

## Selecting outcome measures in mental health: the views of service users

MIKE J. CRAWFORD<sup>1</sup>, DAN ROBOTHAM<sup>2</sup>, LAVANYA THANA<sup>1</sup>,  
SUE PATTERSON<sup>1</sup>, TIM WEAVER<sup>1</sup>, ROSEMARY BARBER<sup>3</sup>, TIL WYKES<sup>4</sup>,  
& DIANA ROSE<sup>4</sup>

<sup>1</sup>Faculty of Medicine, Imperial College London, Centre for Mental Health, London, UK, <sup>2</sup>Mental Health Foundation, London, UK, <sup>3</sup>Section of Public Health, University of Sheffield, Sheffield, UK, and <sup>4</sup>Health Service and Population Research, Kings College London, London, UK

### Abstract

**Background.** Little is known about service users' views of measures used to evaluate treatments in mental health.

**Aims.** To identify the views of people with psychosis and affective disorder about the relevance and acceptability of commonly used outcome measures.

**Methods.** Twenty-four widely used outcome measures were presented to expert groups of service users. Nominal group methods were used to develop consensus about the appropriateness of each measure. Comments made by service users about how outcomes should be assessed were also recorded.

**Results.** Group members expressed concern about the ability of some outcome measures to capture their experiences. Patient-rated measures were assessed as more relevant and appropriate than staff-rated measures, and the need to examine negative as well as the positive effects of treatments was emphasised. Specific concerns were raised about some widely used measures including the Global Assessment of Functioning and the European Quality of Life scale.

**Conclusions.** We consider it essential that service users' views are taken into account when selecting measures to evaluate treatment outcomes. Providing insight into views of users of mental health services, our findings serve as a starting point for discussion.

**Keywords:** Views, attitudes of consumers, user involvement, outcomes, measurement

### Background

The development of effective treatments for health-related problems requires active input from people who use services (Partridge & Scadding, 2004). One area where active service user involvement is essential is in deciding which outcome measures should be used to examine the effectiveness of new treatments (Entwistle et al., 1998; The James Lind Alliance, 2007). Previous studies have demonstrated differences in the way that users and providers of health care judge whether interventions and treatments are

Correspondence: Mike J. Crawford, Reader in Mental Health Services Research, Department of Psychological Medicine, Faculty of Medicine, Imperial College London, Claybrook Centre, 37, Claybrook Road London W6 8LN, UK. Tel: +0207-386-1233. Fax: +0207-386-1216. E-mail: m.crawford@imperial.ac.uk

effective (Cay et al., 1975; Prusoff et al., 1972; Rothwell et al., 1997). These differences may be greater in mental health, where views about the nature of ill health and the desired outcomes of treatment may be more contentious (Faulkner & Thomas, 2002). Differences have been reported between users and providers of mental health services about which outcomes matter most (Crawford et al., 2008; Fischer et al., 2002) and studies that have actively involved service users report changing the way that interventions were subsequently evaluated (Trivedi & Wykes, 2002). Despite these findings, available evidence suggests that most studies examining the effectiveness of interventions and treatments are conducted with little or no involvement from service users (Hanley et al., 2001).

In recent years, outcome measurement has gained an even greater level of importance (Trauer et al., 2009). In addition to the central role that the selection of outcome measures plays in the design of clinical trials, greater emphasis has been placed on findings from trials in determining health policy (Ham et al., 1995). It has been argued that a greater emphasis on routine outcome measurement in clinical services can increase service quality (The Stationery Office, 2008a). In the National Health Service (NHS), funding for services is increasingly linked to the outcomes that services achieve (Department of Health, 2002). Choice of appropriate outcomes is therefore of major importance to both the quality of clinical trials and to efforts to improve the quality of mental health services.

