A systematic review, psychometric analysis and qualitative assessment of generic preference-based measures of health in mental health populations and the estimation of mapping functions from widely used specific measures

John Brazier, Janice Connell, Diana Papaioannou, Clara Mukuria, Brendan Mulhern, Tessa Peasgood, Myfawnwy Lloyd Jones, Suzy Paisley, Alicia O'Cathain, Michael Barkham, Martin Knapp, Sarah Byford, Simon Gilbody and Glenys Parry

NHS

National Institute for Health Research

A systematic review, psychometric analysis and qualitative assessment of generic preference-based measures of health in mental health populations and the estimation of mapping functions from widely used specific measures

John Brazier,[1]* Janice Connell,[1] Diana Papaioannou,[1]
Clara Mukuria,[1] Brendan Mulhern,[1] Tessa Peasgood,[1]
Myfawnwy Lloyd Jones,[1] Suzy Paisley,[1]
Alicia O'Cathain,[1] Michael Barkham,[2] Martin Knapp,[3]
Sarah Byford,[3] Simon Gilbody[4] and Glenys Parry[1]

[1]Health Economics and Decision Science, School of Health and Related Research (ScHARR), University of Sheffield, Sheffield, UK
[2]Centre for Psychological Services Research, University of Sheffield, Sheffield, UK
[3]Centre for the Economics of Mental and Physical Health, King's College London, London, UK
[4]Department of Health Sciences, University of York, York, UK

*Corresponding author

Declared competing interests of authors: John Brazier was a developer of the SF-6D and Michael Barkham was a developer of the CORE-OM.

Disclaimer: This report contains transcripts of interviews conducted in the course of the research and contains language that may offend some readers.

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme or, commissioned/managed through the Methodology research programme (MRP), and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: www.hta.ac.uk/

This report

This issue of the Health Technology Assessment journal series contains a project commissioned/managed by the Methodology research programme (MRP). The Medical Research Council (MRC) is working with NIHR to deliver the single joint health strategy and the MRP was launched in 2008 as part of the delivery model. MRC is lead funding partner for MRP and part of this programme is the joint MRC–NIHR funding panel 'The Methodology Research Programme Panel'

To strengthen the evidence base for health research, the MRP oversees and implements the evolving strategy for high quality methodological research. In addition to the MRC and NIHR funding partners, the MRP takes into account the needs of other stakeholders including the devolved administrations, industry R&D, and regulatory/advisory agencies and other public bodies. The MRP funds investigator-led and needs-led research proposals from across the UK. In addition to the standard MRC and RCUK terms and conditions, projects commissioned/managed by the MRP are expected to provide a detailed report on the research findings and may publish the findings in the HTA journal, if supported by NIHR funds.

Abstract

A systematic review, psychometric analysis and qualitative assessment of generic preference-based measures of health in mental health populations and the estimation of mapping functions from widely used specific measures

John Brazier,[1]* Janice Connell,[1] Diana Papaioannou,[1] Clara Mukuria,[1] Brendan Mulhern,[1] Tessa Peasgood,[1] Myfawnwy Lloyd Jones,[1] Suzy Paisley,[1] Alicia O'Cathain,[1] Michael Barkham,[2] Martin Knapp,[3] Sarah Byford,[3] Simon Gilbody[4] and Glenys Parry[1]

[1]Health Economics and Decision Science, School of Health and Related Research (ScHARR), University of Sheffield, Sheffield, UK
[2]Centre for Psychological Services Research, University of Sheffield, Sheffield, UK
[3]Centre for the Economics of Mental and Physical Health, King's College London, London, UK
[4]Department of Health Sciences, University of York, York, UK

*Corresponding author

Background: Generic preference-based measures of health like the EQ-5D and SF-6D® are increasingly being used in economic evaluation and outcome assessment. However, there are concerns as to whether or not these generic measures are appropriate for use in people with mental health problems.

Objectives: The EQ-5D and SF-36® (including its derivatives the SF-12® and SF-6D) were assessed using the psychometric criteria of validity and responsiveness using quantitative and qualitative methods. Another aim was to estimate mapping functions between the EQ-5D and SF-6D and condition-specific measures, where appropriate.

Design: Four studies were undertaken to examine the appropriateness of the measures: (1) a systematic review of quantitative evidence on validity and responsiveness; (2) a further quantitative assessment of these criteria using existing data sets; (3) a review of qualitative research on the quality of life of people with mental health problems; and (4) qualitative semistructured interviews of people with a full range of problems. A fifth study estimated mapping functions between mental health-specific measures and the EQ-5D and SF-6D.

Setting: A choice of venue was offered for the interviews including the participant's own home, a room at the university or a centre frequently used by mental health services.

Participants: The interviews were undertaken with 19 people with a broad range of mental health problems at varying levels of severity.

Main outcome measures: The reviews included the EQ-5D and SF-36 (and the SF-12 and SF-6D). The psychometric analysis included the Hospital Anxiety and Depression Scale (HADS), Clinical Outcomes in Routine Evaluation – Outcome Measure (CORE-OM), Generalised Anxiety Disorder Assessment (GAD-7), General Health Questionnaire (GHQ-12) and Patient Health Questionnaire (PHQ-9).

Results: (1) and (2) The EQ-5D and SF-36 achieved an adequate level of performance in depression, and to some extent in anxiety and personality disorder. Results from the psychometric analyses in schizophrenia and bipolar disorder have been more mixed. (3) A framework analysis of 13 studies identified six major themes. (4) The interview data fitted the themes from the review well and resulted in minor modifications to the themes. The final set of themes comprised: well-being and ill-being; control, autonomy and choice; self-perception; belonging; activity; hope and hopelessness; and physical health.

Conclusions: The EQ-5D and SF-36 achieved mixed results in the quantitative testing against psychometric criteria. The qualitative analysis suggests this is because they provide a very limited coverage of themes identified by people with mental health problems. Recommendations for future work include the development of new preference-based measures in mental health that are based on, or substantially revise, an existing measure.

Funding: The Medical Research Council.

