy was systematic within the psychiatric literature but not within the broader social science or user movement literature. Therefore other well-developed proposals for outcome domains probably exist that were not identified in this review.

Synthesizing previous work led to the emergence of seven categories of outcome domains: wellbeing, cognition/emotion, behaviour, physical health, interpersonal, society and services. Wellbeing relates to the patient's sense of subjective wellbeing in their life (not about services), and by definition can only be assessed by the patient. This may involve considering individual life domains, or be a single global outcome. The next three categories relate to the patient as an individual – their cognition/emotion, their behaviour and their physical

health. For all three of these the clinician and the patient may have their own assessment, and their assessments may differ. For example, in the cognition/emotion category the self-reported level of depression may not accord with the clinician's assessment of 'objective' signs of depression. The interpersonal category refers to aspects of the patient in relationship to others, both in individual social interactions and in performance of social roles. The society category describes aspects of a patient's mental health problems that may impinge on wider society, both at the individual level of the burden of care, and the macro-level of costs (e.g. welfare benefits, reduced public safety). Finally, the services category emerged as a distinct outcome domain to consider, including both positive and negative aspects of receiving mental health care.

### Internal validity

The internal validity of this review can be considered in terms of the criteria outlined in the Quality of Reporting of Meta-analyses guidelines [37]. The review could be improved in a number of ways. The inclusion criteria could not be formally specified beyond the conceptual level, since relevant principles and outcome domain proposals appeared in different contexts. This difficulty in constructing a precise search strategy for a non-quantitative search has been acknowledged by systematic review specialists: 'when searching for qualitative research for the purpose of systematic reviews, it is often not practicable to construct strategies to capture the many ways in which such research may be described [38].' Similarly, characterizing the identified studies was problematic. The intended type of study was clearly described – related to routine outcome assessment in adult mental health services. In practice, included publications often were not clear about their remit, and hence were difficult to characterize. No flow profile was maintained, to show the points of and reasons for attrition. Approximately 6400 publications matched initial search criteria, but no record was then kept of numbers excluded at each stage (e.g. removal of duplicates, initial screening, removal following retrieval). The rationale for this was that such information would not be relevant to a non-quantitative review, although as the review progressed it became clear that some synthesis was possible, and that the attrition rate would in any case have been of interest for identifying the key points of exclusion. These methodological deficits reduce the replicability of the review.

The quality of the research was not assessed. Some commentators suggest that no ranking of qualitative methods is possible and each article needs to be considered

on its own merits (e.g. [39,40]). Others have developed approaches to judging the quality of quantitative research (e.g. [41,42]), although these relate in the main to methodological standards rather than the evaluation of conceptual work. The requirement for some form of literature review for outcome domains was a minimum quality assurance approach, but this could be strengthened, for example by duplicate reviewing or the involvement of more than one reviewer.

### Generalizability

Is the review externally valid? Eleven (73%) of the sets of principles and six (86%) of the outcome domain proposals come from North American authors, reflecting that purchaser-driven pressures have stimulated more activity in routine outcome assessment there than anywhere else. Report cards and other means of characterizing aspects of the effectiveness of a service are now routinely used in North America, which has had a substantial impact on the types of care available, and the length of time for which it can be offered. The findings of this review are therefore of most relevance to North American settings. However, the Australian, UK and Italian studies were compatible with the North American work, and hence the findings are likely to have external validity more broadly.

Clearly conceptualization of health and illness differs across cultures, with some cultures giving prominence to domains (such as culture or spirituality) that were not located in this study. Some population subgroups are insufficiently researched to allow the identification of valid understandings of what constitutes normal and abnormal within the culture, such as the paucity of knowledge concerning mental health for Aboriginal youth [43]. Therefore, the values and aspirations of individual subgroups (e.g. Maori [44]) have probably not been captured in the outcome domains identified in this study.

Overall, the external validity of this review for different groups is difficult to establish without a broad range of empirical data. However, there is no reason to think that using the seven emergent categories of outcome domains as a starting point for implementing routine outcome assessment would be unwise.

