

**Patient-reported
Health Instruments
Group**

**Instruments for
Mental Health: a
Review**

**Louise J. Schmidt
Andrew M. Garratt
Ray Fitzpatrick**

**Report to the Department of Health
September 2000**



health

Outcome

indicators

PATIENT-REPORTED HEALTH INSTRUMENTS GROUP
(formerly the Patient-assessed Health Outcomes Programme)

INSTRUMENTS FOR MENTAL HEALTH: A REVIEW

**A STRUCTURED REVIEW OF
PATIENT-REPORTED HEALTH INSTRUMENTS**

Louise J. Schmidt
Andrew M. Garratt
Ray Fitzpatrick

*Patient-reported Health Instruments Group
National Centre for Health Outcomes Development (Oxford site)
Unit of Health-Care Epidemiology
Department of Public Health
University of Oxford*

September 2000

This report should be referenced as follows:

Schmidt LJ, Garratt AM, Fitzpatrick,R. *Instruments for Mental Health: a Review Report from the Patient-reported Health Instruments Group (formerly the Patient-assessed Health Outcomes Programme) to the Department of Health, September 2000.*

Copies of this report can be obtained from:

Dr Kirstie Haywood
Co-Director, Patient-reported Health Instruments Group
National Centre for Health Outcomes Development (Oxford site)
Unit of Health-Care Epidemiology
Department of Public Health
University of Oxford
Old Road
Headington
Oxford OX3 7LF

tel: +44 (0)1865 227157

e-mail: kirstie.haywood@uhce.ox.ac.uk

Alternatively, it can be downloaded free of charge from the PHIG website:

<http://phi.uhce.ox.ac.uk/>

LIST OF CONTENTS

Frontispiece	1
List of Contents	2
Executive Summary	3
Chapter 1: INTRODUCTION	6
a) Aims of the review	6
b) Mental illness	6
c) Patient-reported health outcome measures	7
d) Why broaden the approach to mental health outcomes?	8
Chapter 2: METHODS	10
a) Search strategy	10
b) Inclusion/exclusion criteria	10
Chapter 3: RESULTS	11
a) Search results	11
b) Quality of the reviews	11
c) Analysis of the instruments	12
d) Recommendations for specific instruments	15
Chapter 4: DISCUSSION	39
a) Conceptual basis of quality of life	39
b) Subjective or objective evaluation	40
c) Content of quality of life measures	42
Chapter 5: OTHER ISSUES IN INTERPRETING OUTCOMES INFORMATION	43
Chapter 6: CONCLUSIONS	44
Reference List	46
Appendix: psychometric evaluation	51

Executive Summary

Background

The term ‘mental illness’ covers a wide range of mental disorders including severe psychotic illness (such as schizophrenia) and chronic non-psychotic illness (principally depression and anxiety). For both types of mental illness the objectives of care are improved health and social functioning and reduced mortality from suicide and physical illness (Jenkins, 1992). Patient-reported health outcome measures¹ offer a method of assessing the outcome of care or interventions (in terms of symptoms, functioning, subjective well-being or quality of life) from the patient’s perspective, via the use of questionnaires, interviews and other methods.

Research Aims

1. examine why patient-based outcome measures are increasingly used;
2. summarise the range of instruments available for measuring patient-reported health outcomes in the area of mental health;
3. identify any consensus existing in reviews of mental health measures as to the most suitable instruments;
4. identify and discuss conceptual and methodological issues in measuring patient-reported health outcomes in this field.

Methods

Relevant literature was identified via a search of several electronic databases (Embase, Medline, Biological Abstracts, Psychlit, AMED, Econlit, Sociological Abstracts and Cinahl) from 1990 onwards. The literature was searched using (i) a previously developed comprehensive strategy for capturing references relating to patient-reported outcome measures combined with (ii) specific terms relating to mental health/illness and the type of publication we wished to retrieve i.e. reviews. The papers of abstracts relating to reviews of available patient-based outcome instruments for mental health/illness were obtained for examination, although only those giving summary ratings to the instruments’ psychometric properties were included in the final analysis. This report restricts itself to an examination of multi-dimensional measures and those relating to functioning or quality of life. Measures of symptoms are not included.

Key findings

The literature search generated four peer-reviewed structured reviews of patient-reported outcome measures in the field of mental health, which rated instruments according to their psychometric properties. In addition one published compendium of measurement scales was identified as meeting the inclusion criteria, in terms of providing ratings of instruments, and was included as a result.

A total of 85 instruments were identified across these reviews and summary information, together with the ratings assigned by the reviews, are reproduced in this report. Many of the instruments were found to be lengthy and relatively time-intensive. Unlike with the physical diseases, a significant number of measures were completed on

¹ The terms ‘instrument’ and ‘measure’ are used interchangeably throughout this report.

behalf of the patient by proxy. The vast majority of instruments were potentially applicable to a mental health population and were considered potentially suitable for use in either research or clinical practice.

Most of the reviews make their recommendations on the basis of the extent of psychometric evaluation of the instruments although one considered the particular applicability of the instrument to the routine practice setting (Andrews et al, 1994). The Lehman Quality of Life Interview (Lehman, 1983) was recommended by a majority of the reviews (Atkinson & Zibin, 1996; van Nieuwenhuizen et al, 1997; Lehman, 1996).

The Oregon Quality of Life Questionnaire/Scale/Interview (Bigelow et al, 1991) and the Lancashire Quality of Life Profile (Oliver, 1992) were both recommended by two of four reviews (Atkinson & Zibin, 1996; Lehman, 1996, and Lehman, 1996; van Nieuwenhuizen et al, 1997 respectively).

Other instruments singled out as showing promise and/or for use among specific populations include two disease-specific measures: the Quality of Life in Depression Scale (McKenna & Hunt, 1992) and the Quality of Life Scale (Heinrichs et al, 1984), for depression and schizophrenia respectively. Finally the following were also noted for merit: the SF-36 (Ware & Sherbourne, 1992); the Quality of Life, Enjoyment and Satisfaction Questionnaire (Endicott et al, 1993); the Quality of Life Index for Mental Health (Becker et al, 1993); the Quality of Life Self-Assessment Inventory (Skantze & Malm, 1994).

Key conclusions and recommendations

1. In spite of the potential difficulties in obtaining outcomes information from patients with certain types of mental illness, evidence suggests that even among those with chronic and severe mental illness, patients are able to evaluate their condition and the information they provide is unique and invaluable.
2. There is a wealth of measures available (functional, quality of life and multi-dimensional) for measuring patient's perspectives of outcomes and it is clear that no single measure will meet all needs for outcomes data across all settings. The choice of instrument depends upon the application for which the instrument is intended (research, evaluation, individual patient care or population assessment) and the nature of the outcome information desired. It also includes factors such as the nature of the mental illness (for some types of mental illness reliance on self-report alone may be unwise) and whether or not the intervention is multi-disciplinary (which would support the use of a broader definition of outcome, rather than a more restricted health-related one). Regardless of the nature of the intervention or outcome, the instrument chosen for use should, however, have well-established psychometric properties.
3. Several instruments are singled out for merit, a number of which are suitable for use in a research/evaluative context. A few are considered potentially useful for survey use at the population level, although it has been commented that currently available measures tend to focus on psychological ill-health and mental illness at the expense of positive mental health and psychological well-being (Bartlett & Coles, 1998). Several instruments are proposed for potential use in clinical practice however caution must be exercised here since it is generally considered in the patient-

reported health outcomes field that evidence to date does not support the use of such instruments in routine practice (Fitzpatrick et al, 1998). In particular, it has been suggested that instruments do not yet meet the stringent psychometric criteria for use at the individual patient level (Greenhalgh & Meadows, 1999).

4. Finally, unresolved issues remain, including, most fundamentally, a lack of a clear consensus of what quality of life constitutes and how it should be measured. Although it is sometimes suggested that different parties, not just the patient, should evaluate quality of life, it is unclear how these evaluations should be reconciled.

Chapter 1: INTRODUCTION

1 a) Aims

The main aims of the review are to:

1. examine why patient-based outcome measures are increasingly used;
2. summarise the range of instruments available for measuring patient-reported health outcomes in mental health;
3. identify any consensus existing in reviews of mental health measures as to the most suitable instruments;
4. identify and discuss conceptual and methodological issues in measuring patient-reported health outcome in this field.

1 b) Mental health/illness

The term ‘mental disorders’ covers a wide range of mental illness which includes severe psychotic illness (such as schizophrenia) and chronic non-psychotic illness (principally depression and anxiety). For both types of mental illness the objectives of care are improved health and social functioning and reduced mortality from physical illness and suicide (Jenkins, 1992).

According to the Department of Health report ‘Modernising Mental health Services: Safe, sound and supportive’ (1998), mental illness affects about one in six adults at any one time. Mental illness ranges from high levels of anxiety and deep depression to severe illness like schizophrenia (occurring in about one in 100). Depression will affect nearly half of all women and one quarter of all men in the UK before the age of 70. There are over 4,000 deaths from suicide in England each year.

It has been estimated that for every 100 people consulting their GPs with mental health problems, nine will be referred to specialist services. There are approximately 8,000 contacts with mental illness per year per 100,000 population among the formal psychiatric services and approximately 22,000 mental illness contacts per year per 100,000 population among general practitioners. Wing (1994) reported that:

- 26% of attenders in general practice have a mental disorder
- 1% of attenders in consultant out-patient clinics have a mental disorder
- 0.1% of in-patients have a mental disorder

(Charlwood et al (eds), 1999).

Various research studies have suggested that there may be a significant psychiatric component in up to one-third of general practice consultations (Jenkins & Griffiths (eds), 1991).

In terms of current trends, the biggest increase in new contacts with the psychiatric services over the last decade was found to have come from those aged over 75 (ibid). This national report presents data from one district suggesting that schizophrenia is a constant burden whereas the contacts for depression and dementia are increasing. It is estimated that the number of people with dementia alone will rise from the current 665,000 to 855,000 by the year 2020 (DoH 1997, in Selai & Trimble, 1999).

The precise aetiology of many forms of mental illness is unknown, although it is thought that certain factors (genetic, physical, physiological and socio-economic) may pre-dispose individuals to mental illness (Charlwood et al, 1999).

Associations with severe mental illness in the UK are:

- Gender (depression is more common among women)
- Age (the incidence of severe depression increases with age)
- Marital status (strong association between unmarried status and schizophrenia)
- Social environment
- Childhood experience

(ibid)

For people in remission from mental health problems, certain factors can either protect against, or precipitate, relapse including: life events; interpersonal relationships; work, occupation or recreation; and effective interventions, including social support (ibid).

Types of intervention to restore functioning include, for the severely mentally ill, medication, psychotherapies, psycho-educational interventions and strategies for self-management. The range of interventions aimed at promoting independent living and well-being can include:

- long-term care in supported accommodation
- the optimal use of medication
- specific techniques used in rehabilitation programmes
- care management of people living independently
- improved access to work and leisure
- safe environment
- meaningful daytime occupation and welfare advice.

(ibid)

Interventions aimed at improving the quality of life of patients with mental illness can take place in a variety of settings (long-term hospital accommodation, sheltered accommodation, as well as in the community) and can be delivered by a multi-disciplinary team of professionals and voluntary organisations. In fact, the Department of Health's Care Programme Approach stresses the desirability and the necessity of integration between the health and social services, as well as close liaison with the voluntary and private sectors in providing long-term care in the community for people with severe mental illness. It has been suggested that this Care Programme Approach may be improved for users and carers by adopting the service users' view of a good outcome (ibid).

1 c) Patient-reported health outcome measures

The term 'patient-reported health outcome measures' refers to a variety of questionnaires, interviews and other methods of assessing the health of a patient as seen from the patient's perspective. Methods include specific measures of symptoms, health-related quality of life and functional health status, the common feature being that patients themselves assess their situation. Patient-based outcome measures are often used in research (particularly clinical trials) but less often in the area of individual

patient care and for assessing the health care needs of the population. To be suitable for potential applications, patient-reported measures must have evidence of reliability, validity and responsiveness. Reliability is concerned with whether an instrument's scores are reproducible and internally consistent. Validity refers to whether an instrument measures what is intended. Responsiveness addresses whether an instrument is responsive to important changes in health. For a detailed explanation of these terms, the reader is invited to refer to the appendix to this report.

1 d) Why broaden the approach to mental health outcomes?

Although psychiatry may initially have been slow to become involved in quality of life assessment (Stedman, 1996; Gladis et al., 1999), the concept of quality of life, in connection with the process of deinstitutionalisation, has altered the perception of the type of care that should be offered and the objectives of that care. This has been accompanied by a movement, on the one hand, from concentrating on the symptoms associated with the patient's clinical state to looking at the patient's living conditions, and, on the other, a movement from an objective assessment of the patient's needs and services provided to a patient-centred subjective evaluation of needs and outcomes (Mercier, 1994). This movement arose from early empirical research demonstrating a discrepancy between objective conditions such as comfort or health and satisfaction with these conditions (Fabian, 1990).

As to why quality of life measurement was initially slow to catch on this field, the argument has been advanced that psychiatrists have been studying quality of life but it has gone by other names e.g. social support, burden and life events (Lehman, 1997). The literature commonly quotes five reasons for the adoption of quality of life assessments in community psychiatry as proposed by Baker and Intagliata (Mercier, 1994). These are:

- interventions to improve mental health are more directed at comforting or alleviating the burden of mental illness rather than curing patients
- interventions to reduce the morbidity associated with mental illness are often complex
- quality of life assessments have the potential to take client satisfaction into account
- since the WHO's definition of health, health has been increasingly viewed in a holistic manner
- the quality of life theme is politically correct as evidenced by US President Lyndon B. Johnson's statement in the 1960s ("these goals cannot be measured by the size of our bank balances, they can only be measured in the quality of life that our people lead"). (Rescher, 1972 in Katschnig et al, 1997).