Most outcome measures have been developed without direct input from service users and views of service users about the appropriateness of outcome measures that are widely used are largely unknown. We therefore set out to obtain views of service users about the appropriateness of widely used outcome measures. We focussed on outcome measures that have been used to evaluate interventions and services for people with psychosis and mood disorders, due to the prominence of these forms of mental disorder among users of mental health services.

## Method

We compiled a list of outcome measures that have been used to evaluate treatments and services for people with mood disorders and psychosis by searching textbooks and online resources and identifying measures used in published and ongoing clinical trials (American Psychiatric Association, 2005; Groth-Marnat, 2003; National Institute for Mental Health, 2008; Sajatovic & Ramirez, 2003; Thornley & Adams, 1998). These measures were presented to a Project Advisory Group (PAG) for discussion. Views of PAG members were used to draw up a shortlist of outcome measures that were then presented to expert groups of service users. The PAG had nine members with a range of different experiences and qualifications: three were full-time researchers; two members of the group worked as clinical academics; two members worked in research and development within voluntary sector organisations (The Mental Health Foundation and Rethink); and the last two members were service users with experience of undertaking research and service development activities. All members of the group had worked together as part of a Clinical Research Group on Service User Involvement set up by the Mental Health Research Network.

PAG members were presented with a document that provided a description of the measure, data on reliability and validity, and information about how often the measure had been used in previous studies, whether it had been included in recent guidance on outcome measurement in mental health published by the National Institute for Mental Health (2008), and national guidance on the treatment of schizophrenia and depression

(National Institute for Clinical Excellence, 2002, 2004). PAG members were asked to consider the following factors when shortlisting the measures:

1. The frequency with which they have been used in previous studies;
2. Evidence of their currently being used in ongoing research and service evaluations;
3. The psychometric properties of the measure.

Whenever more than one version of the same outcome measure was available, we used the one most frequently used in previously published studies.

### *Expert groups*

Group members were recruited through the Mental Health Research Network, four national mental health charities and two local user groups. We asked for volunteers who had used secondary care mental health services and had direct experience of mood disorder and/or psychosis. We aimed to recruit between 8 and 12 people to each expert group and to ensure representation of men and women of different ages and from different ethnic and cultural backgrounds. Previous involvement in a research study was considered helpful, but was not a requirement to become a group member. The first group meeting was set up to examine outcome measures used in studies on mood disorders and the second focussed on psychosis. In addition to examining measures to assess mood, the first group also assessed general mental health measures and measures assessing social functioning. The second group examined instruments used to assess side effects of medication and quality of life measures in addition to those specifically designed to assess psychotic symptoms.

Potential group members had often experienced both mood disorder and psychosis and chose which group they wanted to attend. Members of the group were sent written information prior to the meeting. At the start of the meeting, group members were asked to provide written informed consent to participate in the study. Members of the expert groups were paid £100 for taking part in each day-long meeting. Ethical approval was obtained prior to the start of the study from Imperial College Research Ethics Committee.

We commenced the expert groups with a general introduction in which we presented the background and aims of the study and described the format of the meeting. We then asked group members to complete a short questionnaire that sought information related to demographics, use of mental health services and previous involvement in research. We used a nominal group technique (Jones & Hunter, 1995) to try to develop consensus about the appropriateness of the different outcome measures presented. We started by presenting each outcome measure by providing a short summary of the background and aims of the measure, what it aimed to assess and how the measure was completed. We gave each group member a copy of the measure and then asked members to read through them and make an initial rating of their appropriateness on an 11-point Likert scale. Each outcome measure was then discussed and feedback on the mean response of the group in round 1 was provided. Members were then asked to rate the appropriateness of each measure again, in light of the initial mean group score. During each group, a service user researcher kept a written record of key discussion points.

### *Data analysis*

A summary measure of the expert groups' initial and final ratings for each outcome measure was calculated using the median rating and the interquartile range. Qualitative data collected

during the group meetings were used to examine views of service users about the use of outcome measures, to help interpret the quantitative data we collected and to describe aspects of outcome measures that service users believe makes them appropriate.