### Clinical implications

What are the clinical implications of this review? It provides a starting point for mental health services that are considering the use of patient-level routine outcome assessment. Such services might work towards decisions about methods and measures in four steps.

#### *Step 1 – realism*

Consider the principles for routine outcome assessment identified in Table 2. They are useful to consider for two reasons. First, they give an indication of the range of issues that will need to be addressed. The psychometric properties of any outcome measures should be established – locally developed assessments are unlikely to be appropriate – and they will need to be administered in a standard way. As a minimum, both patients and staff will be used as informants, and all collected data must either inform the treatment of individual patients or the development and evaluation of services. To make sense of the data, some characterization of the treatment received will be necessary, and mechanisms for analysing the data and producing feedback will be required.

Second, consideration of these issues gives an indication of the resources required. Resources include leadership, expertise, support staff, information technology (e.g. access to computers, easy-to-use software) and clinical time. If these resources are not available then routine outcome assessment should not be undertaken. Starting to use outcome measures without the requisite resources typically results in the haphazard collection of low-quality data that is not analysed or used, until the endeavour is either abandoned or covertly sabotaged (e.g. by 100% non-response rate). Such effort is of no benefit to patients, and creates an unhelpful belief for clinicians that outcome assessment is a clinical burden rather than providing useful data to inform treatment planning.

#### *Clinical time – a key resource*

A particularly valuable resource is the clinician's time – is it best spent assessing outcome? Completing and analysing simple outcome measures in routine clinical practice can add 10% to the time spent by the clinician per patient [45]. At present, outcome measures are not used routinely within mental health care [46], suggesting that clinicians remain unconvinced that this extra time (and the consequent reduction in number of patients they can see) is a price worth paying. Indeed, since seeing patients is often viewed as valuable clinical activity and filling in forms is not seen as 'work', there are in fact active disincentives to staff completing outcome measures [47]. To make routine outcome assessment more realistic, developments may be necessary in the culture of clinical practice, the research base, and the implementation strategies.

A change in the culture of clinical practice may be needed, in which structures (e.g. number of beds) and

processes (e.g. number of clinical contacts) are de-emphasized, and outcomes become the central influence on decision-making about continuing, changing or ending care. This shift would of course have a profound impact on the way mental health services are formed, operated and evaluated. For example, demand for mental health services has increased by a factor of 4.5 from 1971 to 1997 [48], and presumably will continue to rise. If mental health services operate with an increasing and effectively unlimited caseload size, then it is unrealistic to expect any intervention requiring more time to be spent per patient to be implemented, whatever its merits. Alternative models, (e.g. limiting caseload sizes to ensure a defined level of quality of care is possible), might need to be in place before routine outcome assessment could be realistically considered. The resulting population-level health gain from these and other models of service provision could be investigated, and this type of mental health services research is urgently needed. Other examples of approaches to changing the working culture include the introduction of payment incentives for clinicians who collect and use outcome information [47], and the monitoring of outcome data during clinical training [49,50].

Research studies are needed that quantify the effectiveness and cost-effectiveness of implementing routine outcome assessment. This will allow informed discussion about the relative merits of different styles of clinical practice, such as providing care explicitly targeted at improving outcomes versus providing care audited for its conformance to good practice. Several such studies are currently underway across Europe [e.g. 51,52] that will provide this evidence.

Finally, creative approaches will need to be developed that minimize the time spent by clinical staff in collecting and analysing outcome data. For instance, some uses may only require patient-rated data. In these cases, the use of electronic questionnaires for data collection could be considered. For analysis, the use of automatic data entry and analysis or the employment and training of non-clinical staff for this purpose could be evaluated. One approach would be the provision of a computer for a patient to use before his/her clinical meeting, which then analyses and prints out the resulting outcome data (e.g. by using previously entered information to chart progress over time) for the clinician to review at the meeting. Such an approach would raise further questions, such as whether comparison between clinicians or between clinics is appropriate, and how to maintain data quality where the assessment is not undertaken by a clinician. The costs would include the setting up and maintenance of the computer and the clinician's time spent reading the results, and the potential benefits

would be reduced assessment time and more informing of care planning [6,53–55]. Overall, minimizing the burden and maximizing the potential benefits from using outcomes in clinical care will make routine outcome assessment more realistic.