Further reasons are suggested (Oliver et al., 1996) as to why the concept of quality of life has been taken up in the field of mental health and illness namely:

- quality of life measures are positively regarded by patients and relatives
- quality of life measures can enable the evaluation of multi-disciplinary work
- brief quality of life measures are an inexpensive contribution to evaluatory work.

Lastly there is widespread recognition of the fact that routinely used proxy indicators of outcome used in evaluating services, such as mortality and service use, are not reliable

indicators (Repper & Brooker, 1998). The term quality of life appears often in Department of Health policy papers relating to ideal and operational conditions and, throughout the UK, is being translated into local policy (Oliver et al, 1996).

Although this report does not restrict itself to consideration of severe and persistent mental illness, this was the focus of published material retrieved. Mental health research has tended to focus on psychological dysfunction (mental illness) rather than psychological functioning or well-being (Ryff & Singer, 1996) and the outcome measures reflect this.

Chapter 2: METHODS

2 a) Search Strategy

The search strategy used to search the electronic databases (Embase, Medline, Biological Abstracts, Psychlit, AMED, Econlit, Sociological Abstracts and Cinahl from 1990 onwards) was a combination of (i) a generic search strategy developed to capture references relating to patient-reported measures of health outcome (including the development and testing of instruments, reviews of such instruments, and conceptual and methodological issues in measurement) and (ii) reviews in the area of mental illness:

(i) ((acceptability or appropriateness or (component* analysis) or comprehensibility or (effect size*) or (factor analys*) or (factor loading*) or (focus group*) or (item selection) or interpretability or (item response theory) or (latent trait theory) or (measurement propert*) or methodol* or (multi attribute) or multiattribute or precision or preference* or proxy or psychometric* or qualitative or (rasch analysis) or reliabilit* or replicability or repeatability or reproducibility or responsiveness or scaling or sensitivity or (standard gamble) or (summary score*) or (time trade off) or usefulness* or (utility estimate) or valid* or valuation or weighting*) and ((COOP or (functional status) or (health index) or (health profile) or (health status) or HRQL or HRQoL or QALY* or QL or QoL or (qualit* of life) or (quality adjusted life year*) or SF-12 or SF-20 or SF?36 or SF-6) or ((disability or function or subjective or utilit* or (well?being)) near2 (index or indices or instrument or instruments or measure or measures or questionnaire* or profile* or scale* or score* or status or survey*)))) or ((bibliograph* or interview* or overview or review) near5 ((COOP or (functional status) or (health index) or (health profile) or (health status) or HRQL or HRQoL or QALY* or QL or QoL or (qualit* of life) or (quality adjusted life year*) or SF-12 or SF-20 or SF?36 or SF-6) or ((disability or function or subjective or utilit* or (well?being)) near2 (index or indices or instrument or instruments or measure or measures or questionnaire* or profile* or scale* or score* or status or survey*))))))

and

(ii) psychiatric or mental or schizophrenic or bipolar or depression or depressed or affective or psychotic or psychoses or psychosis and critique or review or meta?analysis or comparison

2 b) Inclusion/exclusion criteria

Reviews were included if they met two criteria: (a) they provided details of their inclusion criteria and retrieval methodology, or from which it could be inferred that attempts had been made to ensure that relevant material had been captured and (b) they gave the instruments summary ratings for their psychometric properties, including validity and reliability (please see the appendix for further details).

The review was restricted to measures of well-being, functional status and life satisfaction. Measures of psychiatric illness were excluded as were narrowly defined measures of symptoms. Furthermore, instruments specific to particular forms of mental illness such as dementia and schizophrenia were excluded. The review was restricted to quality of life measures for adults, however the literature on children and adolescents has recently been reviewed (Hunter et al, 1996).

Chapter 3: RESULTS

3 a) Search results

The search strategy produced four reviews from peer-reviewed journals which met the inclusion criteria (van Nieuwenhuizen et al, 1997; Atkinson & Zibin, 1996; Lehman, 1996; Andrews et al, 1994). Two published reviews that did not attempt to rate the psychometric properties of instruments were excluded (Dickerson, 1997; Lehman, 1997). The first of these was a review of community functioning measures which, on closer examination was found to contain only one measure (the Multnomah Community Ability Scale, (Barker et al, 1994)) that had not been identified by reviews meeting the inclusion criteria. The other excluded review also failed to rate the instruments and did not identify any new instruments. A further review was identified which considered the application of instruments measuring psychological health and well-being at the population level (Bartlett & Coles, 1998). This review was excluded because only a few instruments were selected for evaluation. Although the principal focus was on reviews published in peer-reviewed journals, instruments reviewed in a compendium of health status measures, which were rated for psychometric properties, were included (McDowell & Newell, 1996).

3 b) Quality of reviews

Table I shows the methodology underpinning the four reviews/one compendium of quality of life measurement in the area of mental health/mental illness. The headings in the table refer to the subject of the review, the databases and years searched, as well as the search strategy used. Inclusion/exclusion criteria are summarised, where given, as are details of the rating schema used to rate psychometric properties.

1. Of the four relevant reviews identified, Andrews et al (1994), was methodologically the most thorough review, as assessed against established criteria (Sackett et al, 1991; Oxman, 1994). This review identified 64 quality of life, multidimensional or functioning measures via a process of electronic database searching, snowballing (searching the reference lists of identified papers for new references), searching compendiums of measures and surveying clinicians and non-governmental institutions to identify any instruments not retrieved by the other methods. The individual instruments were rated using a three-point rating scale, details of which were obtained from the authors. Reviews were rated in terms of responsiveness (which the other reviews failed to do) and individual ratings for the various sub-components of reliability (item, test-retest and inter-rater) and validity (content, construct and criterion) were provided.
2. The second most thorough review identified was Atkinson & Zibin (1996). This identified 27 measures, 14 of which had not been identified by the above. This review identified instruments only via a process of electronic searching. Instruments were again rated using a three-point rating scheme and the method of ratings is described. Each instrument was given one summary rating for validity and another for reliability.
3. Van Nieuwenhuizen (1997) searched electronic databases over a 15-year period, identifying 14 instruments, of which there was reported to be insufficient published information relating to three (consequently excluded). Two of the instruments were not identified by Andrews et al (1994) or Atkinson & Zibin (1996). Instruments

were ranked via a three-point rating scale using similar criteria to the latter. Instruments were rated in terms of their validity (convergent and divergent) and reliability (internal consistency, test-retest and inter-observer).

4. Lehman (1996) identified ten instruments on the basis of a less methodologically well-documented review, again using a three-point rating system. However, the criteria used were not described. This review identified two new instruments. As with Atkinson & Zibin (1996), one summary rating was provided for validity and one for reliability.
5. Finally although they provide no details of their search methodology, we included the relevant chapter ('Psychological Well-being') from McDowell & Newell (1996) for the reason given above. This source identified three instruments that had not been included in the reviews 1-4. Two summary ratings (on a three-point scale) were given for both reliability and validity, one for the thoroughness of the evaluation and the other for the results.

3 c) Analysis of the instruments

Table II lists the individual instruments identified by the various reviews together with information relating to:

1. number of items
2. dimensions (areas of life) covered by the instrument
3. mode of administration
4. use of proxies
5. completion time
6. populations
7. training requirements
8. potential applications

1. Number of items

The reviews included clear information relating to the number of items within instruments for 58 of the 85 instruments. The category into which the largest single number of instruments fell (where information on the number of items was provided) was the category containing instruments with 100 or more items (n=17). The majority of instruments have over 40 items whilst only eight instruments contain fewer than 20 items.

Number of items	Number of instruments
1-9	3
10-19	5
20-29	8
30-39	11
40-49	4
50-59	2
60-69	3
70-79	2
80-89	1
90-99	2
100+	17
Total number of instruments	58

Note: where the instrument consisted of a number of versions for different types of respondents, we considered the patient self-report version.

2. Dimensions

The most widely measured dimensions that are covered by instruments within the reviews include health, self-esteem/well-being, community/productivity, social/love relations and leisure/creativity (Atkinson & Zibin, 1996). Van Nieuwenhuizen (1997) found that a number of domains were included in nearly all instruments, namely health, work, leisure, living situation, friends/social relations, and family relations.

One review found that domains relating to the family, living situation, finances, religion and psychiatric symptoms were less often included by instruments (Atkinson & Zibin, 1996). Another found that the domains of independence, role fulfilment, autonomy, sense of purpose, inner experience and intimate relationships were covered relatively infrequently, despite research showing that people with mental illness stress the importance of self-worth, self-care, autonomy and coping abilities (Neiwenhuizen, 1997). The majority of instruments do not ask patients to rate how important a particular domain is to their lives, although the importance of taking the mentally ill patient's personal goals into account has been recognised (Atkinson & Zibin, 1996).

3. Mode of administration

Where the instrument is based on the responses of patients (as opposed to proxies), the most common mode of administration is via an interview (n=30), usually administered and rated by a clinician. 23 instruments are based on the use of a self-administered questionnaire or rating scale and 14 instruments could be administered via either method. The use of proxies (clinician-rated rating scales not involving interviews with patients or interviews/questionnaires to relatives or other informants), is the exclusive method of obtaining information for 18 instruments whilst it is a possible or contributory method for 12 instruments. Although these latter instruments are not patient-reported measures they are included since in certain types of severe mental illness the patient may not be able to give an accurate account of his/her situation (this issue is returned to later in the report).

4. Completion time

Where available, the time taken for self-administration was used for completion times and full versions were considered in preference to abbreviated versions. The majority of instruments take 15 minutes or more to complete. 26 instruments could be administered in under 15 minutes whilst 29 instruments require more than 30 minutes for administration. Of the latter, five instruments necessitate more than one hour for full administration.

Number of minutes	Number of instruments
5 or less	10
6-14	16
15-29	21
30-44	13
45-59	11
60+	5
Total number of instruments	76

5. Relevant populations

The table below summarises the populations for which instruments are potentially suitable. As can be seen, the overwhelming majority of instruments is potentially applicable, or has been used among, a mentally ill population; significantly fewer are deemed potentially suitable for use among the general population. Many instruments are considered applicable for use among more than one type of population.

Population	Number of instruments
General	10
General health	5
Primary health	7
Physical health	1
Not mental health	9
Mental health	62
Depression	1
Mentally handicapped	1
Older people	1

6. Training

Training is necessary for 19 instruments whilst no training is needed for 22 instruments. Some training is needed to be able to administer two instruments whilst for a further 11 it is unclear whether general training (such as interviewing skills) or training specific to the instrument is necessary.

7. Application

59 instruments are potentially applicable to or have been used in, routine care whilst 51 are applicable to research situations. 14 instruments could potentially be used in

surveys, whereby survey is taken to mean the application of instruments at the population level for, e.g., needs assessment purposes.

Table III summarises the ratings given by the reviews on the basis of the psychometric properties of the instruments.

All the reviews ranked the instruments' psychometric properties on three point scales which, for analysis purposes, were transposed onto a common system (ranking 1,2, and 3 where 1 relates to poor psychometric properties and 3 to the best). No evidence was available on validity for seven instruments. Of those where information on validity was available, the vast majority of instruments were allocated the midway score of 2 (n=42). More instruments were allocated the worst score of 1 (n=9) than the best score of 3 (n=5). There is no evidence on reliability for 10 instruments. Of those where such information was available, as with validity, most instruments fared middling with a score of 2 (n=41) although, unlike with validity, more instruments had the highest ratings for their psychometric properties (n=11) than the lowest (n=3). Information was unavailable on responsiveness for 22 instruments. For those where information was available, 27 were considered to be responsive whilst 36 were not.

3 d) Recommendations for use of specific instruments

Table IV compares the main recommendations for use of specific measures made by the four reviews (note: McDowell & Newell, 1996, do not make any specific recommendations). Most of the reviews make their recommendations on the basis of the extent of psychometric evaluation of the instruments although Andrews et al (op. cit) also incorporate practicality by considering the length and cost of the measure, particularly as regards routine practice.

The Lehman Quality of Life Interview (Lehman, 1983) was recommended by a majority of the reviews (Atkinson & Zibin, 1996; van Nieuwenhuizen et al, 1997; Lehman, 1996). Andrews et al (1994) however commented that it was impractical for use in routine practice.

The Oregon Quality of Life Questionnaire/Scale/Interview (Bigelow et al, 1991) and the Lancashire Quality of Life Profile (Oliver, 1991) were both recommended by two of four reviews (Atkinson & Zibin, 1996; Lehman, 1996) and (Lehman, 1996; van Nieuwenhuizen et al, 1997) respectively. Andrews et al (1994) considered these too to be impractical for use in routine practice.

Other instruments singled out as showing promise and/or for use among specific populations include two disease-specific measures: the Quality of Life in Depression Scale (McKenna & Hunt, 1992) and the Quality of Life Scale (Heinrichs et al, 1984), for depression and schizophrenia respectively. Finally the following were among those also noted positively: the SF-36 (Ware & Sherbourne, 1992); the Quality of Life, Enjoyment and Satisfaction Questionnaire (Endicott et al, 1993); the Quality of Life Index for Mental Health (Becker et al, 1993); the Quality of Life Self-Assessment Inventory (Skantze & Malm, 1994).