## Results

We identified 132 outcome measures that have been used in studies of people with psychosis and mood disorders. PAG members agreed a shortlist of 22 outcome measures to present to the expert groups. However, PAG members were concerned that our search strategy had excluded measures developed in consultation with service users but not in common use. We therefore included two additional measures – the Warwick-Edinburgh Well-being scale (Tennant et al., 2007) and the Recovery Star (MacKeith & Burns, 2008).

Twenty-five people attended the two nominal groups. Twelve people attended the first group and 15 attended the second (two of them had also participated in the first group). Participants ranged in age from 26 to 66 (mean = 44) years. Eleven (44%) participants were female and 14 (56%) were male. In terms of ethnicity, 15 (60%) participants identified themselves as British white, 5 (20%) as British black, 3 (12%) as British Asian and 2 (1%) as other. All 25 participants had previous contact with secondary care mental health services, with most ( $N=22$ , 88%) having had more than 10 years contact with services. Twenty-two (88%) participants had had some previous involvement in research or service evaluation.

Median ratings for each outcome measure (with interquartile range) in the initial and final rating exercise are presented in Table I. Levels of agreement were higher for the final rating (as indicated by a smaller interquartile range for most outcome measures). After the final rating, eight outcome measures had a median score of 7 or more and five had a median rating of 3 or lower. The highest ratings were for side effects of medication, and members of the expert group on mood disorder stated that they would have liked the chance to see measures that had been used to assess side effects of mood stabilisers and antidepressants. No measure of social functioning achieved a score higher than 5.

### *Comments made by expert group members*

Possible reasons why individual outcome measures achieved the ratings they did emerged from an analysis of the qualitative data that were collected during these two meetings. A series of features were identified as contributing to the appropriateness of an outcome measure. These are summarised in Table II. Group members expressed a strong preference for outcome measures that are rated by service users themselves. They expressed surprise, and in some instances disbelief, that outcome measures based entirely on the judgments of researchers or clinicians could be used to judge a person's response to treatment. Others stated that people may not be able to make ratings without help from others at times when they were most unwell. Some group members supported the idea that outcome measures might be completed by service users working together with staff (when completing the Recovery Star), although others felt that outcome measures should be based entirely on the experiences of the service user.

Several discussions took place about the length of questionnaires or assessment schedules. Concerns were expressed at the length of some measures, but greater concerns were expressed about very short outcome measures such as the European Quality of Life scale (EQ-5D), which group members believed were too short to properly assess the complex outcomes they are designed to measure. An exception to this was measures of side effects of

Table I. Median ratings of the appropriateness of 24 outcome measures.