Contents

List of tables

List of figures

List of abbreviations

AHEAD	Assessing Health Economics of Antidepressants	ICD-10	*International Classification of Diseases*, Tenth Edition
ASSIA	Applied Social Sciences Index and Abstracts	LOWESS	locally weighted scatterplot smoothing
BDI	Beck Depression Inventory	MADD	mixed anxiety and depressive disorders
BIC	Bayesian information criterion	MATISSE	Multicentre study of Art Therapy In Schizophrenia: Systematic Evaluation
BPRS	Brief Psychiatric Rating Scale		
BPRS-E	Brief Psychiatric Rating Scale – Expanded		
CINAHL	Cumulative Index to Nursing and Allied Health Literature	MCS	mental health component summary score
CIS-R	Revised Clinical Interview Schedule	NICE	National Institute for Health and Care Excellence
CMHT	community mental health team	OLS	ordinary least squares
CORE-OM	Clinical Outcomes in Routine Evaluation – Outcome Measure	PANSS	Positive and Negative Syndrome Scale
DSM-IV	*Diagnostic and Statistical Manual of Mental Disorders* – Fourth Edition	PCS	physical health component summary score
		PHQ-9	Patient Health Questionnaire
EMPIRIC	Ethnic Minority Psychiatric Illness Rates in the Community	PoNDER	Psychological Interventions for Postnatal Depression
GAD-7	Generalised Anxiety Disorder Assessment	POPMACT	Prevention of Parasuicide by Manual Assisted Cognitive Behaviour Therapy
GAF	Global Assessment of Functioning	PROM	Patient Reported Outcome Measure
GHQ-12	General Health Questionnaire	PTSD	post-traumatic stress disorder
HADS	Hospital Anxiety and Depression Scale	QALY	quality-adjusted life-year
HADS-A	Hospital Anxiety and Depression Scale – Anxiety	QUATRO	Quality of Life following Adherence Therapy
HADS-D	Hospital Anxiety and Depression Scale – Depression	RCT	randomised controlled trial
		RMSE	root-mean-square error
HADS-T	Hospital Anxiety and Depression Scale – Total score	SCEPTRE	Study of Cost-Effectiveness of Personality disorder TREatment
HRQoL	health-related quality of life	SD	standard deviation
IAPT	Improving Access to Psychological Therapies	SRM	standardised response mean
		TTO	time trade-off

Scientific summary

Background

There has been an increasing use of the EQ-5D and SF-36® (and its derivative the SF-6D®) in the economic evaluation of health-care interventions. These generic measures can produce health-state utility values on a scale from 0 to 1 in order to calculate quality-adjusted life-years (QALYs). These QALY calculations are used by policy-makers [e.g. the National Institute for Health and Care Excellence (NICE)] to assess the cost-effectiveness of interventions in terms of their cost per QALY gained. Generic measures of health are also being used to assess the outcome of care in routine practice (e.g. the NHS Patient Reported Outcome Measures programme). There are concerns that generic measures have been primarily designed for physical health problems and miss important aspects of the impact of mental health problems on the quality of people's lives. Another concern is that even where generic measures are found to be appropriate, they are often not used in clinical studies.

Aims and objectives

This project examines in detail the appropriateness of the EQ-5D and the SF-36 and its derivatives in populations with mental health problems in terms of their validity (construct and content) and responsiveness. This research considers groups of people with mental health problems, including those with common mental health problems (e.g. mild to moderate depression, anxiety, obsessive compulsive disorder and panic disorders), severe and complex non-psychotic disorders (e.g. personality disorder), and schizophrenia and other psychotic disorders. For those conditions where the EQ-5D and/or SF-36 is found to be valid, a further aim is to estimate functions to predict their scores from mental health-specific measures commonly used in clinical trials and other studies (also known as mapping or cross- walking).

Methods and results

The validity of the EQ-5D and SF-36 (including its derivatives, the SF-12® and SF-6D) was examined using the psychometric criteria of validity and responsiveness. Validity is the extent to which an instrument measures the concept of interest, and in this case it is the extent to which EQ-5D and SF-36 (and its derivatives) reflect the impact of mental health problems on people's health-related quality of life, whether or not they use services. Construct and content validity are assessed in this research. Construct validity is assessed in terms of the ability of these measures to reflect known-group differences (or discriminant validity) and converge with other indicators of the concept (convergent validity). Responsiveness is the ability to reflect change in the population over time. Evidence on construct validity was obtained through the reviewing of published evidence and original analysis of seven existing data sets. Content validity assesses the degree to which the items of the measures represent the domains of interest, in terms of the impact of mental health problems on the quality of people's lives. Qualitative evidence on content validity was obtained from a systematic review of the literature and analyses of interviews of people with mental health problems.

In all, four studies were undertaken to examine the appropriateness of the measures, and a fifth study to estimate mapping functions between EQ-5D and SF-6D and commonly used mental health-specific measures.

Study 1: a systematic review of the validity and responsiveness of the EQ-5D, SF-36, SF-12 and SF-6D

The aim of this study was to assess research evidence on the validity and responsiveness of the SF-36, SF-12, SF-6D and EQ-5D in mental health conditions.

Methods
Systematic reviews were undertaken in five mental health conditions. Ten databases were searched from inception to August 2009. Studies were appraised and data extracted. A narrative synthesis was performed on known-group validity, convergent validity (strength of association with related measures such as symptoms or functioning) and responsiveness (e.g. changes in scores in responders/non-responders to treatment and correlation with changes in related measures).

Results
For schizophrenia, the majority of evidence related to the SF-36 (25 studies) and EQ-5D (nine studies). Both measures demonstrated known-group differences but this was mostly limited to differences between individuals with schizophrenia and the general population. Contradictory results were found in studies measuring convergent validity and responsiveness using clinical measures of symptom severity. For bipolar disorder, 23 studies were identified, almost exclusively on the SF-36, which was able to detect known differences in symptom severity and correlated strongly with clinical measures of depression, though weakly with measures of mania. For personality disorders, the majority of studies (six out of nine) related to the EQ-5D, which reflected known-group differences and responsiveness. For depression and anxiety, 23 EQ-5D and eight SF-6D studies were identified. Both measures demonstrated good convergent validity and responsiveness for depression; however, the results on known-group differences may be driven by the presence of comorbid depression in patients with anxiety disorders.