Table I Methodology used in the reviews

Author	Subject of review	Databases searched	Years searched	Search strategy	Inclusion criteria	Exclusion criteria	Rating schema
Andrews et al, 1994	Measures dealing with symptoms, functioning, quality of life, burden and satisfaction in the area of mental health. Focuses on serious mental disorders.	Medline searching, snowballing, compendia of measures and survey of clinicians and non-governmental organisations		Keywords given	A manual or published article describing the measure and evidence on psychometric properties was available	Established measures whose psychometric properties have not been reported.	Rankings: weak (+), adequate (++), excellent (+++) following the review below (McDowell & Newell, 1987).
Atkinson & Zibin, 1996	Quality of life measurement for persons with chronic mental illness	Medline and Psychological abstracts	January 1991- January 1996	Search terms given	Abstracts addressing quality of life measurement, research methods and health policy issues within mental health were retrieved.		'Good': internal consistency >0.85 and test-retest reliability coefficients > 0.75. Good face and content validity, construct validity (stable and coherent scales) and >0.60 convergent validity. 'Fair to good': internal consistency coefficients 0.75-0.85 and/or test-retest reliability of <0.75. Either lacking in one or more of the assessed areas of validity or possessing weaker validity coefficients (0.50-0.60). 'Fair': lacking adequate internal consistency and/or test-retest stability and lacking a clear demonstration of validity.
Lehman, 1996	Quality of life among those with severe and persistent mental illnesses (SPMI)		Up to the end of October 1992		Measures emphasising the patient's quality of life, their perspective and feelings. Measures had to at least assess the domain of subjective well-being.	Measures of clinical symptoms, functional status, family burden and client satisfaction with services.	A rating system was used (+, +/- or -) although the meaning/generation of the ratings is not described.
McDowell & Newell, 1996	Measures of psychological well-being				Measures of psychological well-being for which published evidence was available.	Child health, clinical measures, specific forms of impairment, unpublished measures.	Rated first the thoroughness of reliability and validity testing and then the results. Thoroughness ranked from 0 (no reported evidence of reliability or validity) to +++ (all major forms of reliability/validity testing reported). Results ranked from 0 (no numerical results reported) and ? (results uninterpretable) to +++ (excellent reliability/validity).
van Nieuwenhuizen et al, 1997	Quality of life instruments among those with severe mental illness	Excerpta Medica Psychological Abstracts	1980-1995	Key words given for electronic searching		Inadequate published information available on the measure.	Instruments were ranked on a 4-level system depending on whether the evidence met the criteria that reliability coefficients for internal consistency, test-retest reliability and inter-observer reliability should be >0.80 (or >0.70 for Cronbach's alpha). Convergent validity shown by significant correlations between scale scores and other QOL instruments. Divergent validity shown when low correlations exist between scale scores and other non-QOL measures. Rankings were + (data in correspondence with criteria), ± (data are inconsistent, not methodologically sound, only briefly presented), - (data does not correspond with criteria) and (data unavailable)

Table II Description of individual measures

Instrument <i>Type of instrument (where documented)</i>	Review	Number of items	Dimensions	Mode of administration	Completion time	Population	Training required?	Application
Affect Balance Scale	MN	10	Positive and negative psychological reaction to daily life events Yes/no responses or frequency scale. Variable time frame	Self-administered questionnaire	Few mins	General population		Survey
Behaviour and Symptom Identification Scale BASIS <i>Multidimensional</i>	A	32	Relation to self and others, daily living and role functioning, depression and anxiety, impulsive and addictive behaviour, psychosis. 5-point scale	Interview or self-administered questionnaire	20-30 mins	Mental health	No	Clinical Research
Brief Follow-up Rating <i>Quality of life</i>	A	9	Functioning in the community, occupation, housing, friendships and social activities two years previously and at follow-up	Interview (patient or relative-report) /self-administered questionnaire/ relative-completion questionnaire	1-2 mins	Mental health	No	Clinical
Brier Disability Questionnaire <i>Functioning</i>	A	8 (from SF-20)	Vigorous and moderate activities and performance of expected daily activities. 3-point scale. Covers past month	Self-administered questionnaire (or patient-report interview)	5 mins	Primary health	No	Survey
California Well-being Project Client Interview <i>Quality of life</i>	L	151 (self) 76 (family) 77 (professional)	Happiness, health, income, work/achievement, comfort, social life, spiritual life, resources, food, accommodation, sexual life, creativity, basic human freedoms, warmth and intimacy, safety and more. Mostly Likert scaled questions with some open-ended items. Covers subjective well-being	Interview/ self-administered questionnaire/ group self-administration with interviewer to answer questions		Mental health	Trained interviewers if face-to-face	Research
Classification of Intellectual and Other Psychological Impairments <i>Functioning</i>	A		Global intellectual and cognitive functioning, consciousness, sleep and wakefulness, cognitive processes, reality testing, drives, volition, mood and affect, psychomotor functions. 3/5-point scales. Assesses function for past month and during interview	Clinician-administered & rated interview. Ratings on basis of patient and informant information	15-30 mins	Mental health		Research

Instrument <i>Type of instrument (where documented)</i>	Review	Number of items	Dimensions	Mode of administration	Completion time	Population	Training required?	Application
Client Adjustment Rating Scales <i>Functioning</i>	A	33	Motivation, self-concept, self-reliance, affect and mood, vocational-educational readiness, interpersonal relationships, personal maintenance, community resources and family functioning. 5-point scale	Clinician-rated scale	45 mins	Mental health	Yes	Clinical Research
Client Quality of Life Interview/ Scale <i>Quality of life</i>	V L	46 (self) 19 (interviewer ratings)	L: essentials of life (food, clothing, shelter, health, hygiene, money, safety), job training and education, daily activities and recreation, privacy, social supports, social time, self-reliance, peace of mind. Covers subjective well-being, resources & opportunities and functioning. V: leisure and recreation, education and employment, finances, living situation, family, friends and community, ability to cope. Ratings use fixed, ordinal scales	Structured interview plus observations	30 mins	Mental health	Trained lay interviewer	Clinical
Community Adjustment Form <i>Quality of Life</i>	A L	140	A: living situation, time spent in institutions, employment, leisure activities, social relationships, quality of environment, life satisfaction. L: as above and income, family, legal, self-esteem, medical care, agency utilisation. Covers subjective well-being, resources & opportunities and functioning	Clinician-administered interview	45 mins	Mental health	Not specified	Research
Community Disability Scale <i>Functioning</i>	A	32	Personal hygiene, housework, mobility, leisure, physical functioning, money management. 3-point scale	Clinician-administered & rated interview	20 mins	General population		Survey
Comprehensive Quality of Life Scale for Adults <i>Quality of life</i>	A AZ		A: material well-being, health, productivity, intimacy, safety, place in community, emotional well-being AZ: health, finances, living situation, family, social/love relations, leisure/creativity; community productivity, self-esteem well-being. Objective and subjective quality of life measured for each domain	A: self-administered questionnaire AZ: interview	A: 20 mins AZ: 25 mins	A: General population AZ: Used with a mentally handicapped population but not a mental health population	No	A: Research Survey AZ: Clinical

Instrument <i>Type of instrument (where documented)</i>	Review	Number of items	Dimensions	Mode of administration	Completion time	Population	Training required?	Application
Current and Past Psychopathology Scales <i>Multidimensional</i>	A	171	Psychopathology, social and role functioning (current). Psychopathology, personality characteristics, academic, interpersonal and occupational adjustment (past)	Clinician-administered & rated interview. Can also code directly from patient records	1-2 hours	Mental health	Yes	Research
Denver Community Mental Health Questionnaire <i>Multidimensional</i>	A	79 (interview) or 72 (questionnaire)	Psychological distress, isolation (family, friends), productivity, public system dependency, drug and alcohol use, client satisfaction, aggression with friends, legal difficulties. Assesses current functioning	Clinician-administered interview or self-administered questionnaire	45 mins	Mental health	Yes for interview	Clinical Research
Duke-UNC Health Profile <i>Multidimensional</i>	A	63	Symptom status, physical function, emotional function, social function, symptom function. Varying time periods. 3-point scales (except emotional, 5-point)	Clinician-administered & rated interview or self-administered questionnaire	30 mins (interview) 10 mins (self-report)	Primary health	Yes for lay interviewers	Clinical Research
Duke Health Profile <i>Multidimensional</i>	A	17	Health measures: physical, mental, social, general, self-esteem and perceived health. Dysfunction measures: anxiety, depression, pain and disability. Covers current functioning or past week. 3 point scale	Self-administered questionnaire	5 mins	Primary health	No	Clinical Research
Functional Status Questionnaire <i>Functioning</i>	A	34	Physical, psychological, social role, work situation, confinement to bed, restriction of activities, sexual relationships, perceived health status, social activities. 4, 5 and 6-point scales. Covers past month	Self-administered questionnaire	15 mins	Primary health	No	Clinical Research
General Well-being Index <i>Quality of life</i>	A	22	Feelings about anxiety, depressed mood, positive well-being, self-control, health, vitality. 5-point scale. Covers last week	Self-administered questionnaire	5-10 mins	Mental health	No	Clinical Research
General Well-being Schedule	MN	18	Emotional state, anxiety, personal life, sleep, health, daily activities, energy. 6-point or 0-to-10 rating scale. Covers last month	Self-administered questionnaire	10 mins	General population Mental health		Survey
Global Assessment Scale <i>Functioning</i>	A		Lowest levels of functioning over the past week. Range 1-100	Clinician-rated single scale	Variable	Mental health	Yes	Clinical Research

Instrument <i>Type of instrument (where documented)</i>	Review	Number of items	Dimensions	Mode of administration	Completion time	Population	Training required?	Application
Goal Attainment Scale for Psychiatric Inpatients <i>Functioning</i>	A	37	Includes aggression, clothing, co-operation, empathy, frustration, tolerance, posture and suicidal talk. 5-point scale	Clinician-rated scale	35 mins	Mental health	No	Clinical
Goteborg Quality of Life Instrument <i>Quality of life</i>	AZ		Health, psychiatric symptoms, finances, living situation, family, social/love relations, leisure/creativity, self-esteem, well-being . Subjective	Self-administered questionnaire	10 mins	Not mental health population		Clinical
Groningen Social Disabilities Schedule <i>Functioning</i>	A		Roles: self-care, family, kinship, partner, parental, citizen, social, occupational. 4-point scale	Interview (patient/proxy report)	60 mins	Mental health Primary health	Yes	Clinical Research
Health Measurement Questionnaire <i>Quality of life</i>	AZ		Health, social/love relations, community productivity, self-esteem, well-being. Subjective	Interview (patient-report)	15 mins	Mental health		Clinical
Health of the Nation Outcomes Scale <i>Multidimensional</i>	A		Aggression, self-harm, alcohol and drugs, memory/orientation, physical problems, mood disturbance, hallucinations and delusions, other mental, social relationships, social environment (housing and finance), overall severity	Clinician-rated scale	15-30 mins	Mental health	Yes	Clinical
Independent Living Skills Survey <i>Functioning</i>	A	112	Eating habits, grooming skills, domestic activities, food preparation, health maintenance, public transport, leisure, job-seeking. 5-point scale. Covers last month	Clinician or informant-completion questionnaire	25 mins	Mental health	No	Clinical
Katz Adjustment Scales <i>Multidimensional</i>	A	134 (self) 205 (relative)	Psychiatric symptoms (symptom discomfort too in the patient scale), social adjustment, expected social activities, performance and satisfaction of activities	Self-administered questionnaire with separate scales for patient and relative	20-25 mins (patients) 25-40 mins (relatives)	Mental health	No	Clinical Research

Instrument <i>Type of instrument (where documented)</i>	Review	Number of items	Dimensions	Mode of administration	Completion time	Population	Training required?	Application
Lancashire Quality of Life Profile <i>Quality of life</i>	A L V AZ	A: 105 L: 100	A, V & L: general well-being, work/education, leisure/participation, religion, finances, living situation, legal and safety, family relations, social relations, health, self-conflict. AZ: health, finances, living situation, family, social/love relations, leisure/creativity; community productivity, religion, self esteem well-being. Covers subjective well-being, resources & opportunities & functioning. Variable time frame	Clinician-rated interview	A&V: 36.6 mins AZ: 40 mins L: 1 hour	Mental health	A: No V: only brief training	Clinical Research
Lehman Quality of Life Interview <i>A: Multidimensional AZ, V & L: Quality of life</i>	A L V AZ	143	A & V: living situation, family relations, social relations, leisure, work, finances, safety, health, religion. Core version covers daily activities and functioning instead of finances and leisure. AZ: health, psychiatric symptoms, finances, living situation, family, social/love relations, leisure/creativity, community productivity, religion. Objective and subjective. L: core version covers life satisfaction, living situation, daily activities & functioning, family relations, social relations, finances, work and school, legal and safety, health. Covers subjective well-being, resources & opportunities and functioning. Variable time frame	A: Clinician-rated interview L: trained lay interviewers too	A: 20-30 mins (core version), 45 mins (full version) AZ, V & L: 45 minutes	Mental health	Brief training	Research
Levels of Functioning Scale <i>Multidimensional</i>	A	4	Duration of non-hospitalisation for psychiatric disorder, social contacts, useful employment, absence of symptoms. Revised version 9 items: duration of hospitalisation, social relationships, occupational functioning, ability to meet basic needs, fullness of life, overall level of function. 5-point scale	Clinician-rated scale	5 mins	Mental health	Not specified	Research