Domain	Measure	Version	Median (interquartile range)	
			Initial rating	Final rating
Depression	Beck Depression Inventory	21-Item self administered	6 (2-9)	7 (4-7)
	Hamilton Depression Rating scale	21-Item clinician rated	6 (3-7)	5 (4-7)
	Patient Health Questionnaire (PHQ-9)	9-Item self administered	4 (3-7)	5 (3-7)
Schizophrenia	Psychotic Symptom Rating scales (PSYRATS)	17-Item clinician rated	5.5 (3-6.5)	5 (3-6)
	Positive and Negative Symptoms scale (PANSS)	30-Item clinician rated	7 (5-8)	7 (5.5-8)
	Scale for the assessment of positive (negative) symptoms (SAP/SAN)	50-Items clinician rated	6 (4.5-7.5)	6 (4-6.5)
General mental health	Warwick-Edinburgh Well-being scale (WEWES)	14-Item self administered	7.5 (5-8)	7.5 (5-8)
	CORE outcome measure (CORE-OM)	34-Item self administered	7 (6-8)	7 (6-8)
Social functioning/disability	General Health Questionnaire (GHQ-12)	12-Item self administered	3 (4-6)	4 (3-5)
	Work and Social Adjustment (WASA)	5-Item self administered	5 (3-6)	5 (3-5)
	Social Functioning Questionnaire (SFQ)	8-Item self administered	4 (2-6)	3 (2-5)
	WHO-Disability Assessment Schedule (WHO-DAS II)	12-Item interviewer administered	3 (2-5)	3 (2-4)
Quality of life	Short-form 12-item Health Survey (SF-12)	12-Item self administered	3 (2-4)	3 (2-4)
	World Health Organisation - Quality of Life (WHO-QoL)	26-Item interviewer administered	7.5 (5-8)	7.5 (6-8.5)
	Manchester Short Assessment of Quality of Life (MANSA)	17-Item self administered	7 (5-8)	7 (6-8)
Side effects of medication	Quality of Life Scale (QoL)	16-Item self administered	6 (5-7)	6 (5-7)
	European Quality of Life scale (EQ-5D)	6-Item self administered	3 (2-5)	3 (2-4)
	Liverpool University Neuroleptic Side Effect (LUNTERS)	51-Item self administered	7.5 (5.5-9)	8 (7-8)
	Antipsychotic Non-Neurological Side-Effects Rating scale (ANSSERS)	44-Item clinician administered	5.5 (4.5-7)	7 (5-7)
Other measures	Udvalg for Kliniske Unders Gelser (UKU)	48-Item clinician administered	5.5 (4.5-8)	6.5 (5-8)
	Recovery star	10 Items completed collaboratively	5.5 (4-7.5)	6.5 (5-7.5)
	Clinical global impression	Single question self rated	5 (2-6)	5 (3-6)
	Health of the Nation Outcome scale (HoNOS)	12-Item clinician administered	6 (3-8)	5 (4-8)
	Global Assessment of Functioning (GAF)	Single-item clinician rated	4 (2-5)	3 (3-4)

Table II. Features of an appropriate outcome measure according to group members.

---

Should be based on patient rather than staff-rated judgements
Includes 'positive' as well as 'negative' items
Is comprehensive – neither too long nor too short
Avoids questions that are intrusive about private issues such as sex life
Makes note of the time and place where the outcome is measured
Includes space for 'added comments'
Should be used by staff who have good interpersonal skills and have been properly trained in the use of the outcome measure
If self completed, is presented in a professional manner

---

medication, which group members felt needed to be long in order to fully capture the wide range of negative effects of treatments that group members had experienced.

Another recurrent theme concerned the dominance of what were perceived as too many 'negative' items in most questionnaires. Many group members remarked that they found it upsetting to be asked long lists of questions about difficulties associated with mental ill health. Conversely, The Warwick-Edinburgh Well-being scale was commended for asking questions about aspects of good mental health, where poor emotional health is indicated by not endorsing these 'positive' items.

In both groups, members raised concerns about outcome measures that set out criteria for a 'good' outcome. This applied particularly to measures of quality of life and social functioning. For instance, some group members expressed the view that it should not be assumed that people who got on well with family members had better social functioning, because some people made a conscious choice not to have contact with family members. Some group members considered it inappropriate that measures were grounded in judgements about which aspects of relationships with others or ways that people spend their time indicated a 'better' outcome. Group members suggested, instead, that people should be asked whether people were happy with these aspects of their lives.

Varying views were expressed about inclusion of items on sensitive matters, such as sex. Some group members stated that they would not be happy to answer such questions for the purpose of assessing outcomes and considered the use of 'intrusive' questions in outcome measures such as the Social Functioning Questionnaire likely to impact on people's willingness to take part in research. Others disagreed and felt that issues such as sex were important to quality of life and needed to be included.

Concerning the process of outcome measurement group members repeatedly stated that it was not so much the content or length or an interview or questionnaire that mattered, as much as the personal qualities of the clinician or researcher conducting the assessment. As one group member stated that an outcome measure was '*only as good as the doctor who uses it*'. Sensitivity to the needs of people experiencing mental distress and being treated with dignity and respect were seen as particularly important. At a practical level, group members stressed the importance of presentation of documents. Reporting experiences of being asked to complete poor photocopies of questionnaires, group members stated that poor presentation made them wonder about how seriously their responses would be taken or whether services valued the information they were being asked to provide.