Study 2: assessing the validity and responsiveness of the EQ-5D and SF-6D using existing data sets

The purpose of Study 2 was to generate further evidence on the validity and responsiveness of the EQ-5D and SF-6D in anxiety, depression and schizophrenia samples by undertaking secondary analysis of available data sets.

Methods
Psychometric performance of the generic preference-based measures was assessed in comparison with mental health-specific measures using samples taken from seven patient data sets (total sample size = 5748 patients). For anxiety and depression, the EQ-5D was assessed in comparison with the Hospital Anxiety and Depression Scale (HADS), and the SF-6D was assessed in comparison with the Clinical Outcomes in Routine Evaluation – Outcome Measure (CORE-OM). For the schizophrenia samples, the EQ-5D and SF-6D were both assessed using the Brief Psychiatric Rating Scale – Expanded (BPRS-E) and Positive and Negative Syndrome Scale (PANSS). Tests of acceptability (in terms of missing data and response rates), construct validity and responsiveness were carried out.

Results
Results support the findings from the review that both the EQ-5D and SF-6D are valid in common mental health disorders, particularly in depression but also in anxiety to a lesser extent. There was evidence that both measures were acceptable to respondents, are moderately correlated with the condition-specific instruments used, and can significantly discriminate between severity groups. Both measures also displayed a moderate level of responsiveness to change. The results in schizophrenia were less clear. The EQ-5D score appears able to reflect some symptoms of schizophrenia but evidence for the convergent validity of the SF-6D was limited. Both measures were able to discriminate between severity groups, but neither measure appeared responsive to change.

Study 3: estimating mapping functions between mental health-specific measures and the EQ-5D and SF-6D

The aim of this study was to provide mapping functions to predict EQ-5D or SF-6D values from condition-specific measures used in common mental health conditions, where the generic measures were found to be valid in Studies 1 and 2.

Methods

Using data from four studies ($n = 286$, $n = 475$, $n = 394$ and $n = 213$ patients respectively) available to the authors of this report, mapping functions were estimated by regression between the EQ-5D and the HADS, and between the SF-6D and the HADS, General Health Questionnaire (GHQ-12), Patient Health Questionnaire (PHQ-9), Generalised Anxiety Disorder Assessment (GAD-7) and CORE-OM. EQ-5D and SF-6D scores and dimension-level response were used as independent variables. For each mental health-specific measure, total scores, dimension scores and item response were entered as dependent variables. For models with dimension scores and item-level responses, additive models were fitted to the data, with interactions and squared terms explored for models with dimension scores. For the EQ-5D and SF-6D scores, ordinary least squares regression was used to estimate the models, and Tobit models were explored for the EQ-5D (owing to the presence of a large proportion of responders scoring 1). Multinomial models were also fitted to the models with dimension-level response as independent variables. The performance of the models was examined using R^2 (where appropriate), Bayesian information criterion, root-mean-square error and the ability to predict scores across severity groups.

Results

The results indicated that commonly used condition-specific measures for depression and anxiety are not good predictors of EQ-5D and SF-6D scores. The HADS was found to predict EQ-5D scores poorly in two data sets; it substantially overpredicted at the worst severity and underpredicted at the lowest severity. The HADS and other measures, notably the CORE-OM, were better able to predict the SF-6D, though the mapping functions continued to suffer from some degree of over- and underprediction towards the ends of the ranges.

Study 4: a synthesis of qualitative research on the quality of life of people with mental health problems

The aim of study 4 was to find out from people with mental health problems how their problems impact on their quality of life, using a review of the qualitative literature. This review provided the basis for the topic guide used in semistructured in-depth interviews and an initial framework for their analyses.

Methods

Our research firstly involved a systematic review of qualitative studies undertaken with people with mental health problems. Studies were only included where they reported the views of people with mental health problems and supported these with participant quotes. Participants had to be either diagnosed or attending an establishment for people with mental health problems. Extensive searching was undertaken using an iterative approach up to April 2012. Framework analysis was used to allow the identification of common and variable patterns of themes within and across different studies. The searching and reviewing was undertaken by two researchers. A multidisciplinary team met regularly, in addition to meetings with clinicians and a user representative to discuss and challenge the inclusion and exclusion criteria, thematic framework, and conceptual interpretations and conclusions.

Results

A framework analysis of 13 studies revealed six major themes: well-being and ill-being; control, autonomy and choice; self-perception; belonging; activity; and hope. In particular, it underlined the complexity of the measurement of quality of life in people with mental health problems and the difficulty in separating

the symptoms of mental health from other aspects of life quality. One of the strongest themes was the importance of a sense of belonging, derived from good quality relationships. However, relationships and activity can also have a negative impact. A limitation of the review was that available studies focused on quality of life of people with severe and enduring mental health problems, particularly schizophrenia rather than affective disorder. This is addressed in the interview study.

Study 5: results from qualitative semistructured interviews of people with mental health problems

The purpose of the fifth study was to examine further the domains of quality of life that are important to people with mental health problems through interviews, and to confirm or otherwise add to the themes found in the review. To address one of the limitations of the literature review, care was taken to recruit people with a full range of problems including those with mild to moderate affective problems.

Methods
Qualitative semistructured interviews were conducted with 19 people with a broad range of mental health problems at varying levels of severity. Participants were recruited through the local Improving Access to Psychological Therapies (IAPT) service for milder depression and anxiety, and through two community mental health teams who identified those with more severe problems, including complex non-psychotic disorders and psychotic disorders. The interviews were analysed thematically using framework analysis. The themes from the review made up the initial themes of the framework.