Instrument <i>Type of instrument (where documented)</i>	Review	Number of items	Dimensions	Mode of administration	Completion time	Population	Training required?	Application
Life-as-a-Whole Index <i>Quality of life</i>	AZ	Uniscale	Health, self-esteem/well-being Subjective	Interview (patient-report)	< 5 mins	Not used with mental health population		
Life Experiences Checklist <i>Quality of life</i>	AZ		Finances, living situation, social/love relations, leisure/creativity. Objective	Self-administered questionnaire and proxy completion	10 mins			Clinical
Life Satisfaction Index <i>Quality of life</i>	AZ MN	20	Health, psychiatric symptoms, finances. Subjective	Self-administered questionnaire	10 mins	Mental health		Clinical
Life Skills Profile <i>Multidimensional</i>	A	39	Self-care, non-turbulence, social contact, communication, responsibility. Assesses general functioning over last 3 months. 4-point scale	Clinician or family member-completion questionnaire (must be familiar with patient)	20-25 mins	Mental health	No	Clinical Research
Major Problem Rating Scale <i>Multidimensional</i>	A	280	Psychiatric and psychosocial problems grouped into functioning. Indicates at intake whether items are problems and these form the basis of the follow-up items. 4-point scale	Self-rated computerised rating scale	40-50 mins (at intake) 10-15 mins (at follow-up)	Mental health	No	Clinical Research
McMaster Health Index Questionnaire <i>Multidimensional</i>	A	59	Assesses performance of physical, social and emotional functioning. Varying time frames	Self-administered questionnaire (can also be interviewer-administered over the telephone)	20 mins	Mental health Primary health	No	Research
Mental Health Inventory <i>Multidimensional</i>	A	38	Anxiety, depression, behavioural/emotional control, general positive affect, emotional ties. Rated on a 6-point scale. Covers last month. Other versions exist.	Self-administered questionnaire	10-15 mins	Mental health General population	No	Survey
Morningside Rehabilitation Status Schedule <i>Multidimensional</i>	A		Dependency, inactivity, social integration/isolation, effects of current symptoms and behaviour. Rated 0 (independent functioning) to 7 (hospital supervision)	Clinician-rated scale	5 mins (if know patient); 30 mins (if not)	Mental health	Not specified	Clinical Research
MRC Needs for Care Assessment <i>Multidimensional</i>	A		Assesses functioning and interventions for problems. Needs then classified on basis of answers. 7-point scale assessing appropriateness of interventions	Clinician-rated scale	Variable	Mental health	Yes (extensive)	Clinical Research

Instrument <i>Type of instrument (where documented)</i>	Review	Number of items	Dimensions	Mode of administration	Completion time	Population	Training required?	Application
MRC Social Role Performance Schedule <i>Functioning</i>	A	8 ratings	Household management, employment, management of money, child care, intimate relationship with spouse or close friend, other relationship, social presentation of self, coping with emergencies. 4-point scale	Clinician-administered & rated interview	Unknown	Mental health	Not specified	Research
Multifaceted Lifestyle Satisfaction Scale <i>Quality of life</i>	AZ		Living situation, social/love relations, leisure/creativity, community/productivity, self-esteem/well-being. Subjective	Interview (patient report)	20 mins	Not used with a mental health population		Clinical
Nottingham Health Profile <i>A: Multidimensional AZ: Quality of life</i>	A AZ	38	A: energy, emotional reaction, social isolation, sleep, pain, physical mobility. AZ: health, social/love relations, self-esteem/ well-being. Objective & subjective. Yes/no answers	Self-administered questionnaire	A: 10 mins AZ: 15 mins	A: General health & population AZ: Mental health	No	A: Clinical & Research AZ: Survey & Screening
Oregon Quality of Life Questionnaire/Interview (Bigelow)* <i>Quality of life</i>	A L V	A: 246 L & V: 263 (self) V: 141 (interview) L: 146 (interview)	A: well-being, stress, basic needs satisfaction, independence, interpersonal adjustment, social support, productivity, leisure time, civic adjustment, use of drugs and alcohol, use of community resources. L: psychological distress, and well-being, stress, basic needs satisfaction, independence, interpersonal interactions, spouse role, social support, work at home, employability, work on the job, meaningful use of time, alcohol and drug use. V: as A and psychological distress, spouse role, work at home and job, employability, self and home, maintenance, finances, physical health, psychiatric medication. Previous week and month evaluated in self-report version; here and now in interview version. Covers subjective well-being, resources & opportunities and functioning. 4-point scale	A: Clinician-administered & rated interview V: self-administered questionnaire also possible	A & V: 45-90 mins L: 45 mins	A: Mental health L: Mental health and general population	A: No L: trained interviewer (not necessarily clinical)	Clinical Research

Instrument <i>Type of instrument (where documented)</i>	Review	Number of items	Dimensions	Mode of administration	Completion time	Population	Training required?	Application
(Oregon) Quality of Life Questionnaire/Interview (Bigelow)* <i>Quality of life</i>	A AZ	263	A: psychological distress, well-being, toleration and coping with stress, basic need satisfaction, independence, interpersonal interaction, spouse role, social support, work (job and home), employability, leisure time, alcohol and drugs. AZ: health, psychological symptoms, finances, living situation, social/love relations, community/productivity, self-esteem/well-being. Objective	Clinician-administered & rated interview	A: 30-40 mins (community) up to 3 hours (mental health) AZ: 35 mins	Mental health	Yes	Clinical Research
Philadelphia Geriatric Center Morale Scale	MN	22	Emotional reaction, energy, social relationships, accommodation, activity, anxiety, health, anger, life satisfaction. Dichtomous responses	Self-administered questionnaire		Community and hospital samples of older people		Clinical Survey
Psychiatric Evaluation Form <i>Multidimensional</i>	A	27	Psychopathological scales and role impairment. Covers a one-week period	Clinician-rated scale (including from patient records) or interview	20-40 mins (interview) 2-4 mins (ratings)	Mental health	Yes	Clinical
Quality of Life Checklist <i>Quality of life</i>	A L V	A & L: 93 N: 95	Housing and household, knowledge and education, relationships, dependency, inner experience, medical care, leisure, work, religion. Covers subjective well-being, resources & opportunities and functioning. Scored as “satisfactory”, “unsatisfactory” or on a 10-point scale. Time frame is previous month	Clinician-administered & rated interview	1 hour (interview) 5-10 mins (rating)	Mental health	A: No L & V: trained interviewer	Clinical
Quality of Life Enjoyment and satisfaction Questionnaire <i>Quality of life</i>	A AZ	93	A: physical health, subjective feelings, leisure activities, social relationships, general activities, work, household duties, school /coursework. AZ: health, living situation, social/love relations, leisure/creativity, community productivity, self-esteem/well-being. Subjective. Uses a 5-point scale, covers last week	Self-administered questionnaire	30 mins	Mental health	No	Clinical Research

* it is unclear whether or not these are two separate measures

Instrument <i>Type of instrument (where documented)</i>	Review	Number of items	Dimensions	Mode of administration	Completion time	Population	Training required?	Application
Quality of Life in Depression Scale <i>Quality of life</i>	A V AZ	34	A: covers how respondents are feeling at the moment, their needs and satisfaction. AZ: health, living situation, family, social/love relations, leisure/creativity, community/productivity, self-esteem/well-being. Subjective. Previous few days evaluated. Answers true/false or yes/no	Self-administered questionnaire	A: 5-10 mins AZ: 15 min V: not specified	Mental health	No	Research Clinical
Quality of Life Index <i>Quality of life</i>	A AZ		A: activity, daily living, health, support, outlook. AZ: health, family, social/love relations, community productivity, self-esteem/well-being. Objective. 3-point scale	A: self/clinician/lay -rated scale AZ: proxy report too	A: 1 min AZ: < 5 mins	A: Primary health AZ: Not used with mental health population	Not specified	A: Clinical & Research AZ: Screening
Quality of Life Index for Mental Health A: <i>multidimensional</i> AZ & V: <i>Quality of life</i>	A V A&Z	113 V: various versions of 42, 68 and 28 items	A and V: satisfaction of objective QOL indicators, occupational activities, psychological well-being, physical health, social relations, economics, activities of daily living, symptoms, goal attainment. AZ: health, psychiatric symptoms, finances, living situation, social/love relations, leisure/creativity, community productivity, self-esteem well-being. Subjective & objective. Variable time period	A: Self-rated questionnaire (or interview) AZ: professional rating V: self-administered questionnaire to patient/clinician/family	A&V: 20-30 mins (self) 10-20 mins (clinicians) AZ: 30 mins (self) 15 mins (profs)	Mental health	No	Clinical Research
Quality of Life Interview Schedule <i>Quality of life</i>	AZ	87	Health, psychiatric symptoms, finances, living situation, social/love relations, leisure/creativity, self-esteem/well-being. Objective	Interview (patient-report)	35 mins	Mental health		Clinical
Quality of Life Inventory <i>Quality of life</i>	AZ		Health, finances, living situation, family, social/love relations, leisure/creativity, community productivity, religion, self-esteem/well-being. Subjective	Self-administered questionnaire	10 mins	Mental health		Clinical
Quality of Life Questionnaire <i>Quality of life</i>	AZ		Living situation, social/love relations, leisure/creativity, community productivity, self-esteem/well-being. Subjective & objective	Professional-administered & rated interview	20 mins	Not used with a mental health population		Clinical

Instrument <i>Type of instrument (where documented)</i>	Review	Number of items	Dimensions	Mode of administration	Completion time	Population	Training required?	Application
Quality of Life Scale <i>Quality of life</i>	A L V AZ	21	A and V: intrapsychic foundations, interpersonal relations, instrumental role, common objects and activities. AZ: psychological symptoms, social/love relations, leisure/creativity, community/productivity, self-esteem/well-being. L: commonplace activities, occupational role, work, possession of objects, interpersonal relations, sense of purpose, motivation, curiosity, anhedonia, aimless inactivity, empathy, emotional interaction. V: focuses specifically on deficit symptoms. Subjective and objective. Covers subjective well-being, resources & opportunities, and functioning. Covers past month, rated on a 7-point scale	Clinician-administered & rated interview	A, V & L: 45 mins AZ: 40 mins	Mental health	A: No L: trained clinician	Clinical
Quality of Life Self-Assessment Inventory <i>Quality of life</i>	AZ V	100	AZ: health, finances, living situation, social/love relations, leisure/creativity, community productivity, religion, self-esteem/well-being. V: housing, environment, knowledge and education, contacts, dependence, inner experience, mental health, physical health, leisure, work, religion. Subjective. Evaluated time period is now. Ratings are satisfactory or unsatisfactory	Interview (self-report)	AZ: 30 mins V: 10 mins (self) + 40-50 mins (interview)	Mental health	Not specified	Clinical
Quality of Life Systemic Inventory <i>Quality of life</i>	AZ		Health, family, leisure/creativity, community productivity Subjective	Interview (patient-report) using visual representations	45 mins	Not used with a mental health population		Clinical

Instrument <i>Type of instrument (where documented)</i>	Review	Number of items	Dimensions	Mode of administration	Completion time	Population	Training required?	Application
Quality of Well-being Scale <i>A: multidimensional</i> <i>AZ: quality of life</i>	A AZ	≥18	A: 3 dimensions of mobility, physical activity and social activity. Function level assigned. AZ: health, psychiatric symptoms, social/love relations, community productivity. Objective. Covers previous 8 days	A: Clinician-administered & rated interview AZ: Interview & professional/ proxy ratings	10-15 mins	A: General health AZ: Not used with a mental health population		A: Research AZ: Survey
Questionnaire for Life Satisfaction Assessment <i>Quality of life</i>	V	32	Friends, leisure, health, finances, work, living situation, family and children, partner and sexuality 4-point scale, dissatisfied-very satisfied over the previous four weeks	Self-administered questionnaire	10 mins			
Rating of Social Disability <i>Multidimensional</i>	A	208	Physical, behavioural and social disabilities. Item is marked if disability is present. Assessment made 6 months after onset of disability	Clinician-rated scale	20-25 mins	Mental health General health	Yes	Clinical Research
Rehabilitation Evaluation <i>Functioning</i>	A	23	Deviant behaviour, social and everyday behaviour. 3-point scales. Need 1 week observation period	Clinician-rated scale	10 mins	Mental health	Yes	Clinical Research
Resource Associated Functional Level (Revised) <i>Multidimensional</i>	A	15	Life skills (antisocial, self-care, withdrawal, compliance, bizarre), illness, accommodation, employment, social relationships. Functioning rated true/false	Clinician-rated scale	5 mins		Yes	Research
Resource Associated Functional Level Scale <i>Functioning</i>	A		7- point scale representing functioning and independence from mental health system	Clinician-rated scale	1 min	Mental health	Yes	Research
Role Functioning Scale <i>Functioning</i>	A	4 single scales	Working, independent living and self-care, immediate social network relationships, extended social network relationship. 7-point scale	Clinician-rated scale	Few mins	Mental health	Yes	Research

Instrument <i>Type of instrument (where documented)</i>	Review	Number of items	Dimensions	Mode of administration	Completion time	Population	Training required?	Application
Satisfaction with Life Domains Scale <i>Quality of life</i>	A L V	15	A: house, neighbourhood, food, clothing, health, people living with, friends, family, relationships, job, leisure, community activities and facilities, finances, current home compared to hospital. V: as above and work/day programming. L: as above and fun. Covers subjective well-being. 7-point delighted-terrible faces scale. Unspecified time frame	A: self-administered rating scale L: patient-rated scale administered by an interviewer	Less than 10 mins	Mental health	A: no L: trained interviewer V: not known	Clinical
Satisfaction with Life Scale <i>Quality of life</i>	AZ		Self-esteem/ well-being. Subjective	Self-administered questionnaire	5 mins	Mental health		Clinical
Schedule for the Evaluation of Individual Quality of Life <i>Quality of life</i>	AZ		User-defined. Subjective	Interview (patient-report)	40 mins	Not used with a mental health population		Clinical
SF-36 A: <i>Multidimensional</i> AZ: <i>Quality of life</i>	A AZ	36	A: physical functioning, physical and emotional role limitations, bodily pain, mental health, social functioning, vitality, general health perceptions, reported health transition. AZ: health, family, social/love relations, leisure/creativity, community productivity, self-esteem/well-being. Objective and subjective. 4 week or 1 week timeframe	A: Self- (or proxy) administered questionnaire (or patient-report interview) AZ: interview (patient report)	A: 5-10 mins AZ: 20 mins	Depression General health	Minimal training required	Clinical Research Survey
Sickness Impact Profile A: <i>multidimensional</i> AZ: <i>Quality of life</i>	A AZ	136	A: sleep and rest, eating, work, home management, recreation and pastimes, ambulation, mobility, body care, movement, social interaction, alertness, emotional behaviour, communication. Respondents endorse statements about functioning for the day, related to health. AZ: health, living situation, social/love relations, leisure/creativity, community productivity, self-esteem/well-being. Subjective & objective	Self-administered questionnaire (also possible as interview)	A: 20 mins (self) 20-30 mins (interview) AZ: 30 mins	Mental health Physical health		Clinical Research Survey