## Discussion

By using nominal group methods with expert groups comprising users of mental health services, we have generated data on service user perspectives on the appropriateness of

widely used outcome measures. Group members were clear that the outcomes of interventions and treatments needed to be assessed using patient-rated measures. They highlighted the importance of assessing negative as well as positive effects of interventions and treatments and were critical of outcome measures that presented people with only negatively phrased questions. Many comments made by service users concerned the process of collecting outcome measures rather than the content of individual measures. The personal qualities and training of staff using questionnaires and checklists were highlighted as important to the experience, as were the presentation and format of questionnaires.

Although measures of medication side effects assessed by the group were all relatively highly regarded, it is of concern that some of the most widely used outcome measures in mental health, including the EQ-5D and the Global Assessment of Functioning received very low ratings. In addition to being staff rated, the Global Assessment of Functioning was criticised for conflation of functioning with mental health in a single score. Some group members commented that there had been times when functioning was reasonably good despite poor mental health, and others reported continuing to function poorly even after symptoms of emotional distress had subsided. Group members questioned whether it was possible to assess these outcomes with a single score. This concern has also been raised by clinicians and researchers using the Global Assessment of Functioning (Williams, 2005).

### *Study strengths and limitations*

Although we were able to include many measures in commonly used to assess outcomes of people with mood disorders and psychosis in studies in the UK, many others were not included. Moreover, several of the assessed outcome measures are available in differing formats. We attempted to use the versions of these outcome measures that have been most widely used, for example, we asked group members to consider the appropriateness of the 12-item General Health Questionnaire, rather than the 28- or 60-item version of this questionnaire (Goldberg & Williams, 1991) and the 34-item version of the CORE-OM but do not know how different versions would have been rated (Evans et al., 2002). In a similar vein, while a self-completed version of Psychotic Symptom Rating Scales is available we presented group members with the staff-rated version. Given the preference of group members for patient-rated outcome measures, the self-completed version of the measure may have received a higher median rating.

The criteria established for developing our shortlist were designed to ensure that we examined outcome measures that have been widely used. This meant that measures that have only recently been developed were generally excluded. An unfortunate consequence of this was that a range of recently developed outcome measures that have been developed in partnership with service users were excluded (Allott, 2005; Rose et al., 2009). The only measure that we examined which made explicit reference to incorporating views of service users during its development was the Warwick-Edinburgh Well-being scale (Tennant et al., 2007). It is therefore of note that this outcome measure achieved one of the highest rating by group members.

Although differences in the median level of appropriateness of outcome measures were seen in this study, we have not used statistical tests to assess the strength of the difference. This is because the study was not designed to try to make decisions about whether a particular outcome measure was 'appropriate' or not, but to provide a description of the views of people with psychosis and affective disorder about the relevance and acceptability

of commonly used outcome measures. For instance, we do not know whether users of mental health services judge the Warwick-Edinburgh Well-being scale a more relevant and acceptable assessment of general mental health than the General Health Questionnaire, and apparent differences in the median ratings for these two measures could have arisen by chance. However, we did find that service users who took part in the expert groups liked the Warwick-Edinburgh Well-being scale and valued the way that it asks about positive aspects of mental health.

The 24 outcome measures we assessed were divided between the two expert panels. While it made sense to ask people with experience of being treated for depression to rate outcome measures for mood disorders, and for people with psychosis to rate outcome measures in schizophrenia, either panel could have been asked to rate generic outcome measures of quality of life and social functioning, etc. We do not know whether ratings of these measures would have differed had they been presented to a different panel.

Finally, the qualitative data we collected were in the form of contemporaneous notes. We did not use formal qualitative research methods to explore service user views about outcome measurement. While we were able to capture the main themes of the discussions that took place at the two meetings, we believe that future qualitative research should be conducted to help develop a better understanding of service user views of outcome measurement in mental health.