### *Step 2 – outcome domains*

Identify what outcome domains are appropriate to monitor in the service. This decision will be informed by considering the seven emergent categories identified in Figure 1. Differing understandings of mental health problems lead to disagreement about the outcome domains to consider. These understandings can be conceptualized as lying on a spectrum, from patient-defined to professionally defined [56]. At the extreme of the patient-defined end lies an understanding that emphasizes the importance and uniqueness of individual experience, and accords no value to comparison of one person with another. At the extreme of the professionally defined end lies an understanding that emphasizes the importance of using scientific knowledge to understand abnormal mental experiences, and accords no value to the meaning attached to these experiences by the patient. Most clinical practice, of course, takes place within these extremes, but the point on the continuum will influence the outcome domains selected for routine assessment. A service operating towards the patient-defined end will be more interested in outcome domains related to individual phenomenological change, and how the health care service is experienced. A service operating towards the professionally defined end will be more interested in outcome domains related to symptoms and functioning, and ensuring that the interventions given are in accord with research evidence. One approach to reconciling the conflicting interests of staff and patients is to assess all seven domains identified in this review, which runs the risk of being impractical. Another approach is to identify one domain for assessment, which runs the risk of being insufficiently meaningful. Most approaches to implementing routine outcome assessment involve assessing between two and four domains.

### *Step 3 – technicalities*

Consider how these outcome domains will be measured within the service. This will require consideration of several questions [33,57]. What constitutes a good outcome for a patient who is not expected to recover? Is the goal to show that the treatment caused improvement, or just to show that improvement occurred (without reference to treatment)? Whose outcome is being considered? Cost containment, adherence to clinical protocols,

reduced symptomatology and reduced visibility of the mentally ill are all outcomes from different perspectives. Is the focus just on outcome for the patient, or also for their relatives or carers in their own right? Is the focus on individual change, or aggregating data to investigate changes in groups? The data required for individual or group-level analysis may be very different. Are direct measures (e.g. from the patient or carer) or indirect measures (e.g. from staff or service usage) to be favoured? Is equal weighting given to externally observable measures and private, non-observable experiences of the patient? Are global (single-score) or multiple item measures preferred? Are generic measures (applicable to broad groups) or specific measures (for highly characterized conditions) preferred? Are individualized (tailored to the individual) or standardized measures (which can be compared to group norms) preferred? Should assessment be undertaken at 'important' times during the patient's pathway through care, or at predetermined fixed time periods? How should 'conflicting' changes, such as increased symptoms accompanied by increased quality of life, be interpreted? The answers to these questions will reflect underlying principles of the service.

#### *Step 4 – outcome measures*

Identify the outcome measure(s) that most meet the requirements that have been identified in Steps 1–3. This should be the final step. Several collations of outcome measures exist (e.g. [36,58–61]), although these tend to evaluate their suitability for use in research studies. Measures for routine use also need to be 'feasible' [1], for example by being brief, simple, acceptable, available, relevant and valuable [47]. It is possible to evaluate the feasibility of outcome measures [62].

By way of example, an evaluation of routine outcome assessment in adult mental health services is currently taking place in London. Patients with any diagnosis are included providing they are aged 18–65 and have been in contact with the mental health service for at least 3 months. The optimal frequency and format of assessment and feedback for different patient populations remains an open question, but in this study both staff and patients are asked to complete outcome measures every month, and treatment-level graphical feedback is provided to both people every three months. The goal of the intervention is to beneficially influence the process and content of care [2]. Six criteria were used for choosing outcome measures: (i) the measure either assesses a desired outcome (needs, quality of life) or process measure (therapeutic alliance), allows explicit comparison between staff and patient views, or is a severity measure leading to a desirable focus on outcome; (ii) the

measure has peer-reviewed published evidence of acceptable psychometric properties; (iii) the measure is designed specifically for a mental health population; (iv) the measure is brief to administer (arbitrarily chosen as an administration time of less than 15 min); (v) there is no charge to use the measure; and (vi) there is no training required to use the measure. Using these criteria, the measures chosen for staff and patient were the Helping Alliance Scale (HAS) to assess therapeutic alliance [63], the Camberwell Assessment of Need Short Appraisal Schedule (CANSAS) to assess needs [64] and the Threshold Assessment Grid to assess severity [65]. The patient measures chosen were the HAS, the CANSAS and the Manchester Short Assessment to assess quality of life [66].