Instrument <i>Type of instrument (where documented)</i>	Review	Number of items	Dimensions	Mode of administration	Completion time	Populations	Training required?	Application
Single-item Indicators of Well-being <i>Quality of life</i>	A MN	Several scales (1 item per scale)	7-point delighted-terrible faces scale and 9-point ladder scale (best life to worst life)	Self-rated scale	A: 5 mins per item MN: < 1 min	General population	Yes	Survey
SmithKline Beecham Quality of Life Scale <i>Quality of life</i>	AZ	28	Health, finances, social/love relations, leisure/creativity; community productivity; religion. Subjective. Rating via “self now”, “ideal self”, “sick self”	Self-administered questionnaire	45 mins	Mental health		Clinical
Social Adjustment Scale <i>Functioning</i>	A	42	Work, social and leisure, relationship with extended family, parental/marital role, economic independence. 5/ 7-point scale. Covers 2 months	Trained carer-administered interview	50 mins	Mental health General population	Yes	Research
Social Adjustment Scale for Schizophrenics <i>Functioning</i>	A	52	Work, household, external family, social/leisure activities, personal well-being. 5-point scale. Covers past 2 months	Clinician-administered & rated interview	15-60 mins	Mental health	Not specified	Research
Social Adjustment Scale Self-report <i>Functioning</i>	A	42	Work, social and leisure, relationships with family, marital, parental, family unit. 5-point scale. Covers previous 2 weeks	Self-administered questionnaire (or relative-completed)	20 mins	Mental health	No	Clinical Research
Social Behaviour Assessment Schedule <i>Multidimensional</i>	A	239	Demographics, patient’s behaviour, patient’s social performance, adverse effects on others, concurrent event, support to informant/informant’s housing situation. 3/5-point scales.	Clinician-administered & rated interview with informant	45-75 mins	Mental health	Yes	Research
Social Behaviour Schedule <i>Functioning</i>	A	21 ratings	Communication, social mixing, panic attacks and phobias, acting out bizarre ideas, depression, personal appearance and hygiene. Additional items cover occupation, leisure, restrictions on activity, unrealistic aims, reason for being in setting, most difficult problem, handicaps, work quality, attitudes. 4/5-point scales, covers past month	Clinician-administered & rated interview with staff member about the patient	20 mins	Mental health	Not specified	Clinical Research
Social Dysfunction Rating Scale <i>Quality of life</i>	A	21	Self-system, interpersonal system, performance system (personal satisfaction and self-fulfilment). Rated not present (1) to severe (6)	Clinician-administered & rated interview	15 mins	Mental health	No	Research

Instrument <i>Type of instrument (where documented)</i>	Review	Number of items	Dimensions	Mode of administration	Completion time	Population	Training required?	Application
Social Functioning Scale <i>Functioning</i>	A	78	Social engagement/withdrawal, interpersonal, pro-social, recreation, independence-competence, independence-performance, employment. 4/5-point scales	Self-administered or relative-completion questionnaires	20 mins	Mental health	No	Clinical
Social Functioning Schedule <i>Multidimensional</i>	A	121	Employment, household, money, self-care, marital relationship, child-care, parent-child relationships, extra-marital relationships, social contacts, hobbies, spare time activities. Analogue scale used. Covers past month	Clinician-administered and rated interview	10-20 mins	Mental health	Interviewing skills needed	Clinical Research
Social Maladjustment Schedule <i>Multidimensional</i>	A	26 pages & 42 ratings	Housing, occupation, social role, economic situation, leisure and social activities, family relationships, marriage. 4-point scale	Clinician-administered & rated interview. Relatives' opinions can be incorporated	45 mins	Mental health General population	Yes	Clinical Survey
Standardised Social Schedule <i>Quality of life</i>	L	48	Housing, occupation/social role, economic situation, leisure/social activities, family and domestic relationships, marital situation. Covers subjective well-being, resources and opportunities and functioning	Interviewer-administered & rated interview	45 mins	Mental health	Trained interviewer	
Structured and Scaled Interview to Assess Maladjustment <i>Functioning</i>	A	60	Work, social, family, marriage, sex. 11-point scale	Clinician-administered & rated interview	30 mins	Mental health	Yes	Clinical Research
Subjective Well-being Inventory <i>Quality of life</i>	A	40	Satisfaction with life on a 3-point scale	Self-administered questionnaire or interview	10 mins	General health		Clinical Research

Instrument <i>Type of instrument (where documented)</i>	Review	Number of items	Dimensions	Mode of administration	Completion time	Population	Training required?	Application
Vermont Community Questionnaire <i>Multidimensional</i>	A	VCQ-C = 233 VQL-L = 156	VCQ-C (cross-sectional): residence, work, finances, intimate relationships, family information, social support, weekly activities, self-care, use of treatment services, contact with criminal justice system, community involvement, satisfaction, environmental stressors, competence, psychopathology. VQL-L (longitudinal): medications, residence, hospitalization, work, income, important personal relationships, deaths of important people, other life events, use of community support systems, physical health. Covers past 20 years	Clinician-administered & rated interview	VQC-C = 75-90 mins VQL-L = 75 mins	Mental health	Not specified	Research
WHO Psychiatric Disability Schedule <i>Functioning</i>	A		Social dysfunction and adjustment. 6-point scale. Covers past month	Clinician-administered & rated interview with relative as main informant but patient also interviewed	30 mins	Mental health	Yes (extensive)	Clinical Research
WHO Quality of Life assessment <i>Quality of life</i>	A	278	Physical, psychological, levels of independence, social relationships, environment, spiritual. 5-point scale. Covers past 2 weeks	Self-administered questionnaire		Being evaluated	No	Clinical Research Survey

KEY:	
A	Andrews et al (1994)
AZ	Atkinson & Zibin (1996)
L	Lehman (1996)
MN	McDowell & Newell (1996)
V	van Nieuwenhuizen et al (1997)

Table III Review of instruments' psychometric properties

Instrument	Validity	Reliability	Responsiveness
Affect Balance Scale	Thoroughness ++; Results ++ ^{MD}	Thoroughness ++; Results ++ ^{MD}	
Behaviour and Symptom Identification Scale	Content-adequate, construct-adequate, criterion-adequate ^A	Item-adequate, test-retest adequate ^A	Yes ^A
Brief Follow-up Rating	Construct-weak ^A	No evidence ^A	No ^A
Brier Disability Questionnaire	Construct-adequate ^A	Item-adequate ^A	No ^A
California Well-being Project Client Interview	NA ^L	NA ^L	
Classification of Intellectual and Other Psychological Impairments	Construct-adequate ^A	Item-adequate, inter-rater adequate ^A	No ^A
Client Adjustment Rating Scales	Criterion-weak ^A	Test-retest adequate, inter-rater adequate ^A	No ^A
Client Quality of Life Interview	^L Convergent Validity..., Divergent Validity ... ^V	NA ^L Internal Consistency.. , Test-retest.., Interobserver reliability... ^V	
Community Adjustment Form	No evidence ^A NA ^L	No evidence ^A NA ^L	No ^A
Community Disability Scale	Construct-adequate, criterion-adequate ^A	Item-excellent ^A	No ^A
Comprehensive Quality of Life Scale for Adults	Construct-adequate ^A Fair ^{AZ}	Item-adequate; Test-retest-adequate ^A Fair ^{AZ}	No ^A
Current and Past Psychopathology Scales	Construct-weak, criterion-adequate ^A	Item-weak, inter-rater adequate ^A	No ^A
Denver Community Mental Health Questionnaire	Construct-adequate, criterion-weak ^A	Item-adequate, inter-rater adequate ^A	Yes ^A
Duke-UNC Health Profile	Content-weak, construct-adequate ^A	Item-adequate, test-retest adequate ^A	No ^A
Duke Health Profile	Content-weak, construct-adequate ^A	Item-adequate, test-retest adequate ^A	No ^A
Functional Status Questionnaire	Content-adequate, construct-weak ^A	Item-adequate ^A	No ^A
General Well-Being Index	Construct-adequate; Content-adequate ^A	Item-excellent; Test-retest-excellent ^A	Yes ^A

General Well-being Schedule	Thoroughness +++; Results ++ ^{MN}	Thoroughness +++ ; Results +++ ^{MN}	
Global Assessment Scale	Construct-adequate, criterion-adequate ^A	Inter-rater adequate ^A	Yes ^A
Goal Attainment Scale for Psychiatric Inpatients	Content-adequate ^A	Inter-rater adequate ^A	No ^A
Goteborg Quality of Life Instrument	Fair to good ^{AZ}	Fair to good ^{AZ}	
Groningen Social Disabilities Schedule	Construct-adequate ^A	Inter-rater adequate ^A	Yes ^A
Health Measurement Questionnaire	Fair to good ^{AZ}	Fair to good ^{AZ}	
Health of the Nation Outcome Scales	nil ^A	Item-adequate, inter-rater adequate ^A	Yes ^A
Independent Living Skills Survey	Content-adequate, criterion-adequate ^A	Item-adequate ^A	No ^A
Katz Adjustment Scales	Construct-adequate, criterion adequate ^A	Item-adequate ^A	No ^A
Lancashire Quality of Life Profile	Content-adequate; Construct-weak; Criterion-adequate ^A + ^L Convergent Validity +, Divergent Validity ... ^V Fair to good ^{AZ}	Item-adequate; Test-retest-weak; Inter-rater-weak ^A + ^L Internal Consistency +; Test-retest ± ; Interobserver reliability - ^V Fair to good ^{AZ}	Yes ^A
Lehman Quality of Life Interview	Content-adequate, construct-adequate, criterion-adequate ^A + ^L Convergent Validity +, Divergent Validity + ^V Good ^{AZ}	Item-adequate, test-retest adequate ^A + ^L Internal Consistency +, Test-retest ±, Interobserver.. ^V Fair to good ^{AZ}	No ^A
Levels of Functioning Scale	Construct-weak ^A	Item-weak, inter-rater excellent ^A	Yes ^A
Life-as-a-Whole Index	Fair ^{AZ}	Fair ^{AZ}	
Life Experiences Checklist	Fair ^{AZ}	Fair ^{AZ}	
Life Satisfaction Index	Fair to good ^{AZ} Thoroughness +++; Results ++ ^{MN}	Fair to good ^{AZ} Thoroughness ++; Results ++ ^{MN}	
Life Skills Profile	Content-adequate, construct-adequate, criterion-adequate ^A	Item-adequate, test-retest excellent, inter-rater adequate ^A	Yes ^A
Major Problem Rating Scale	Content-adequate, construct-adequate ^A	Item-adequate, test-retest adequate ^A	Yes ^A
McMaster Health Index Questionnaire	Content-adequate, construct-adequate, criterion-adequate ^A	Test-retest adequate ^A	Yes ^A

Mental Health Inventory	Content-adequate, construct-adequate ^A	Item-excellent, test-retest adequate ^A	Yes ^A
Morningside Rehabilitation Status Schedule	Construct-adequate, criterion-adequate ^A	Inter-rater adequate ^A	Yes ^A
MRC Needs for Care Assessment	Content-adequate, criterion-adequate ^A	Inter-rater excellent ^A	No ^A
MRC Social Role Performance Schedule	Construct-adequate ^A	Inter-rater excellent ^A	No ^A
Multifaceted Lifestyle Satisfaction Scale	Fair to good ^{AZ}	Fair to good ^{AZ}	
Nottingham Health Profile	Content-excellent, construct-adequate, criterion-adequate ^A Fair ^{AZ}	Test-retest adequate ^A Fair to good ^{AZ}	Yes ^A
Oregon Quality of Life Questionnaire	Construct-adequate; Criterion-adequate ^A + ^L Convergent Validity Divergent Validity ... ^V	Item-adequate; Test-retest-weak ^A + ^L Internal Consistency ±; Test-retest ±; Interobserver reliability ± ^V	Yes ^A
(Oregon) Quality of Life Questionnaire	Construct-adequate, criterion-adequate ^A Good ^{AZ}	Item-adequate ^A Fair to good ^{AZ}	Yes ^A
Philadelphia Geriatric Center Morale Scale	Thoroughness ++ ; Results ++ ^{MN}	Thoroughness ++ ; Results ++ ^{MN}	
Psychiatric Evaluation Form	Construct-adequate, criterion-adequate ^A	Item-adequate, inter-rater adequate ^A	Yes ^A
Quality of Life Checklist	Nil ^A NA ^L Convergent Validity ..., Divergent Validity ... ^V	Nil ^A NA ^L Internal Consistency ..., Test-retest ..., Interobserver reliability ... ^V	No ^A
Quality of Life Enjoyment and Satisfaction Questionnaire	Construct-adequate, criterion-adequate ^A Good ^{AZ}	Item-excellent, Test-retest adequate ^A Good ^{AZ}	Yes ^A
Quality of Life in Depression Scale	Content-adequate, construct-adequate, criterion-forthcoming ^A Convergent Validity +, Divergent Validity - ^V Good ^{AZ}	Item-excellent, test-retest excellent ^A Internal Consistency +, Test-retest ±, Interobserver reliability ... ^V Good ^{AZ}	No ^A
Quality of Life Index	Content-excellent, construct-adequate ^A Fair to good ^{AZ}	Inter-rater adequate ^A Fair to good ^{AZ}	No ^A
Quality of Life Index for Mental Health	Content-adequate, construct-weak, criterion-adequate ^A Convergent Validity ±, Divergent Validity ... ^V Fair to Good ^{AZ}	Test-retest adequate ^A Internal Consistency ..., Test-retest +, Interobserver... ^V Good ^{AZ}	Forthcoming ^A
Quality of Life Interview Schedule	Fair to good ^{AZ}	Fair to good ^{AZ}	
Quality of Life Inventory	Good ^{AZ}	Fair to good ^{AZ}	