### *Implications of findings*

One of the clearest findings of the study was the preference of group members for patient-rated outcome measures. This finding contrasts with extensive reliance on staff-rated outcome measures seen in mental health services (Gilbody et al., 2002). Concerns have been raised about the reliability and validity of patient-rated outcome measures in mental health (Becker et al., 1993), and current efforts to examine service quality are often centred on staff-rated outcome measures such as the Health of the Nation Outcome Scale (HONOS) (Fairbairn, 2007). This contrasts with other areas of health care where patient-rated outcome measures are the norm. The service users who took part in this study challenged the idea that the quality of services or outcomes of interventions could be measured appropriately by using staff-rated measures and we believe that these data strengthen the argument for using patient-rated outcome measures in mental health. Ongoing research to develop and refine a patient-rated HONOS is therefore to be welcomed (Trauer & Callaly, 2002).

Our finding that service users are concerned with assessment of side effects of treatment accords with other studies (Lee et al., 2000; Shumway et al., 2003). We extend this however beyond side effects of medication to include examination of the negative effects of psychological treatments. Members of the group focussed on mood disorder stated that they would have liked to have seen an outcome measure assessing the side effects of mood stabilisers and antidepressants. These findings support the inclusion of measures of negative as well as positive effects of interventions and treatment used in mental health (Rose et al., 2008).

Group members were also concerned about the length of many of the outcome measures they were asked to examine. Similar concerns were also raised by lay people, patients and health care professionals who took part in a previous survey of the acceptability of self-reported outcome measures used to assess the health of people with personality disorder (Blount et al., 2002). In that study, 52 people including 7 lay people and 14 ex-users of a residential treatment service for people with personality disorder were asked to comment on



five outcome measures. While shorter measures were generally rated more positively than longer ones, both lay people and professionals who took part in this study criticised the two shortest outcome measures as ‘too simplistic’ (Blount et al., 2002).

Concerns have long been expressed about the validity and utility of widely used measures of quality of life (The Stationery Office, 2008b). While outcome measures such as the EQ-5D have been specifically developed to allow comparison across medical settings, members of our expert groups expressed the view that this measure was insufficiently detailed to capture the quality of life of people with mental disorders. They preferred longer measures such as the World Health Organisation – Quality of Life (WHO-QoL) that includes items addressing a broader range of factors judged by service users to be important when assessing quality of life.

Group members rated measures relating to social functioning particularly poorly. Service users had identified social functioning as a domain that was particularly important to assess when judging the outcomes of interventions and treatments (Perry & Gilbody, 2009). However, group members were critical of the normative nature of these measures. Group members told us that they wanted to be asked about whether they were satisfied with different aspects of their social functioning rather than having assumptions made that fewer contacts with others or involvement in fewer occupational activities was necessarily an indicator of a poorer outcome. Some group members also raised concerns about inclusion of ‘intrusive’ items such as questions about sex life in such outcome measures. While there are already a large number of outcome measures used in mental health, we believe that these findings suggest that a measure of social functioning that focuses on areas of function that service users consider appropriate and relevant and does not use normative standards is required.

## Conclusions

Asked to assess various outcome measures widely used in studies of people with psychosis and mood disorder, users of mental health services considered some to be more appropriate than others. Our data emphasise the importance service users place on patient-rated outcome measures. Service users who took part in our expert groups stated that studies should examine the negative as well as the positive effects of interventions and treatments. Concerns were raised about some widely used outcome measures such as EQ-5D and Global Assessment of Functioning (GAF) and about all the four measures of social functioning that we examined.