Results
Despite widening the types and severity of mental health problems studied, our interview data fitted well with the themes from the review. Any differences tended to be within the themes and related to the degree of impact of the themes on different levels of severity, chronicity and diagnosis. With some exceptions, those with severe chronic difficulties were more likely to talk about losses and things that took quality away from life, whereas those with moderate or relatively short-lived problems spoke of the things that added quality to life. It was also found that depression had a greater impact on quality of life than other problems such as anxiety, though coping mechanisms could have a detrimental effect on other areas of life (e.g. avoidance). The only change made to the themes was that physical health was found to be more important among the interviewees than suggested by the review, so this was added as a seventh theme, as opposed to a minor sub-theme as in the review. Another important finding was a conflict between the short-term stress and effort required for a fuller life (e.g. work or social activity) and the longer-term benefits.

The content of the EQ-5D and SF-6D has been reviewed against the seven themes identified in the qualitative research and the need to cover both positive and negative aspects. In summary, the EQ-5D covers little of the content of these seven themes because of its focus on physical health. Only physical health is covered, and, to some extent, activity, which is included in a rather crude way through usual activities. The EQ-5D covers ill-being in terms of depression and anxiety but not well-being. The SF-36 covers more, through having a multi-item dimension on mental health and a vitality dimension that includes more aspects of well-being and ill-being, and some aspects of social functioning. On the other hand, like the EQ-5D, it fails to include the psychological responses of people with mental health problems to occupational and social activities such as stress, self-perception and control, autonomy and choice, and hope and hopelessness. These generic measures do not contain a sufficient proportion of the domains identified by people with mental health problems in the qualitative research.

Conclusions

The EQ-5D and SF-36 achieve an adequate level of performance in terms of the classical psychometric tests for validity and responsiveness in depression and, to some extent, in anxiety and personality disorder. The results of the mapping work, however, raise some concerns as to whether or not the EQ-5D in particular is able to fully reflect the specific impact of common mental health problems. Results from the psychometric analyses were more mixed in schizophrenia and bipolar disorder, with a suggestion that the EQ-5D and SF-36 may be reflecting depression rather than other consequences of these conditions. Although important issues were identified within the quantitative evidence, it was unclear why the evidence appears so mixed in relation to the EQ-5D and SF-36. This was the rationale for undertaking the qualitative research as this provides evidence on the content validity of these measures in people with mental health problems. This found that the generic measures did not cover many of the theories identified in the qualitative research.

Although this project has been comprehensive in the way it has adopted a multimethod approach and sought to use the best available evidence, it does have a number of limitations. Many of the quantitative studies reviewed were limited in terms of size and coverage for conditions. The studies are normally designed for other purposes, and the measures used to assess construct validity and responsiveness were often clinical measures of symptoms that may not provide a good indicator of the concepts of interest. Although the qualitative work was intended to provide a more in-depth picture of the problem, the population of people with mental health problems in the qualitative research was not comprehensive despite an extensive review of the literature and an attempt to recruit across the spectrum of conditions for the interviews. Furthermore, the samples were influenced by the clinicians used to help recruit participants.

Future research recommendations

- Further testing of construct validity and responsiveness of these generic measures should be undertaken using a larger number of data sets.
- The analysis of content validity should be extended to existing condition-specific outcome measures used in mental health research. This is important for guiding the choice of mental health measures for use in research.
- There is a case for the development of a preference-based mental health measure. This could be an enhanced version of an existing generic measure, created by adding extra dimensions (i.e. 'bolt-ons'), or a new specific mental health measure for calculating QALYs, which might include one or two physical domains but would focus on the themes identified in this research.
- The qualitative research needs to be extended to those conditions not well represented (such as obsessive compulsive disorder) and should involve participants recruited through different channels to avoid the risk of selection bias that can result from a reliance on professionals.

Funding

Chapter 1 Background and introduction

There has been a shift in mental health services from an emphasis on treatment focused on reducing symptoms based on a narrow notion of health and disease, to a more holistic approach which takes into consideration both well-being and functioning.[1,2] Mental health services in the UK, for example, are now including the routine assessment of patient-reported outcomes in psychological services [e.g. the Improving Access to Psychological Therapies (IAPT) initiative and the Department of Health's Patient Reported Outcome Measures (PROMs) programme]. At the same time, there has been an increasing use of generic measures of health, such as the EQ-5D, SF-36® (and its derivatives) and HUI3 in the economic evaluation of health-care interventions. These generic preference-based measures produce health-state utility values on a scale from 0 to 1 in order to calculate quality-adjusted life-years (QALYs). These QALY calculations are used by policy-makers including the National Institute for Health and Care Excellence (NICE) to assess the cost-effectiveness of interventions in terms of their cost per QALY gained.[3–10] These generic measures of health are designed for both physical and mental health problems; however, some argue that they are not suitable for people with more severe mental health problems, including psychosis.[11,12]

In this report we present the findings of a research project examining the appropriateness of these generic measures for assessing the impact of mental health problems on quality of life. This chapter begins by describing the EQ-5D and SF-36 (and its derivative, the preference-based SF-6D® index). It goes on to consider the different definitions of quality of life and the problems of measuring such a subjective concept, in order to provide a broader context for the work being presented in this report. It then considers how the key measurement concept of validity is approached in the literature before providing an overview of the research undertaken for the report.

Generic preference-based measures

Many agencies around the world (e.g. NICE) recommend assessing the efficiency of new health-care interventions in terms of their incremental (or extra) cost per QALY gained.[4–10] The QALY combines survival and health-related quality of life into a single measure of value. The number of QALYs is calculated by multiplying survival in life-years by a value assigned to those years, which is known as a health-state utility value. Full health has a value of 1 and states equivalent to being dead are given a value of 0, with states worse than dead being negative.[13] While the QALY was initially used mainly to assess the benefits of life-saving treatments such as coronary artery bypass graft,[14] over time it has been applied to assess the benefits of interventions aimed primarily at improving quality of life rather than quantity, including interventions for people with mental health conditions. The purpose of using a measure such as the QALY is to inform comparisons not only between treatments, but also between programmes of care across physical and mental health conditions.