Routine assessment of outcome in mental health services can indirectly benefit patients, by informing the development of programmes and systems [8]. It also has the potential to provide valuable treatment-level information which directly benefits patients [45]. The long-term goal is for routine outcome assessment to become an integral component of clinical care, rather than an administrative burden added on to the 'real' work of clinicians [5]. Carefully thought-out and well-resourced approaches to collecting and using outcome information are therefore needed, to avoid wasting effort and clinical goodwill. This review provides a method for service managers and clinicians who want to assess the impact of care on people using routine adult mental health services.

#### **Acknowledgements**

Thanks to Graham Thornicroft and Elizabeth Kuipers for their valuable advice. This work was supported by a Medical Research Council Clinician Scientist Fellowship (G108/390).

#### **References**

1. Andrews G, Peters L, Teesson M. Measurement of consumer outcome in mental health. In: CRUFAD, eds. *Clinical research unit for anxiety disorders. A report to the National Mental Health Information Strategy Committee*. Sydney: CRUFAD, 1994.
2. Slade M. Routine outcome assessment in mental health services. *Psychological Medicine* (in press).
3. Burns T, Priebe S. Mental health care systems and their characteristics: a proposal. *Acta Psychiatrica Scandinavica* 1996; 94:381–385.
4. Harrison G, Eaton W. From research world to real world. *Current Opinion in Psychiatry* 1999; 12:187–189.
5. Ellwood P. Outcomes management – a technology of patient experience. *New England Journal of Medicine* 1988; 318:1549–1556.