Quality of Life Questionnaire	Fair to good ^{AZ}	Fair to good ^{AZ}	
Quality of Life Scale	Construct-adequate ^A NA ^L Convergent Validity \pm , Divergent Validity ... ^V Fair to good ^{AZ}	Inter-rater adequate ^A + ^L Internal Consistency +, Test-retest +, Interobserver reliability + ^V Good ^{AZ}	No ^A
Quality of Life Self-Assessment Inventory	Convergent validity \pm , divergent validity ... ^L Fair to good ^{AZ}	Internal consistency..., test-retest +, interobserver reliability.. ^L Fair to good ^{AZ}	
Quality of Life Systemic Inventory	Fair to good ^{AZ}	Fair to good ^{AZ}	
Quality of Well-being Scale	Content-adequate, construct-adequate ^A Good ^{AZ}	Test-retest excellent, inter-rater adequate ^A Good ^{AZ}	No ^A
Questionnaire for Life Satisfaction Assessment (FLZ)	Convergent validity..., divergent validity... ^V	Internal consistency..., test-retest..., interobserver reliability.. ^V	
Rating of Social Disability	Criterion-adequate ^A	Item-adequate, inter-rater excellent ^A	No ^A
Rehabilitation Evaluation	Content-adequate, construct-adequate ^A	Item-adequate, test-retest adequate, inter-rater adequate ^A	Yes ^A
Resource Associated Functional level (Revised)	Not yet assessed ^A	Not yet assessed ^A	No ^A
Resource Associated Functional level Scale	Construct-adequate ^A	Inter-rater adequate ^A	Yes ^A
Role Functioning Scale	Construct-adequate, criterion-adequate ^A	Item-excellent, test-retest adequate, inter-rater adequate ^A	No ^A
Satisfaction with Life Domains Scale	Criterion-weak ^A +/- ^L Convergent Validity +, Divergent Validity ... ^V	Nil ^A NA ^L Internal Consistency +, Test-retest ..., Interobserver reliability ... ^V	No ^A
Satisfaction with Life Scale	Good ^{AZ}	Good ^{AZ}	
Schedule for the Evaluation of Individual Quality of Life	Good ^{AZ}	Good ^{AZ}	
SF-36	Content-excellent, construct-excellent, criterion-adequate ^A Good ^{AZ}	Item-adequate, test-retest adequate ^A Good ^{AZ}	Yes ^A
Sickness Impact Profile	Content-excellent, construct-adequate, criterion-adequate ^A Good ^{AZ}	Item-excellent, test-retest adequate ^A Good ^{AZ}	Yes ^A

Single Item Indicators of Well-being	Construct-adequate ^A Thoroughness + ; Results ++ ^{MN}	Test-retest adequate ^A Thoroughness ++ ; Results ++ ^{MN}	No ^A
SmithKline Beecham Quality of Life Scale	Good ^{AZ}	Good ^{AZ}	
Social Adjustment Scale	Construct-adequate ^A	Inter-rater excellent ^A	No ^A
Social Adjustment Scale for Schizophrenics	Construct-weak ^A	Nil ^A	No ^A
Social Adjustment Scale Self-report	Construct-adequate, criterion-adequate ^A	Item-adequate, test-retest adequate ^A	Yes ^A
Social Behaviour Assessment Schedule	Content-adequate ^A	Item-adequate, inter-rater adequate ^A	Yes ^A
Social Behaviour Schedule	Construct-adequate, criterion-adequate ^A	Item-adequate, test-retest adequate, inter-rater adequate ^A	Yes ^A
Social Dysfunction Rating Scale	Construct-adequate, criterion-adequate ^A	Inter-rater adequate ^A	No ^A
Social Functioning Scale	Construct-adequate, criterion-adequate ^A	Item-adequate, inter-rater excellent ^A	Yes ^A
Social Functioning Schedule	Construct-adequate ^A	Inter-rater weak ^A	Yes ^A
Social Maladjustment Schedule	Construct-weak ^A	Inter-rater adequate ^A	No ^A
Standardised Social Schedule	+ ^L	+ ^L	
Structured and Scaled Interview to Assess Maladjustment	Content-adequate, construct-adequate, criterion-adequate ^A	Item-adequate, inter-rater adequate ^A	No ^A
Subjective Well-being Inventory	Construct-weak ^A	Nil ^A	No ^A
Vermont Community Questionnaire	Content-adequate ^A	Inter-rater adequate ^A	No ^A
WHO Psychiatric Disability Schedule	Construct-adequate ^A	Item-adequate, inter-rater excellent ^A	No ^A
WHO Quality of Life Assessment	Under testing ^A	Under testing ^A	No ^A

KEY:	AUTHOR:	RATING:
^A	Andrews et al (1994)	weak, adequate, excellent, nil
^{AZ}	Atkinson & Zibin (1996)	fair, fair to good, good
^L	Lehman (1996)	+, +/-, -, NA
^{MN}	McDowell & Newell (1996)	For reliability and validity, thoroughness score (0, +, ++ or +++) and results score (0, ?, +, ++ or +++)
^V	van Nieuwenhuizen et al (1997)	+ data in correspondence with criteria ± data are inconsistent, not methodologically sound, only briefly presented, - data does not correspond with criteria, data unavailable

Table IV Main recommendations for specific instruments

Author	Andrews et al (1994)	Atkinson & Zibin (1996)	Lehman (1996)	van Nieuwenhuizen et al (1997)	Measure reviewed by
Basis of recommendations	<i>Length, cost, psychometric properties, use in mental health population and potential for use in routine clinical practice</i>	<i>Extent of use and validation among a mixed mental health population</i>	<i>Comprehensiveness and level of psychometric evaluation</i>	(a) <i>Psychometric features & feasibility</i> (b) <i>Individual's needs and satisfaction</i> (c) <i>Ratings from several perspectives</i>	
Behaviour and Symptom Identification Scale (BASIS) (Eisen et al, 1994)	√				A
Health Measurement Questionnaire (Gater et al, 1995)		√			AZ
Health of the Nation Outcome Scales (HoNOS) (Wing, 1994)	√				A
Lancashire Quality of Life Profile (Oliver, 1992)	Impractical for routine use due to length		√	√(a)	A, AZ, L, V
Lehman Quality of Life Interview (Lehman, 1983)	Impractical for routine use due to length	√	√	√(a)	A, AZ, L, V
Life Skills Profile (Rosen et al, 1989)	Good psychometric properties but expensive				A
Mental Health Inventory (MHI) (Veit & Ware, 1983)	√				A
(Oregon) Quality of Life Questionnaire/Interview (Bigelow et al, 1991)	Impractical for routine use due to length	√	√		A, AZ, L, V
Quality of Life in Depression Scale (McKenna & Hunt, 1992)	Condition-specific	Shows promise with specific populations			A, AZ, V
Quality of Life Enjoyment and Satisfaction Questionnaire (Endicott et al, 1993)	√(although length may limit use)	Shows promise with specific populations			A, AZ
Quality of Life Index for Mental Health (Becker et al, 1993)	Impractical for routine use due to length	Shows promise		√(c)	A, AZ, V
Quality of Life Interview Schedule (Holcomb et al, 1993)		√			AZ

Author	Andrews et al (1994)	Atkinson & Zibin (1996)	Lehman (1996)	van Nieuwenhuizen et al (1997)	Measure reviewed by
Basis of recommendations	<i>Length, cost, psychometric properties, use in mental health population and potential for use in routine clinical practice</i>	<i>Extent of use and validation among a mixed mental health population</i>	<i>Comprehensiveness and level of psychometric evaluation</i>	(a) <i>Psychometric features & feasibility</i> (b) <i>Individual's needs and satisfaction</i> (c) <i>Ratings from several perspectives</i>	
Quality of Life Inventory (Frisch et al, 1992)		√			AZ
Quality of Life Scale (Heinrichs et al, 1994)	Impractical for routine use due to length		√ (for schizophrenia)	√ (for schizophrenia)	A, AZ, L, V
Quality of Life Self-Assessment Inventory (Skantze et al, 1992)		Shows promise		√ (b)	AZ, V
Role Functioning Scale (Goodman et al, 1993)	√				A
Satisfaction with Life Domains Scale (Baker & Intagliata, 1982)	Little psychometric evidence available		√ (but only measures life satisfaction)		A, L, V
SF-36 (Ware & Sherbourne, 1992)	√	Shows promise with specific populations			A, AZ
Sickness Impact Profile (Bergner et al, 1976)	Impractical for routine use due to length	Shows promise with specific populations			A, AZ

Chapter 4: DISCUSSION

4 a) Conceptual basis of quality of life

Most instruments currently in use in medicine today were developed from either the generic quality of life framework or the health-related quality of life framework (Gladis et al, 1999). The generic framework, which includes measures of well-being and patient satisfaction, mainly concerns subjective perceptions and seeks to relate aspirations with reality. It includes a variety of domains, not just those related to health. The health-related quality of life framework, as its name suggests, is concerned with documenting the burden of illnesses and their interventions and can either be generic (used across disease groupings) or disease-specific. Although some health-related quality of life instruments do have summary questions concerning overall life satisfaction, they mainly concentrate on symptoms and functioning, impairment and disability (ibid).

In the field of mental health, it is suggested that two dominant and competing models in assessing quality of life in mental health have emerged (ibid). These are the Life Satisfaction Model (arising from a generic framework which considers a very broad aspect of quality of life, equating quality of life with satisfaction in the various life domains) and the Three Component Model (i.e. satisfaction, functioning and socio-economic conditions, arising from the health-related framework). The Satisfaction Model takes the more patient-centred subjective approach and is possibly exemplified by Frisch's the Quality of Life Inventory (Frisch et al, 1992). However even instruments with this conceptual underpinning typically include some objective and functioning items such as Lehman's Quality of Life Interview (Lehman, 1983). The Three Component Model derives from the adaptive functioning model which takes as its basic premise the concept that life satisfaction depends on one's ability to master the external environment (Fabian, 1990). This model adds adaptation to the objective and subjective indicators in an attempt to capture that aspect of quality of life most amenable to the influence of mental health professionals, namely adapting to the given socio-economic conditions. Although the two models seem distinct, evidence exists suggesting a common construct is being measured with some overlap between self-reported satisfaction and measures of adaptive functioning (ibid).

There seems to be a general consensus in much of the literature that empirical and theoretical justification exists for a broad definition of quality of life, encompassing both objective and subjective factors assessing satisfaction, functioning and material circumstances (Mercier, 1994; Oliver et al, 1996; Gladis et al, 1999), although the appropriate quality of life framework depends on the outcome one is focussing on. It would seem that many quality of life instruments in use at present are based on a broad definition of quality of life and try to measure at least two of the three components of satisfaction, functioning and material conditions (Gladis et al, 1999; Stedman, 1996). However for others, it would seem quality of life research is still dominated by the subjective approach to the exclusion of contextual factors (Katschnig, 1997). In some non-severe and persistent populations, such as depression and anxiety, measures tend to be disease-specific and focus primarily on functioning (Lehman, 1997; Gladis et al, 1999). Given the differences between mental and somatic illness, doubts have been expressed in the literature concerning whether more narrowly defined health-related quality of life measures, with their emphasis on

physical disability and symptoms, are appropriate to people with mental illness (ibid). Positive results have however been reported from use of the SF-36 (Ware & Sherbourne, 1992).

4 b) Subjective or objective evaluation

The issue of subjective evaluation (the patient assesses or evaluates his/her feelings about aspects of life) versus objective evaluation (external, observable variables form the basis of quality of life assessment) is a major theme in the literature relating to the measurement of quality of life among mentally ill patients.

It has also been shown that it is the patients' subjective well-being rather than objective medical conditions that determines their behaviour, compliance and evaluation of treatment. Evidence among the general population suggests that, even after controlling for objective measures, self-perceived health correlates with mortality and recovery rates (Jenkins, 1992). There are some objective factors that are associated with quality of life perceptions among the mentally ill, notably finances, health status, employment status and frequency of interpersonal relationships (Fabian, 1990; Mercier, 1994). However evidence suggests that patients provide information about outcomes that cannot be obtained through looking at these factors or symptoms and functioning alone (Repper & Brooker, 1998).

However there are concerns in the mental health literature concerning the use of subjective evaluation as the basis for quality of life assessment. This centres around two related, but distinct, issues: first, is someone with a mental illness able, in terms of possessing the necessary cognitive functioning, to evaluate his or her own feelings and perceptions and second, would the resulting evaluations be in accordance with what social norms would predict?

The problems that might be inherent in the mentally ill's assessment of their own condition have been summarised as: the affective fallacy; the cognitive fallacy; and the reality distortion fallacy (Katschnig, 1997). The former occurs when patients view their situation more negatively than they appear to an independent observer (occurring particularly among depressed patients) or when patients rate their situation more favourably than independent observation would suggest (as with some manic patients). Cognitive fallacies occur when patients do not possess the intellectual functioning to assess their lives, as is the case among patients with dementia or mental retardation. Lastly, reality distortion concerns patients whose hallucinations or delusions distort their perceptions of their selves and their surroundings.

Lehman (Huxley, 1986; Fabian 1990) investigated whether the patient's clinical state, as shown by their symptomatology, influences patients' quality of life assessments. His work found that patients' clinical state did not interfere with their ability to provide subjective assessments of quality of life, with the exception of satisfaction with health (Huxley, 1986) and only certain types of psychopathology (depression and anxiety) correlated only moderately with subjective quality of life indicators (Fabian, 1990). Others have confirmed that mentally ill patients in a variety of settings are able to voice preferences and feelings (Huxley, 1986) and that quality of life research among the mentally ill has not demonstrated substantially different results when compared with the general population (Fabian et al, 1990). Subjective indicators most

closely associated with overall quality of life include satisfaction with health, social relationships, occupation (work or principal occupation or leisure) and finances (Mercier, 1994).