This report appraises two generic preference-based measures used to generate values for health-related quality of life (HRQoL) on the zero-to-one QALY scale, and considers whether or not they are appropriate in people with mental health problems. The most commonly used measures for valuing health states in order to generate QALYs are the generic preference-based measures of health, such as the EQ-5D,[15] SF-6D,[16] HUI3[17] and Quality of Well-Being Scale.[18] Generic preference-based measures have two parts: one is a classification for describing health across a set of dimensions or domains, and the second is an algorithm for assigning values to each health state defined by the classification. The EQ-5D, for example, has five dimensions, each with three levels that together define 243 possible health states. A respondent is assigned a health state through their completion of a short questionnaire in which they indicate the level best describing their health on each dimension. The scoring algorithm is provided by valuations obtained from members of the general public using a valuation technique. In the case of the EQ-5D, the valuation technique is called time trade-off (TTO), whereby respondents are asked how many years in full health are equivalent to a longer period in an ill-health state.[9] The focus for this report, however, is not the method

of scoring per se, but the description of health used by these instruments, as this largely determines the appropriateness of their use in mental health.

The EQ-5D has been selected as the most widely used preference-based measure and the one that is preferred by NICE in health technology assessment submissions and the Department of Health PROMs programme. The SF-36 and its derivatives (the SF-12® and the preference-based SF-6D) have also been selected because the SF-36 is another widely used generic measure of health, and the SF-6D is accepted by a number of other agencies around the world.[3–10]

EQ-5D

The EQ-5D questionnaire comprises five dimensions: mobility, self-care, usual activities, pain and anxiety/depression (*Table 1*). These can be seen as a combination of physical functioning (mobility, self-care), mental health (depression and anxiety), social functioning (which may be included in usual activities) and symptoms (pain and discomfort). Respondents are asked to report their level of problems (no problems, some/moderate problems or extreme problem/unable to do) on each dimension to provide a position on the EQ-5D health-state classification system. Responses can be converted into one of 243 different health states [ranging from no problems on any of the dimensions (11111) to severe problems on all five dimensions (33333)], each with its own preference-based score. Preference-based scores are determined by eliciting strength of preference data, that is, by establishing which health states are preferred from a sample of the general public. Preferences are elicited using the TTO technique, which involves asking participants to consider the relative lengths of time in full health (i.e. number of life-years) they would be willing to sacrifice to avoid poorer health states. The scoring algorithm, or social tariff, for the UK is based on TTO responses of a random sample ($n = 2997$) of non-institutionalised adults. Values are anchored by '1' representing full health and '0' representing the state 'dead', with states 'worse than death' bounded by '−1'. Utility values from the UK EQ-5D tariff range from −0.59 to 1.[15] The EQ-5D is often administered with the EQ visual analogue scale (EQ-VAS), which asks respondents to indicate their health state on a rating scale from worst health imaginable to best imaginable.

TABLE 1 EQ-5D classification

Dimension	Level	Description
Mobility	1	No problems walking about
	2	Some problems walking about
	3	Confined to bed
Self-care	1	No problems with self-care
	2	Some problems washing or dressing self
	3	Unable to wash or dress self
Usual activities	1	No problems with performing usual activities (e.g. work, study, housework, family or leisure activities)
	2	Some problems with performing usual activities
	3	Unable to perform usual activities
Pain/discomfort	1	No pain or discomfort
	2	Moderate pain or discomfort
	3	Extreme pain or discomfort
Anxiety/depression	1	Not anxious or depressed
	2	Moderately anxious or depressed
	3	Extremely anxious or depressed

Source: Dolan 1997.[15]

SF-36 and SF-12

The SF-36 is a generic health status profile measure consisting of eight dimensions of general health, bodily pain, physical functioning, role-physical, mental health, vitality, social functioning and role-emotional (*Table 2*).[19,20] These eight dimensions produced separate scores by taking a simple summation of item responses and applying a linear transformation to place them onto a 0 to 100 scale. There is also an alternative normalised scoring system generating a mean normative score of 50, with 10 points representing a standard deviation (SD) in a normal population. Physical and mental health component summary scores (PCS and MCS) can also be generated from these dimension scores.[10] The SF-12[11] is a shortened version of the SF-36, containing 12 SF-36 items, and also produces PCS and MCS scores. These measures are two of the most widely used generic measures of health and have been validated in many conditions and settings. However, the SF-36 and SF-12 cannot be used in their standard form in economic evaluation because the scoring method is not preference-based and, therefore, cannot be used to generate QALYs. Furthermore, they produce a profile of scores and cannot be combined with survival, so they provide no single measure of effectiveness required for assessing cost-effectiveness. For this reason, a preference-based measure called the SF-6D has been derived from the SF-36.

SF-6D

The SF-6D provides a means of translating the SF-36 or the SF-12 into a preference-based single index.[16,21] The SF-6D reduces the eight dimensions of the SF-36 to six: physical functioning, role limitations, social functioning, pain, mental health and vitality (*Table 3*). Each dimension has four, five or six levels, giving a total of 18,000 possible health states. The values attached to each level and dimension generated by the

TABLE 2 Content of the SF-36

Dimension	No. of items	Summary of content	No. of response choices	Range of response choice
Physical functioning	10	Extent to which health limits physical activities such as self-care, walking, climbing stairs, bending, lifting, and moderate and vigorous exercises	3	'Yes limited a lot' to 'no, not limited at all'
Role limitations – physical	4	Extent to which physical health interferes with work or other daily activities, including accomplishing less than wanted, limitations in the kind of activities, or difficulty in performing activities	2	Yes/No
Bodily pain	2	Intensity of pain and effect of pain on normal work, both inside and outside the home	5 and 6	'None' to 'very severe' and 'not at all' to 'extremely'
General health	5	Personal evaluation of health, including current health, health outlook, and resistance to illness	5	'All of the time' to 'none of the time'
Vitality	4	Feeling energetic and full of life vs. feeling tired and worn out	6	'All of the time' to 'none of the time'
Social functioning	2	Extent to which physical health or emotional problems interfere with normal social activities	5	'Not at all' to 'extremely' and 'All of the time' to 'none of the time'
Role limitations – emotional	3	Extent to which emotional problems interfere with work or other daily activities, including decreased time spent on activities, accomplishing less and not working as carefully as usual	2	Yes/No
Mental health	5	General mental health, including depression, anxiety, behavioural–emotional control, general positive affect	6	'All of the time' to 'none of the time'