6. McKenzie N, Blanes T, Marks IM. Computerised clinical benefit-cost audit of mental health care II: time input, costs, patient satisfaction. *Journal of Mental Health* 1995; 1:71–78.
7. Walter G, Kirkby KC, Marks IM, Whiteford H, Andrews G, Swinson R. Outcome measurement – sharing experiences in Australia. *Australasian Psychiatry* 1996; 4:316–318.
8. Lyons JS, Howard KI, O'Mahoney MT, Lish JD. *The measurement and management of clinical outcomes in mental health*. New York: Wiley, 1997.
9. Dickey B, Azeni H. Data watch: impact of managed care on mental health services. *Health Affairs* 1992; 11:197–204.
10. Department of Health. *Mental health national service framework*. London: HMSO, 1999.
11. Wing JK, Beevor AS, Curtis RH, Park SB, Hadden S, Burns A. Health of the Nation Outcome Scales (HoNOS). *British Journal of Psychiatry* 1998; 172:11–18.
12. Stedman T, Yellowlees P, Mellisop G, Clarke R, Drake S. *Measuring consumer outcomes in mental health*. Canberra: Department of Health and Family Services, 1997.
13. NHS Centre for Reviews and Dissemination. *Undertaking systematic reviews of research on effectiveness: CRD's guidance for carrying out or commissioning reviews*, 2nd edn. York: University of York, 2001.
14. Shern DL, Flynn LM. Outcomes roundtable: developing, evaluating and disseminating outcomes monitoring. *Behavioural Healthcare Tomorrow* 1996; 5:25–30.
15. Smith GR, Manderscheid R, Flynn LM, Steinwachs DM. Principles for assessment of patient outcomes in mental health care. *Psychiatric Services* 1997; 48:1033–1036.
16. Campbell J. *Consumerism, outcomes and satisfaction: a review of the literature*. [Cited 16 August 2002.] Available from URL: <http://www.networksplus.net/fhp/madnation/citations/consumerism.htm>
17. Ciarlo JA, Edwards DW, Kiresuk TJ, Newman FL, Brown TR. *The assessment of client/patient outcomes techniques for use in mental health programs*. Washington, DC: National Institute of Mental Health, 1981.
18. Attkisson C, Cook J, Karno M *et al*. Clinical services research. *Schizophrenia Bulletin* 1992; 18:561–626.
19. Burnam MA. Measuring outcomes of care for substance use and mental disorders. *New Directions in Mental Health Services* 1996; 71:3–17.
20. Smith GR, Rost KM, Fischer EP, Burnam MA, Burns BJ. Assessing the effectiveness of mental health care in routine clinical practice. *Evaluation and the Helping Professions* 1997; 20:65–80.
21. Green RS, Gracely E. Selecting a rating scale for evaluating services to the chronically mentally ill. *Community Mental Health Journal* 1987; 23:91–102.
22. Wenger N, Mattson M, Furberg C. *Assessment of quality of life in clinical trials of cardiovascular therapies*. New York: LeJacq, 1984.
23. Bergner M, Rothman M. Health status measures: an overview and guide for selection. *American Review of Public Health* 1987; 8:191–210.
24. Hargreaves W, Shumway M. Effectiveness of mental health services for the severely mentally ill. In: Taube C, Mechanic D, Hohmann A, eds. *The future of mental health services research*. Washington, DC: US Government Printing Office, 1989; 253–284.
25. Ware J. Measuring health and functional status in mental health services research. In: Taube C, Mechanic D, Hohmann A, eds. *The future of mental health services research*. Washington, DC: US Government Printing Office, 1989; 289–302.
26. Rosenblatt A, Attkisson C. Assessing outcomes for sufferers of severe mental disorder: a conceptual framework and review. *Evaluation and Program Planning* 1993; 16:347–363.
27. Ruggeri M, Tansella M. Evaluating outcome in mental health care. *Current Opinion in Psychiatry* 1995; 8:116–121.
28. Sederer LI, Hermann MD, Dickey B. The imperative of outcome assessment in psychiatry. *American Journal of Medical Quality* 1995; 10:127–132.
29. Cook J, Jonikas J. Outcomes of psychiatric rehabilitation. *New Directions in Mental Health Services* 1996; 71:33–47.
30. McGlynn E. Setting the context for measuring patient outcomes. *New Directions in Mental Health Services* 1996; 71:19–32.
31. Schlosser B. New perspectives on outcomes assessment: the philosophy and application of the subjective health process model. *Psychotherapy* 1996; 33:284–304.
32. Clifford P. M is for outcome: the CORE outcomes initiative. *Journal of Mental Health* 1998; 7:19–24.
33. Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. *Health Technology Assessment* 1998; 2:14.
34. Jennings BM, Staggers N, Brosch LR. A classification scheme for outcome indicators. *Image Journal of Nursing Scholarship* 1999; 31:381–388.
35. Thornicroft G, Tansella M. *The mental health matrix*. Cambridge: Cambridge University Press, 1999.
36. Thornicroft G, Tansella M, eds. *Mental health outcome measures*, 2nd edn. London: Gaskell, 2000.
37. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *The Lancet* 1999; 354:1896–1900.
38. Khan KS, ter Riet G, Popay J, Nixon J, Kleijnen J. Phase 5: study quality assessment. In: Khan KS, ter Riet G, Glanville J, Sowden A, Kleijnen J, eds. *NHS Centre for Reviews and Dissemination Report No. 4. Undertaking systematic reviews of research on effectiveness*, 2nd edn. York: University of York, NHS CRD, 2000; 37–50.
39. Howe K, Eisenhart M. Standards for qualitative (and quantitative) research: a prolegomenon. *Educational Researcher* 1990; 19:2–9.
40. Buchanan DR. An uneasy alliance: combining qualitative and quantitative research methods. *Health Education Quarterly* 1992; Spring:117–135.
41. Oakley A. *Experiments in knowing, gender and method in social sciences*. Cambridge: Polity, 2000.
42. Seale S, Silverman D. Ensuring rigour in qualitative research. *European Journal of Public Health* 1997; 7:379–384.
43. Hunter E. 'Freedom's just another word': Aboriginal youth and mental health. *Australian and New Zealand Journal of Psychiatry* 1995; 29:374–384.
44. Durie M. Mental health and Maori development. *Australian and New Zealand Journal of Psychiatry* 1999; 33:5–12.
45. Marks I. Overcoming obstacles to routine outcome measurement. *British Journal of Psychiatry* 1998; 173:281–286.
46. Gilbody SM, House AO, Sheldon TA. Psychiatrists in the UK do not use outcome measures: national survey. *British Journal of Psychiatry* 2002; 180:101–103.
47. Slade M, Thornicroft G, Glover G. The feasibility of routine outcome measures in mental health. *Social Psychiatry and Psychiatric Epidemiology* 1999; 34:243–249.
48. Verhaak PF, Van de Lisdonk EH, Bor JH, Hutschemaekers GJ. GPs' referral to mental health care during the past 25 years. *British Journal of General Practice* 2000; 50:307–308.
49. McDonald R, Blizard R. Quality assurance of outcome in mental health care: a model for routine use in clinical settings. *Health Trends* 1988; 20:111–114.
50. Duggan C, Marks I, Richards D. Clinical audit of behaviour therapy training of nurses. *Health Trends* 1993; 25:25–30.