## Acknowledgments

We are grateful to Angela Sweeny, Alison Faulkner, Joanna Fox and Jayasree Kalathil for their help with data collection and analysis and to Janey Antoniou, Neil Armstrong, Valerie Baker, Terry Bowyer, Humphrey Greaves, Michael Knight, Gary Molloy, Augusto Monteiro, Ros Newnham, Graham Peacock, Kay Sheldon, Roger Smith, Caroline Thomas, Jennifer Trite, Lauren Wright and others who attended the expert group meetings.

The study was funded by a project grant from the Mental Health Research Network.

**Declaration of interest:** The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

## References

- Allott, P. (2005). Recovery. In D. Sallah & M. Clarke (Eds.), *Research and Development in Mental Health: Theory, Framework and Models*. Oxford: Elsevier Science.
- American Psychiatric Association. (2005). *Handbook of Psychiatric Measures*. Washington DC: APA.
- Becker, M., Diamond, R., & Sainfort, F. (1993). A new patient focused index for measuring quality of life in persons with severe and persistent mental illness. *Quality of Life Research*, 2, 239–251.
- Blount, C., Evans C., Birch, S., Warren, F., & Norton K. (2002). The properties of self-report research measures: Beyond psychometrics. *Psychology and Psychotherapy: Theory, Research and Practice*, 75, 151–164.
- Cay, E.L., Small, W.P., Neilson, J., & Henderson, A. (1975). Patients assessment of the result of surgery for peptic ulcer. *Lancet*, 1, 29–31.
- Crawford, M.J., Price, K., Rutter, D., Moran, P., Tyrer, P., Bateman, A., et al. (2008). Dedicated community-based services for adults with personality disorder: Delphi study. *British Journal of Psychiatry*, 193, 342–343.
- Department of Health. (2002). *Reforming NHS Financial Flows: Introducing Payment by Results*. London: DoH.
- Entwistle, V.A., Renfrew, M.J., Yearley, S., Forrester, J., & Lamont, T. (1998). Lay perspectives: Advantages for health research. *British Medical Journal*, 316, 463–466.
- Evans, C., Connell, J., Barkham, M., Margison, F., McGrath, G., Mellor-Clark, J., & Audin, K. (2002). Towards a standardised brief outcome measure: Psychometric properties and utility of the CORE-OM. *British Journal of Psychiatry*, 180, 51–60.
- Fairbairn, A. (2007). Payment by results in mental health: The current state of play in England. *Advances in Psychiatric Treatment*, 13, 3–6.
- Faulkner, A., & Thomas, P. (2002). User-led research and evidence-based medicine. *British Journal of Psychiatry*, 180, 1–3.
- Fischer, E.P., Shumway, M., & Owen, R.R. (2002). Priorities of consumers, providers, and family members in the treatment of schizophrenia. *Psychiatric Services*, 53, 724–729.
- Gilbody, S.M., House, A.O., & Sheldon, T.A. (2002). Psychiatrists in the UK do not use outcomes measures. National survey. *British Journal of Psychiatry*, 180, 101–103.
- Goldberg, D., & Williams, P. (1991). *A User's Guide to the General Health Questionnaire*. Windsor: NFER-Nelson.
- Groth-Marnat, G. (2003). *Handbook of Psychological Assessment*. Hoboken, NJ: John Wiley & Sons.
- Ham, C., Hunter, D.J., & Robinson, R. (1995). Evidence based policymaking. *British Medical Journal*, 310, 71–72.
- Hanley, B., Truesdale, A., King, A., Elbourne, D., & Chalmers, I. (2001). Involving consumers in designing, conducting, and interpreting randomised controlled trials: Questionnaire survey. *British Medical Journal*, 322, 519–523.
- Jones, J., & Hunter, D. (1995). Consensus methods for medical and health services research. *British Medical Journal*, 311, 380.
- Lee, T.T., Ziegler, J.K., Sommi, R., Sugar, C., Mahmoud, R., & Lenert, L.A. (2000). Comparison of preferences for health outcomes in schizophrenia among stakeholder groups. *Journal of Psychiatric Research*, 34, 201–210.
- MacKeith, J., & Burns, S. (2008). *Mental Health Recovery Star*. London: Mental Health Providers Forum.
- National Institute for Clinical Excellence. (2002). *Clinical Guideline 1. Schizophrenia: Core Interventions in the Treatment and Management of Schizophrenia in Primary and Secondary Care*. London: NICE.
- National Institute for Clinical Excellence. (2004). *Depression: Management of Depression in Primary and Secondary Care*. London: NICE.
- National Institute for Mental Health. (2008). *Mental Health Outcomes Compendium*. London: NIMHE.
- Partridge, N., & Scadding, J. (2004). The James Lind Alliance: Patients and clinicians should jointly identify their priorities for clinical trials. *Lancet*, 364, 1923–1924.
- Perry, A., & Gilbody, S. (2009). User-defined outcomes in mental health: A qualitative study and consensus development exercise. *Journal of Mental Health*, 18, 415–423.
- Prusoff, B.A., Klerman G.L., & Paykel, E.S. (1972). Concordance between clinical assessments and patients' self-report in depression. *Archives of General Psychiatry*, 2, 546–552.
- Rose, D., Sweeney, A., Leese, M., Clement, S., Jones, I.R., Burns, T., et al. (2009). Developing a user-generated measure of continuity of care: Brief report. *Acta Psychiatrica Scandinavica*, 119, 320–324.
- Rose, D., Wykes, T., Farrier, D., Doran, A.-M., Sporle, T., & Bogner, D. (2008). What do clients think of cognitive remediation therapy? A consumer-led investigation of satisfaction and side effects. *American Journal of Psychiatric Rehabilitation*, 11, 181–204.
- Rothwell, P.M., McDowell, Z., Wong, C.K., & Dorman, P.J. (1997). Doctors and patients don't agree: Cross sectional study of patients' and doctors' perceptions and assessments of disability in multiple sclerosis. *British Medical Journal*, 314, 1580–1583.
- Sajatovic, M., & Ramirez L.F. (2003). *Rating Scales in Mental Health* (2nd Ed.). Hudson, Ohio: Lexi-Comp.