Taken from Brazier J, Deverill M, Harper R, Booth A. A review of the use of health status measures in economic evaluation. *Health Technol Assess* 1999;**3**(9).[20]

TABLE 3 The SF-6D

Dimension	Level	Description
Physical functioning	1	Your health does not limit you in *vigorous activities*
	2	Your health limits you a little in *vigorous activities*
	3	Your health limits you a little in *moderate activities*
	4	Your health limits you a lot in *moderate activities*
	5	Your health limits you *a little in bathing and dressing*
	6	Your health limits you *a lot in bathing and dressing*
Role limitations	1	You have *no* problems with your work or other regular daily activities as a result of your physical health or any emotional problems
	2	You are limited in the kind of work or other activities as a result of your physical health
	3	You accomplish less than you would like as a result of emotional problems
	4	You are limited in the kind of work or other activities as a result of your physical health and accomplish less than you would like as a result of emotional problems
Social functioning	1	Your health limits your social activities *none of the time*
	2	Your health limits your social activities *a little of the time*
	3	Your health limits your social activities *some of the time*
	4	Your health limits your social activities *most of the time*
	5	Your health limits your social activities *all of the time*
Pain	1	You have *no* pain
	2	You have pain but it does not interfere with your normal work (both outside the home and housework)
	3	You have pain that interferes with your normal work (both outside the home and housework) *a little bit*
	4	You have pain that interferes with your normal work (both outside the home and housework) *moderately*
	5	You have pain that interferes with your normal work (both outside the home and housework) *quite a bit*
	6	You have pain that interferes with your normal work (both outside the home and housework) *extremely*
Mental health	1	You feel tense or downhearted and low *none of the time*
	2	You feel tense or downhearted and low *a little of the time*
	3	You feel tense or downhearted and low *some of the time*
	4	You feel tense or downhearted and low *most of the time*
	5	You feel tense or downhearted and low *all of the time*
Vitality	1	You have a lot of energy *all of the time*
	2	You have a lot of energy *most of the time*
	3	You have a lot of energy *some of the time*
	4	You have a lot of energy *a little of the time*
	5	You have a lot of energy *none of the time*

classification system were derived from standard gamble valuations for a sample of 249 of these health states. Face-to-face interviews were conducted with a representative sample of 611 members of the UK population.[16] States were valued by asking respondents to choose between each of five SF-6D states (imagining they remained in those states for the rest of their lives), versus a gamble between the best and 'pits' health states. Respondents were then asked to value the worst state in relation to immediate death. The valuations for the SF-6D were derived from a linear random-effects model, and ranged from 0.29 to 1.0 (full health) (see www.shef.ac.uk/scharr/sections/heds/mvh/sf-6d).[16]

What is 'quality of life'?

The question addressed in this report is whether the EQ-5D or SF-6D are able to reflect the way mental health problems impact on the quality of people's lives. Before addressing this question it is helpful to consider more broadly the way quality of life can be defined. Quality of life has been described as an amorphous concept for which there is no widely accepted theory or measurement instrument.[22] The area is bedevilled by the absence of any accepted gold standard. There are many different views or models of quality of life and the more influential ones in health include: objective indicators, needs satisfaction, subjective well-being, capabilities, psychological health and subjective health or HRQoL. These views are not independent and often overlap. This section provides a brief description of these different accounts in order to understand where EQ-5D and SF-36 fit into the literature, and to provide some background for interpreting some of the findings presented from the qualitative research reported in *Chapters 5* and *6*.

Objective indicators

A tension exists in quality-of-life measurement over whether it should have a subjective or objective orientation. A more objective approach may place its emphasis on income, living conditions, access to resources, participation in occupational and social roles,[23,24] or the presence or absence of a medical condition or symptom. While objective measures have an important place in the broader quality-of-life literature, within health there has been an increasing emphasis on the importance of the patient's perspective and this has been assumed to imply a move away from objective measures. This has been partly because many of the commonly used objective indicators like income have been found to be only weakly related to well-being. It is also because objective indicators, by their nature, take a top-down, paternalistic approach, rather than reflecting what individuals might perceive to be important to their quality of life. There has been a similar movement in health care, with a move away from using clinical indicators to measure the outcome of health care towards a more user-focused approach using subjective measures of health and well-being.

Needs satisfaction

There has been a tradition of defining quality of life in terms of the extent to which human needs are met. This is based on Maslow's[25] hierarchy of needs necessary for human existence (e.g. food, drink and shelter, social and belonging, status and self-esteem and self-actualisation). Once the basic needs for food, drink and shelter are met, then human beings can look to fulfil higher needs such as control, autonomy, pleasure and self-realisation.[26,27] These have been found to correlate highly with life satisfaction. This approach continues to be influential in the development of measures, including recent work on an outcome measure for social care.[28]

Subjective well-being

It can be argued that in the developed world, where basic needs have been met, a focus on objective or needs-based measures should be replaced by perceived well-being.[29] Economists in the 19th century saw utility as a cardinal indicator of happiness and this comes from the utilitarian Benthamite tradition concerned with maximising the happiness of the greatest number.[30] There has been something of a revival in using happiness and other measures of well-being in public policy in psychology and economics, such as through the work of Kahneman and others.[31] In its modern formulation, subjective well-being is seen to have two components. One is a hedonistic view based on how an individual feels in any given moment. It has typically been assessed using simple items asking people to rate their current level of happiness.[32]

The second is a more evaluative concept based on a reflection on how satisfied we are with our lives, and this may include past and present happiness and future expectations, and embrace notions of self-fulfilment, realisation or actualisation. This has typically been assessed using items that ask people to rate how satisfied they are with their life and aspects of it.[33]

The revival of subjective well-being in economics has in part been a response to criticisms of measures of benefit used by economists in economic evaluations such as the QALY. By the start of the 20th century, mainstream economics had moved away from an experiential view of utility to one based on preference satisfaction.[34] This is best seen in the marketplace where individuals make choices between goods and services on the basis of what they most desire. They are regarded as autonomous and perfectly informed individuals who will choose the mix of goods and services that maximises their well-being. Where such 'revealed' preferences are not available, as is often the case for health, respondents are asked to state strength of preference for different states of the world. This is a preference-based or decision-based notion of utility. However, there is a large body of evidence that consumers do not necessarily know what will maximise the well-being that they eventually experience.[31] While this debate has not been important in the general quality-of-life literature, it has been highly controversial in health economics since the EQ-5D and SF-6D are scored using the preferences of the general public.[35]

Capabilities

The notion of capability was developed by Sen[36] in response to what he regarded as the narrow perspective of economists in the way they assessed value. He argues that subjective well-being and preference-based utility both fail to consider all the factors that matter for informing public policy. Again, this is an economics-focused debate, but one that has important implications later in this report for the interpretation of qualitative data on what seems to matter to people with mental health problems (see *Chapters 5* and *6*).