51. Slade M. The use of patient-level outcomes to inform treatment. *Epidemiologia e Psichiatria Sociale* 2002; 11:20–27.
52. Ruggeri M, Biggeri A, Rucci P, Tansella M. Multivariate analysis of outcome of mental health care using graphical chain models. The South-Verona outcome project 1. *Psychological Medicine* 1998; 28:1421–1431.
53. Marks I. Rapid audit of clinical outcome and cost by computer. *Australian and New Zealand Journal of Psychiatry* 1995; 29:32–37.
54. Marks IM, Blanes T, McKenzie N. Computerised clinical benefit-cost audit of mental health care I: theoretical and practical issues. *Journal of Mental Health* 1995; 4:63–70.
55. Marks I. A computerized system to measure treatment outcomes and cost. *Psychiatric Services* 1996; 47:811–812.
56. Long A, Dixon P. Monitoring outcomes in routine practice: defining appropriate measurement criteria. *Journal of Evaluation in Clinical Practice* 1996; 2:71–78.
57. Trauer T. Issues in the assessment of outcome in mental health. *Australian and New Zealand Journal of Psychiatry* 1998; 32:337–343.
58. Barry MM, Zissi A. Quality of life as an outcome measure in evaluating mental health services: a review of the empirical evidence. *Social Psychiatry and Psychiatric Epidemiology* 1997; 32:38–47.
59. Dickerson E. Assessing clinical outcomes: the community functioning of persons with serious mental illness. *Psychiatric Services* 1997; 48:897–902.
60. Bartlett CJ, Coles E. Psychological health and well-being: why and how should public health specialists measure it? Part 2: stress, subjective well-being and overall conclusions. *Journal of Public Health Medicine* 1998; 20:288–294.
61. American Psychiatric Association. *Handbook of psychiatric measures*. Washington, DC: American Psychiatric Association, 2000
62. Slade M, Cahill S, Kelsey W, Powell R, Strathdee G, Valiakalayil A. Threshold 3: the feasibility of the Threshold Assessment Grid (TAG) for routine assessment of the severity of mental health problems. *Social Psychiatry and Psychiatric Epidemiology* 2001; 36:516–521.
63. Priebe S, Gruyters T. The role of the helping alliance in psychiatric community care: a prospective study. *Journal of Nervous and Mental Diseases* 1993; 181:552–557.
64. Phelan M, Slade M, Thornicroft G *et al*. The Camberwell Assessment of Need: the validity and reliability of an instrument to assess the needs of people with severe mental illness. *British Journal of Psychiatry* 1995; 167:589–595.
65. Slade M, Powell R, Rosen A, Strathdee G. Threshold Assessment Grid (TAG): the development of a valid and brief scale to assess the severity of mental illness. *Social Psychiatry and Psychiatric Epidemiology* 2000; 35:78–85.
66. Priebe S, Huxley P, Knight S, Evans S. Application and results of the Manchester Short Assessment of quality of life. *International Journal of Social Psychiatry* 1999; 45:7–12.