- Shumway, M., Saunders, T., Shern, D., Pines, E., Downs, A., Burbine, T., et al. (2003). Preferences for schizophrenia treatment outcomes among public policy makers, consumers, families, and providers. *Psychiatric Services, 54*, 1124–1128.
- Tennant, R., Hiller, L., Fishwick, R., Platt, S., Joseph, S., Weich, S., et al. (2007). The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): Development and UK validation. *Health and Quality of Life Outcomes, 5*, 63.
- The James Lind Alliance. (2007). The research cycle. Retrieved from <http://www.lindalliance.org/>
- The Stationery Office. (2008a). *High Quality Care for All: NHS Next Stage Review Final Report*. London: The Stationery Office.
- The Stationery Office. (2008b). *House of Commons Health Committee: National Institute for Health and Clinical Excellence. First Report of Session 2007–08*. London: The Stationery Office.
- Thornley, B., & Adams, C. (1998). Content and quality of 2000 controlled trials in schizophrenia over 50 years. *British Medical Journal, 317*, 1181–1184.
- Trauer, T., & Callaly, T. (2002). Concordance between mentally ill patients and their case managers using the Health of the Nation Outcome Scales (HoNOS). *Australasian Psychiatry, 10*, 24–28.
- Trauer, T., Callaly, T., & Herrman, H. (2009). Attitudes of mental health staff to routine outcome measurement. *Journal of Mental Health, 18*, 288–297.
- Trivedi, P., & Wykes, T. (2002). From passive subjects to equal partners: Qualitative review of user involvement in research. *British Journal of Psychiatry, 181*, 468–472.
- Williams, J.B.W. (2005). Mental health status, functioning and disabilities measures. In American Psychiatric Association. (Ed.), *Handbook of Psychiatric Measures* (pp. 93–115). Washington DC: American Psychiatric Association.