Sen[36] offers a framework for moving from the purchase of goods and services (including health care) to utility or well-being. He proposes that along the way individuals are transformed by the benefits of the goods or service; for example, a bicycle will confer improved functioning in terms of mobility.[36,37] Economics usually ignores the impact of what a person has on a person's functioning, and focuses on utility as reflected in their choices. Sen takes it one step further for policy-making, as he argues that public policy is not concerned with functioning per se; so, in the case of the bike, public policy is concerned not with whether or not the person chooses to use the bike, but simply the fact that they are *able* to function in a particular way. This idea of capability offers someone the choice to be able to 'do' or to 'be'. The object of public policy is to provide opportunities and not to make people do one thing or another. Sen accepts that for some basic capabilities this distinction between capability and functioning may not be important as individuals have little choice, but in large areas of public policy it will be important to separate out these concepts.[36]

Sen did not prescribe a particular set of capabilities or a way to operationalise the concept. There have been attempts by others, with the most common being attempts by experts to construct a list of those functionings they think are important, covering basic capabilities through to higher-level issues around human rights and well-being.[38] These tend to be more like the objective lists described earlier and there has been little attempt to score them. In health economics, examples of capability measures include the recently developed ICECAP-O and ICECAP-A, though these use stated preferences to generate the scores. The items were developed from interviews with members of the general public, and attempts to capture the idea of capability by using terms such as 'I am able to be . . .' or 'I can have . . .'.[39,40] Although there may be doubts about whether or not this is truly a measure of capability (including the way it uses preferences in the scoring), it is an important development in terms of going beyond health. This measure will be considered again in the light of the findings of this research in *Chapter 6*.

Psychological well-being

This is a broad category of constructs that covers aspects of personality that include morale, self-esteem and self-concept, sense of coherence, self-efficacy, autonomy and control. This list often includes aspects of subjective well-being, but we have presented them separately as they can be seen as mediating variables rather than necessarily determining quality of life in their own right. Indeed, it has been proposed that there should be a model of quality of life based on these psychological factors and the way they explain variations in subjective well-being through cognitive mechanisms.[41,42] They clearly have an important role in the way people cope with and adapt to problems with their health. Knowledge of the precise mechanisms and the support for different models, however, are limited. It is interesting to note that these variables have also been found in measures of quality of life and capabilities as dimensions in their own right.[27,39,40]

Health and health-related quality of life

There has been a separate tradition of measuring perceived health in health services research and health economics, and the EQ-5D and SF-36 have come out of this tradition.[43] Health can be seen as a part of quality of life, and is usually seen as narrower than quality of life. However, some definitions of health are very broad, such as the widely cited World Health Organization (WHO) statement, set out in its original constitution, that health is: 'A state of complete physical, mental and social well-being, and not merely the absence of disease and infirmity'.[44] This too is a rather amorphous construct. The field is made more complicated by the idea of HRQoL, whereby the concern is not with health per se but those aspects of quality of life that are related to a person's health. Given the central role of health in many people's lives, this is likely to cover many different aspects of quality of life. A definition of HRQoL provided by researchers, representing what many people in the field mean by HRQoL, is: 'A person's subjective perception of the impact of health status, including disease and treatment, on physical, psychological and social functioning and well-being'.[45] This is also a broad definition, and as a result, the content of HRQoL measures varies widely and is no less wide-ranging in coverage than many measure purporting to measure quality of life; indeed, they are often wider.

Discussion of views of quality of life and their implications for this report

This section has demonstrated that there are many different and often competing views of quality of life: objective versus subjective accounts; what people do and feel versus capability; well-being versus psychological well-being (e.g. self-esteem, autonomy, choice and so forth); and narrow versus broader notions. No measure can capture all these elements and ultimately there is a value judgement to be made regarding what matters to policy-makers. It is also the case that these views are not entirely separable, so, for example, needs satisfaction covers objective aspects alongside those concerned with psychological well-being. In practice, measures often contain items from more than one of these views of quality of life. The EQ-5D, for example, covers negative aspects of well-being (i.e. depression and anxiety) and objective aspects such as mobility and self-care. Items of the SF-36 cover well-being (i.e. mental health and vitality), objective aspects (i.e. physical function) and psychological health (i.e. general health perception).

The lack of an agreed view or clear definition means that any attempt to measure quality of life and mental health-related quality of life is fraught with difficulty. This rules out the application of criterion validation in this field, whereby a measure is compared to some accepted gold standard. This has been a source of frustration for researchers and has led to some scepticism as to the value of trying to measure quality of life.[46] However, the primary aim of health services is to improve the quality (health related) and quantity of life, so it is imperative for policy-relevant research to have measures of what matters in the lives of people who use health services. The aim of this report is to examine the extent to which the EQ-5D and SF-36 achieve this through the application of qualitative and quantitative methods.