However clinical populations may show more positive perceptions of quality of life than the general population, indicating they may be satisfied with less (ibid) or it may indicate their accommodation to adversity (Lehman, 1996). This has been termed “standard drift fallacy”: “if one cannot possibly achieve one’s aims, these aims are changed” (Katschnig, 1997). Since people with mental illness may be satisfied with what most in society would consider to be unsatisfactory environmental dimensions of quality of life, asking purely about subjective indicators may be unwise (Oliver et al, 1996).

The danger of exclusively focussing on subjective evaluations, arising from the issue of relationship between symptomatology and subjective quality of life assessment, is described:

“if we regard quality of life judgements as consisting *only* of these perceptions, then quality of life will probably not prove to be that distinct from symptom status” (author’s emphasis, Gladis et al, 1999).

The degree to which the assessments of the general population and those with mental illness converge or diverge may depend on the domain under consideration. It has been suggested that domains consistently rated more negatively among the clinical population than the general population include work or daily activities, leisure and social relations, and family relations. Other domains demonstrate striking similarity in assessments such as finances (Mercier, 1994).

On the issue of whether the severely and persistent mentally ill may perceive the quality of their lives differently than social norms would predict, Lehman (1996) argues that (a) the psychometric properties for the better measures are comparable to the general population and (b) any inconsistent quality of life valuations may offer valuable information for clinical interventions and service planning. For instance, if a patient is dissatisfied with one area of life, which clinicians and the family are satisfied with, the relationship between the patient and care-givers may deteriorate if the patient’s concerns are not acknowledged. The importance of taking into account whether the individual wishes to perform up to the certain level, particularly among role functioning measures, has been stressed (Bowling, 1997).

Findings are therefore equivocal but a general conclusion might be that some psychiatric diagnostic categories influence subjective well-being, especially those related to affective state (e.g. depression). However, even for these groups, researchers agree that measuring quality of life or subjective well-being represents a viable means of measuring client change or programme outcomes (Fabian, 1990) with results showing that even in the presence of psychiatric symptoms, individual’s responses to questions about their well-being and satisfaction can be meaningful (Stedman, 1996). In an effort, however, to minimise the confounding of quality of life assessments by symptoms and psychological state, one practical step might be to ensure that quality of life measures are accompanied by an explicit assessment of psychopathologic symptoms (Lehman, 1996).

4 c) Content of quality of life measures

The outcome domains relevant to the mental health population have been summarised as: measures of clinical status (signs and symptoms); rehabilitative domain (adaptation and functional capacity); and humanitarian domain (subjective well-being, consumer satisfaction, quality of life and the impact on carers), (Repper and Brooker, 1998). Others have summarised the areas of importance as including functional status, access to resources and opportunities, subjective well-being, family burden and community safety (Lehman, 1996). It has been noted that consumers of mental health services are rarely involved in identifying relevant issues or designing appropriate instruments (Nocon & Qureshi, 1996).

From the perspective of using quality of life measures as an outcome tool with which to assess interventions and care, one might wonder whether it is useful or relevant for such measurement tools to include domains which, on the face of it, seem unrelated to health care interventions. Indeed one may question the remit health care is expected to have and wonder whether an improvement in symptoms as a result of treatment might not be considered a sufficiently good outcome, without negating the need to broaden the definition of clinical significance (Gladis et al, 1999).

It has, however, been argued that the very nature of mental illness (the multiple problems involved and often tenuous life circumstances), and the fact that income, freedom, social support etc are all related to psychopathology, necessitate that outcomes are seen in the broadest sense (Katschnig, 1997). Proponents therefore argue that attention should not only be given to the direct, intended effects of the intervention (the “proximal” outcomes), but also to the indirect, unintended effects (the “distal” outcomes), both positive and negative (Slaughter et al, 1991), especially given the unwanted and severe side-effects of some pharmacological interventions. In addition people with mental health problems can receive a range of, often multi-disciplinary, interventions simultaneously, the interactive effects of which need to be taken into account (Repper & Brooker, 1998). With mental illness, where there may be no “outcome” in terms of an end result free of symptoms, the goals of an intervention may include the maintenance of current functioning and continued contact with services and an instrument would need to be chosen that reflected this.

Although overlap exists between measures of quality of life, functioning and patient satisfaction, there are large differences in the conceptualisation with no universally accepted operational definition of quality of life in existence (Charlwood et al, 1999). This lack of an agreed definition, unlike with cancer care for instance, makes it difficult to assess whether instruments are covering all the relevant domains. Although quality of life and functioning measures address issues of importance to people, the way the questions are structured are based on clinical impressions, rather than reflecting the individual’s conceptualisation (ibid).

Chapter 5: OTHER ISSUES IN INTERPRETING OUTCOMES INFORMATION

- a) Katschnig (1997) draws attention to the different time frames associated with the various quality of life approaches. For instance, subjective well-being can fluctuate quickly, functioning in social roles may take longer and socio-economic conditions may change only very slowly. There may also be a trade-off between better quality of life now and better quality of life tomorrow as in the case of pharmacological interventions with adverse effects.
- b) Floor effects are frequently encountered when using normative quality of life measures among the severely mentally ill, especially in the role-functioning domains (e.g. spouse, parent, employment). There is also a need to pay attention to instrument sensitivity (Lehman, 1996; Dickerson, 1997).
- c) Many of the severely mentally ill have difficulty completing questionnaires, hence the prevalence of interviews (Lehman, 1996). Mental health patients' functional limitations (judgement, comprehension and affective state) often compound difficulties routinely associated with self-report measurement (Atkinson & Zibin, 1996). Means of improving self-report include giving very clear instructions, having someone on hand to answer queries and making the items clear and easy to understand.
- d) People make assessments based on comparison of their social group so if everyone experiences the same (negative) impact of an intervention, the intervention may not in fact be viewed negatively. This social desirability issue is particularly a problem if the individual only has contact with one social group, which might occur among people with mental illness (Fabian, 1990).
- e) It has been argued that adjustment and coping are important modifiers of outcome and are worth including in instruments (Bowling, 1997). However whether or not these have possible confounders that need to be taken into account depends on the application of the outcome measure. In the case of a randomised controlled trial it would not be necessary.
- f) Evidence shows discrepancies between patients' ratings and informants' ratings concerning patients' symptoms, psychopathology and functioning in specific areas. Stedman (1996), for instance, draws upon evidence suggesting that although there may be agreement regarding patients and staff concerning symptoms and functioning, there is less agreement in areas such as social support and occupation. Van Nieuwenhuizen et al (1997) suggests that quality of life ratings given by others such as caregivers show little to no correlation with the quality of life ratings given by people with mental illness. It seems generally accepted that quality of life has a major subjective element that can only be assessed by self-report of people themselves so that if proxies are to be used, it is preferable that they are carefully chosen and that they rate objective, observable matters.

Chapter 6: CONCLUSIONS

Most of the reviews documented make their recommendations on the basis of the extent of psychometric evaluation of the instruments, although one (Andrews et al, 1994) considered the particular applicability of the instrument to the routine practice setting. The Lehman Quality of Life Interview (Lehman, 1983) was recommended by a majority of the reviews (Atkinson & Zibin, 1996; van Nieuwenhuizen et al, 1997; Lehman 1996, op. cit). The Oregon Quality of Life Questionnaire (Bigelow et al, 1991) and the Lancashire Quality of Life Profile (Oliver, 1992) were both recommended by two of four reviews (Atkinson & Zibin, 1996; Lehman, 1996) and (Lehman, 1996; van Nieuwenhuizen, 1997) respectively.

Other instruments singled out as showing promise and/or for use among specific populations include two disease-specific measures: the Quality of Life in Depression Scale (McKenna & Hunt, 1992) and the Quality of Life Scale (Heinrichs et al, 1984), for depression and schizophrenia respectively. Finally the following were also noted for merit: the SF-36 (Ware & Sherbourne, 1992); the Quality of Life, Enjoyment and Satisfaction Questionnaire (Endicott et al, 1993); the Quality of Life Index for Mental Health (Becker et al, 1993); the Quality of Life Self-Assessment Inventory (Skantze & Malm, 1994).

The issues raised in the preceding discussion concerning the definition and remit of quality of life measures, the role of subjective evaluation and the use of proxies testify to some of the difficulties in measuring patient-based outcomes. However in spite of these difficulties, it is apparent that even among those with a chronic and severe mental illness, patients are able to evaluate their condition and the information they provide is unique and invaluable.

Although some instruments may be suitable, in terms of their brevity, for use in routine clinical practice, many of the instruments do not have established reliability and validity for use at an individual patient level². For research purposes, however, there are several instruments which have adequate properties, from a psychometric standpoint, and could be recommended for use. The particular choice of instrument needs to be assessed in the light of the question to be answered and consideration of how the outcomes information will be used. A general quality of life measure may raise issues that health care cannot address, such as housing and employment, and may be so broad that it becomes relatively insensitive to health care improvements. A disease-specific approach on the other hand may be sensitive but can lack generalisability to other disease groupings and may fail to capture all the relevant outcomes of a multi-disciplinary intervention. Health-related quality of life measures like the SF-36 (Ware & Sherbourne, 1992) may, due to a concentration on aspects of physical functioning, not be relevant to people with specific mental disorders. Functioning relates the individual's functioning to cultural norms and expectations of functioning and it is this area that health professionals can perhaps most impact upon. However, functioning can exist in a vacuum when one does not know the socio-

² Although recent evidence has been provided in support of the routine use of the Health of the Nation Outcome Scale (McCelland et al, 2000), this is a health professional-rated rating scale rather than a patient-reported measure

economic environment within which a level of functioning is achieved or whether the individual is satisfied with the level of functioning.

We have raised several issues arising with some types of mental illness which potential users of instruments would need to take into account. These include, most fundamentally, the ability of the patient to evaluate his/her condition (and the associated use of proxies) and issues concerning the interpretation of information, such as the existence of floor effects.

In terms of research needs, a consensus regarding the conceptual basis of instruments and their content in certain situations would be desirable. Also, although it is sometimes suggested that different parties, not just the patient, should evaluate quality of life, it is unclear how these evaluations, which may significantly differ, should be reconciled. Furthermore data is needed comparing quality of life data from mentally ill populations with other non-psychiatric groups to establish norms and providing information on the sensitivity of instruments. Lastly, models for integrating patient-reported outcomes information with other outcomes information are also in need of development (Lehman, 1996).

REFERENCES

- Andrews, G., Peters, L., and Teesson M. (1994) Measurement of consumer outcome in mental health. Clinical Research Unit for Anxiety Disorders, University of New South Wales and St Vincent's Hospital, Australia.
- Anderson, R. M., Sullivan, F., and Usherwood, T. P. (1990). The Medical Outcomes Study Instrument (MOSI)- use of a new health status measure in Britain. *Fam. Pract.* **7**, 205-218.
- Atkinson, M. J., and Zibin, S. (1996). Quality of life measurement among persons with chronic mental illness: A critique of measures and methods. Ottawa, Ontario, Minister of Supply and Services.
- Avis, N. E., and Smith, K. W. (1994). Conceptual and methodological issues in selecting and developing quality of life measures. In "Advances in medical sociology" (R. Fitzpatrick, Ed), pp255-280. JAI Press Inc., London.
- Baker, F., and Intagliata, J. (1982). Quality of life in the evaluation of community support systems. *Evaluation and Program Planning* **5**, 69-79.
- Barker, S., Barron, N., and McFarland, B. H. (1994) A community ability scale for chronically mentally ill consumers: I. Reliability and validity. *Community mental Health Journal* **30**, 363-379.
- Bartlett, C. J., and Coles, E. C. (1998). 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* **2**, 288-294.
- Becker, M., Diamond, R., and 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.
- Bergner, M., Bobbitt, R. A., Kressel, S., Pollard, W. E., Gilson, B. S., and Morris, J. R. (1976). The Sickness Impact Profile: Conceptual formulation and methodology for the development of a health status measure. *International Journal of Health Services* **6**, 393-415.
- Bigelow, D. A., McFarland, B. H., and Olson, M. M. (1991). Quality of life in community mental health program clients: Validating a measure. *Community Mental Health Journal.* **43**, 43-55.
- Bowling, A. (1997). "Measuring health: A review of quality of life measurement." Open University Press, Buckingham.
- Charlwood, P., Mason, A., Goldacre, M., Cleary, R., and Wilkinson, E. (eds). (1999) Health Outcome Indicators: Severe mental Illness. Report of a working group to the Department of Health. Oxford: National Centre for Health Outcomes Development.

Department of Health. (1988) Modernising Mental Health Services. Safe, sound and supportive.

Dickerson, F. B. (1997). Assessing clinical outcomes: the community functioning of persons with serious mental illness. *Psychiatr.Serv.* **48**, 897-902.

Eisen, S. V., Dill, D. L., and Grob, M. C. (1994). Reliability and validity of a brief patient-report instrument for psychiatric outcome evaluation. *Hospital and Community Psychiatry* **45**, 242-247.

Endicott, J., Nee J., Harrison, W., and Blumenthal, R. (1993). Quality of life enjoyment and satisfaction questionnaire: a new measure. *Psychopharmacology Bulletin* **29**, 321-326.

Fabian, E. S. (1990). Quality of life: A review of theory and practice implications for individuals with long-term mental illness. *Rehabilitation-Psychology* **35**, 161-170.

Fitzpatrick R, Davey C., Buxton M.J., and Jones D.R. (1998). Evaluating patient-based outcome measures for use in clinical trials. *Health Technology Assessment* **2** (14).

Frisch, M. B., Cornell, J., Villanueva, M., and Retzlaff, B. J. (1992) Clinical validation of the Quality of Life Inventory: A measure of life satisfaction for use in treatment planning and outcome assessment. *Psychological Assessment* **4**, 92-101.

Gater, R. A., Kind, P., and Gudex, C. (1995). Quality of life in liaison psychiatry: A comparison of patient and clinician assessment. *British Journal of Psychiatry* **166**, 515-520.

Gladis, M. M., Gosch, E. A., Dishuk, N. M., and Crits, C. P. (1999). Quality of life: Expanding the scope of clinical significance. *J.Consult Clin.Psychol.* **67**, 320-331.

Goodman, S. H., Sewell, D. R., Cooley, E. L., and Leaviii, N. (1993). Assessing levels of adaptive functioning: The Role Functioning Scale. *Community Mental Health Journal* **29**, 119-131.

Greenhalgh, J., and Meadows, K. (1991). The effectiveness of the use of patient-based measures of health in routine practice in improving the process and outcomes of patient care: a literature review. *Journal of Evaluation in Clinical Practice* **5**, 401-416.

Heinrichs, D. W., Hanlon, T. E., and Carpenter, W. T. (1984). The quality of life scale: An instrument for rating the schizophrenic deficit syndrome. *Schizophrenia Bulletin* **10**, 388-398.

Holcomb, W. R., Morgan, P., Adams, N. A., Ponder, H., and Farrel, M. (1993). Development of a structured interview scale for measuring quality of life of the severely mentally ill. *Journal of Clinical Psychology* **49**, 830-840.

- Hunter, J., Higginson, I., and Garralda, E. (1996). Systematic literature review: outcome measures for child and adolescent mental health services. *J Public Health Med.* **18**, 197-206.
- Huxley, P. (1986). "Quality measurement in mental health services: A discussion paper on quality of life measurement." *Good practices in mental health*, London.
- Jenkins, R. (1992). Health targets. In "Measuring mental health needs" (G. Thornicroft and C. R. Brewin, Eds.), pp. 18-41. Gaskell, London.
- Jenkins, R., and Griffiths, S. (eds) (1991) *Indicators for Mental Health in the Population*. Department of Health. London: HMSO.
- Jenkinson, C. (1995). Measuring the outcomes of medical care: possibilities and limitations. *Soc. Sci. Med.* **41**, 1395-1401.
- Katschnig, H. (1997). How useful is the concept of quality of life in psychiatry? In "Quality of life in mental disorders" (H. Katschnig, H. Freeman and N. Sartorius, Eds), pp. 3-16. John Wiley and Sons Ltd., Chichester.
- Lehman, A. F. (1983). The well-being of chronic mental patients. *Archives of General Psychiatry* **40**, 369-373.
- Lehman, A. F. (1996). Measures of quality of life among persons with severe and persistent mental disorders. In "Mental health outcome measures" (G. Thornicroft and M. Tamselba, Eds), pp. 74-92. Springer-Verlag, Berlin.
- Lehman, A. F. (1997). Instruments for measuring quality of life in mental illnesses. In "Quality of life in mental disorders" (H. Katschnig, H. Freeman, and N. Sartorius, Eds.), pp. 79-94. John Wiley and Sons Ltd., Chichester.
- McDowell, I., and Jenkinson, C. (1996). Development standards for health measures. *J Health Serv Res Policy* **1**, 238-246.
- McDowell, I. and Newell, C. (1996). "Measuring Health: A guide to rating scales and questionnaires." Oxford University Press, New York.
- McClelland, R., Trimble, P., Fox, M. L., Stevenson, M. R., and Bell, B. (2000) Validation of an outcome scale for use in psychiatric practice. *Quality in Health Care* **9**, 98-105.
- McKenna, S. P., and Hunt, S. M. (1992). A new measure of quality of life in depression: Testing the reliability and construct validity of QLDS. *Health Policy* **22**, 321-330.
- Mercier, C. (1994). Improving the quality of life of people with severe mental disorders. *Social-Indicators-Research* **33**, 165-192.

Nocon, A., and Qureshi, H. (1996). People with mental health problems. *In* "Outcomes of Community Care for Users and Carers" (A. Nocon and H. Qureshi), pp. 89-104. Open University Press, Buckingham.

Oliver, J. P. J. (1992). The social care directive: Development of a quality of life profile for use in community services for the mentally ill. *Social Work & Social Sciences Review* **3**, 5-45.

Oliver, J., Huxley, P., Bridges, K., and Mohamad, H. (1996). Definitions and conceptual issues concerning quality of life. *In* "Quality of life and mental health services." (J. Oliver, P. Huxley, K. Bridges, and H. Mohamad, Eds.), pp. 15-47. Routledge, London and New York.

Oxman A. (1994). Checklists for review articles. *British Medical Journal* **309**, 648-651.

Repper, J. and Brooker, C. (1998). Difficulties in the measurement of outcome in people who have serious mental health problems. *J.Adv.Nurs.* **27**, 75-82.

Rosen, A., Hadzi-Pavlovic, D., and Parker, G. (1989). The Life Skills Profile: A measure assessing function and disability in schizophrenia. *Schizophrenia Bulletin* **15**, 325-337.

Ryff, C. D. and Singer, B. (1996). Psychological well-being: Meaning, measurement, and implications for psychotherapy research. *Psychother.Psychosom.* **65**, 14-23.

Sackett, D., Haynes, B., Guyatt, G., and Tugwell P. (1991). How to read reviews and economic analyses. *In* "Clinical Epidemiology: A basic science for clinical medicine". Little, Brown & Company.

Selai, C. and Trimble, R. (1999). Assessing quality of life in dementia. *Aging and Mental Health* **3**, 101-111.

Skantze, K., Malm, U., Dencker, S. J., May, P. R. A., and Corrigan, P. (1992). Comparison of quality of life with standard of living in schizophrenia outpatients. *British Journal of Psychiatry* **161**, 797-801.

Slaughter, J. G., Lehman, A. F., and Myers, C. P. (1991). Quality of life of severely mentally ill adults in residential care facilities. *Adult-Residential-Care-Journal* **5**, 97-111.

Stedman, T. (1996). Approaches to measuring quality of life and their relevance to mental health. *Australian-and-New-Zealand-Journal-of-Psychiatry* **30**, 731-740.

Van Nieuwenhuizen, C., Schene, A. H., Boevink, W. A., and Wolf, J. R. L. M. (1997). Measuring the quality of life of clients with severe mental illness: A review of instruments. *Psychiatric-Rehabilitation-Journal* **20**, 33-41.

Veit, C. T., and Ware, Jnr, J. E. (1983). The structure of psychological distress and well-being in general populations. *Journal of Consulting and Clinical Psychology* **51**, 730-742.

Ware J. E. & Sherbourne C. D. (1992). The MOS 36-item short-form health survey (SF-36). *Medical Care* **30**, 473-483.

Wing, J. K. (1994). Health of the Nation Outcome Scales: HoNOS Field Trials. London: Royal College of Psychiatrists Unit.

Appendix: psychometric evaluation

Patient-reported health outcome measures can be generic (i.e. they can be used across different diseases, as well as in healthy populations) or disease specific (developed for use in a specific patient group). Generic instruments are designed to measure aspects of health that are of universal importance and are therefore suitable for comparisons between different groups of patients or healthy populations. They can take account of the influence of co-morbidity on health and have the potential to capture the side-effects of an intervention. This makes them potentially useful for assessing the impact of new health care technologies when the therapeutic effects are uncertain. Specific instruments can be selected that reflect the areas considered by patients or clinicians to be of greatest importance and can be specific to a particular disease, population, function, condition, or problem. The narrow focus of specific instruments has the potential to make them more responsive to changes in health. The possible complementary role played by generic and specific instruments has led to recommendations that the two approaches be used in conjunction.

Patient-reported instruments can produce a profile of scores that relate to different 'dimensions' of health or a single index of health. Instruments that produce a health profile are usually based on the psychometric approach to instrument development. Although some single indexes are based on the psychometric approach they are usually based on approaches derived from economics and decision theory.

Health profiles measure health across a number of distinct 'dimensions' such as physical functioning, mental health and role limitations. The items within health profiles are scored and summed to reflect individual dimensions but sometimes also produce a single index.

The development of patient-reported health outcome measures involves devising, scaling and testing the items or questions that form the content of an instrument. Instrument content can be derived from literature reviews, theory and interviews of focus groups conducted with patients or experts involved in patient care; instrument content is often based on some combination of all three. If a patient-reported instrument is to have content validity as a measure that is relevant to the recipients of care then patients should be involved in the derivation of items (Fitzpatrick et al, 1998). Following derivation items are then scaled, the most common approach being adjectival scales, including Likert scaling. Following construction it is normal to test an instrument on a small but representative sample of patients to check for any wording difficulties or ambiguities.

Before an instrument can be recommended for application the measurement properties of reliability, validity and responsiveness should be assessed. The data quality, scaling assumptions and dimensionality of the instrument should be assessed concurrently. Individual items should be assessed for levels of missing data and response frequencies. Items with relatively large amounts of missing data should be removed from the instrument. Items with large end effects at the floor or ceiling are poor discriminators and should be removed from the instrument. In deriving items, instrument authors often hypothesise dimensions that may be based on their own experience or theory. These dimensions can be assessed empirically through the statistical techniques of factor analysis and principal component analysis (PCA)

which is sometimes referred to as internal validation. PCA groups items together that measure the same underlying construct and provides evidence for the underlying dimensionality of an instrument. Item-total correlation is usually undertaken after the application of PCA and assesses the strength of association between an item and the remainder of the dimension.

Reliability is concerned with whether an instrument is internally consistent or reproducible. Tests of internal consistency are appropriate for multi-item scales. Cronbach's alpha assesses the overall level of correlation between items within a scale. It is equivalent to the average level of correlation between all the possible halves of a scale. Internal consistency can be assessed with a single administration of an instrument but it is only applicable to multi-item scales. Instruments that do not have multi-items scales must be assessed for test-retest reliability. Internal consistency should not be the sole criterion of reliability because it cannot take account of variation over time in patients whose health has not changed. Test-retest reliability is designed to take account of variation over time. It assesses the level of association between two sets of instrument scores from the same group of patients on two different occasions. There is no real agreement on the length of time between administrations of test and retest questionnaires but it should not be so short that patients can recall their previous responses, nor should it be so long that health may have changed. In practice the use of postal surveys means that two weeks is the minimum time between test and retest. It is common for such studies to include a health transition question to identify patients who do not change between administrations. These questions assess whether there has been any change in the aspects of health that are the focus of the instrument being assessed.

Standards for the reliability coefficient are dependent on whether the instrument is intended for use with groups or individual patients. For decisions about groups of patients a reliability coefficient of 0.7 is recommended while for decisions about individual patients the more stringent criterion of 0.9 is recommended (ibid).

Validity is concerned with whether an instrument is measuring what is intended. There are a number of ways of assessing the validity of instruments that include both qualitative and quantitative methods. Validity is not a fixed property that can be ascertained from a single experiment and should be assessed in relation to the application of an instrument (Jenkinson, 1995). New instruments, refined instruments and instruments that are being used in a new setting should be tested for validity. Face and content validity are qualitative matters of judging whether an instrument is suitable for its proposed application. Face validity concerns judgements about individual items after an instrument has been constructed whereas content validity is concerned with judgements about how well the domain of interest has been sampled for items. These forms of validity are assessed through inspection of the instrument and together they assess whether the items adequately relate to the domain of interest and whether the domain is sufficiently covered by the items in their entirety (Fitzpatrick et al, 1998).

Patient-reported health outcome measures are concerned with the measurement of variables that are not directly observable which are referred to as 'hypothetical constructs'. Drawing on theoretical or empirical work, patient-reported measures can be expected to have quantifiable relationships with other constructs. For example,

patients with restricted physical functioning may take more days off work; patients with severe pain may take more analgesics. Construct validity involves comparing instrument scores to other variables. Construct validation is not necessarily concerned with strong relationships, but whether an instrument measures something in a way that is predicted by theory or established evidence. Thus validity testing should not be seen as a once and for all exercise but an ongoing process of accumulating evidence arising from a number of tests.

The statistical methods usually involve correlation but if groups are being compared t-tests or non-parametric equivalents are used. The method of groups is used when, according to theory, one group of patients possesses more or less of the construct being measured. For example, in assessing the MOS Short-form General Health Survey scale of mental health for construct validity, Anderson et al (1990) compared the scores for subjects with a mental illness with scores for subjects who were free from mental illness. Subjects with mental illness were found to have a significantly lower mean score.

There are no agreed criteria for levels of correlation between an instrument and other variables used in assessing construct validity (Avis and Smith, 1994). Correlations should not be too high as this would imply that the new instrument is measuring the same thing as the variables used for testing which undermines the role of the new instrument. McDowell and Jenkinson (1996) recommend that expected correlations should be specified a priori. Developers must think about the variables being used for assessing validity and expected levels of correlation rather than simply correlating available data, which could produce a spurious result.

Responsiveness refers to the ability of an instrument to measure important changes in health. This is an important measurement property of any instrument that is to be used for measuring outcomes. Responsiveness is assessed by looking at changes in instrument scores for groups of patients whose health is known to have changed. Two approaches are applied in selecting patients for studies of responsiveness. In the first, the patient is given an intervention of known efficacy. The responsiveness of instruments is assessed for patients who are expected to have improved. In the second, patient or physician ratings of change are used to select patients (Fitzpatrick et al, 1998). There are a number of statistics that can be used for quantifying the responsiveness of instruments, many of which are referred to as effect sizes. Responsiveness is rarely defined in terms of importance to patients, but as has been stated, this is an important consideration in relation to patient-reported measures (ibid). Few studies have assessed whether the level of change being demonstrated by instruments is indeed of a level that patients consider important.