The appropriateness of the EQ-5D and SF-36 in mental health

Measures of health or HRQoL can be generic, or specific to a patient group defined in various ways (e.g. by a medical condition). The EQ-5D and SF-36 are generic measures that are intended to reflect a core set of domains relevant to all groups of patients (such as mobility, self-care or pain). The more specific measures cover a narrower set of domains and tend to be more symptoms-orientated. Mental health research has tended to be dominated by measures specific to mental health, such as the Patient Health Questionnaire (PHQ-9), Generalised Anxiety Disorder Assessment (GAD-7), General Health Questionnaire (GHQ-12), Beck Depression Inventory (BDI), Lancashire Quality of Life Scale, Health of the Nation Outcome Scales (HoNOS), Clinical Outcomes in Routine Evaluation – Outcome Measure (CORE-OM), Camberwell Assessment of Need (CAN) and many more. Generic measures of health are rarely used in mental health research and this has been seen as a major limitation to conducting economic evaluations in mental health.[1] This has been either because these measures are not regarded as valid in the patient group or investigators are keen to keep patient burden down and prefer to use a more specific measure.

There has been some scepticism about the appropriateness of the EQ-5D in mental health, given that it has only one dimension specifically addressing mental health problems,[11,47–49] though some have argued it is appropriate in common mental health conditions like depression.[50] Recent years have seen the development of preference-based condition-specific measures,[51] including one for common mental health problems.[52] However, the pressure to use generic measures remains and has started to impact on the scales used in clinical studies and, most recently in the UK, in routine patient-reported outcome measurement. The appropriateness of a measure depends on meeting criteria such as reliability, validity and responsiveness. This report is concerned with bringing together quantitative and qualitative evidence on these criteria.

Assessing appropriateness

This report brings together important new evidence on the appropriateness of the two most widely used generic preference-based measures in mental health in the UK (the EQ-5D and SF-6D), and assesses whether or not it is necessary to adapt them in some way, or consider developing a new preference-based measure for mental health. It uses a combination of quantitative and qualitative evidence within the psychometric framework to assess the validity of the measures in mental health populations.[53,54]

Validity

Validity is defined as the extent to which an instrument measures what it is intended to measure and is the core criterion for assessing psychometric performance. The problem for assessing validity is that there is not a gold standard for comparison, as the concept of quality of life, and, therefore, mental health-related quality of life, is fundamentally subjective. Nonetheless, various indirect ways of establishing validity have been developed in the health field and elsewhere, including content validity and construct validity.[55,56]

Content validity 'refers to the degree to which the content of the items reflects the content of the domains of interest'.[57] There are quantitative methods for assessing content validity, based on the extent of agreement between experts regarding the extent to which an item taps the domain. However, these techniques presuppose that the domains are known, which may not always be the case. To address this gap, researchers in the health field have increasingly been applying qualitative methods to obtain the views of people with a given condition to develop and validate the content of the instrument.[56] Indeed, the U.S. Food and Drug Administration[58] regards this as an essential step in the process of validating a patient-reported outcome measure. Qualitative methods provide a depth of understanding that is rarely possible with conventional psychometric techniques. We obtained evidence on the appropriate content of a measure of the impact of mental health problems on quality of life by reviewing qualitative studies of people with mental health problems using in-depth interviews and focus groups. The findings were further tested by conducting in-depth interviews with mental health service users in order to obtain data on a broader range of mental health problems and to refine the emerging themes.

The more conventional approach to validation is to examine construct validation. This can be broken down into two types of tests: discriminant or known-group differences validity and convergent validity. Known-group validity requires data sets with indicators or measures of severity that can be used to define the groups. It should be noted that the usefulness of these tests is limited by sample size (especially as studies are usually not powered on a preference-based measure), the appropriateness of the clinical groups defined, and exogenous factors that may influence quality of life. For instance, groups may be defined by specific clinical symptoms such as mania, but whether or not these result in differences in quality of life is less clear. In mental health there is also the problem of comorbidities, whereby differences between groups may arise due to another condition (such as physical health problems or depression) rather than the one being considered (e.g. schizophrenia).

The other subset of construct validity examined in this report, convergent validity, is defined as the extent to which one measure correlates with another measure of what is purported to be a similar concept. This can be the extent to which generic measures correlate with each other and/or with measures of mental health, but again this depends on the extent to which the latter reflects genuine differences in quality of life.

Responsiveness

Responsiveness is the ability of an instrument to measure clinically significant changes in health over time and can be seen as another form of validity. Responsiveness is usually assessed statistically using measures such as the 'effect size' (where the mean change in score is divided by either the SD at baseline or the SD of the change). A common assumption in the assessment of responsiveness is that for a given health change, the measure with the largest effect size is the better measure.[59] Where the objective is to minimise sample size, this makes sense and maximises the ability to detect significant differences. For economic evaluation, this is less important and, instead, it is a case of establishing that the descriptive content is able to detect change of significance to service users. Related to this is the concept of precision, which is concerned with the ability of a measure to detect changes over the range experienced by patients being assessed.[53] This requires the items and levels of the domains to be well spread over the range of the measure.

The problem for assessing validity and the related responsiveness is that it is rarely possible to prove that a measure is valid or that it is not valid. The nature of the evidence is rarely that conclusive. It is more a question of bringing together evidence, both quantitative and qualitative, that either supports or refutes the claim of validity in comparison to other measures and indicators used in the field.

Outline of report

The aim of this research is to assess the appropriateness of the EQ-5D and the SF-36 (and its derivatives the SF-12 and the preference-based SF-6D) in people with mental health problems, in terms of their validity and responsiveness. These criteria are assessed using evidence from a range of different sources: the literature, a number of existing data sets and interviews with users. This research considers all the main groups of people with mental health problems, including those with common mental health problems (e.g. mild to moderate depression, anxiety, obsessive compulsive disorder and panic disorders), severe and complex non-psychotic disorders (e.g. personality disorder), and schizophrenia and other psychotic disorders. The ways the groups are defined is to a large extent determined by the literature being reviewed, the data available or, in the case of the interviews, by the way services for people with mental health problems are organised. Although we must adopt diagnostic and service groupings in parts of the report, such groupings are not mutually exclusive and, furthermore, many users have more than one diagnosis.

In all, there were four studies undertaken to examine the appropriateness of the measures, and a fifth study to estimate mapping functions between the EQ-5D and SF-6D and commonly used mental health-specific measures. *Chapters 2* and *3* are concerned with the quantitative evidence on the construct validity (i.e. known-group differences and convergent validity) and responsiveness, with *Chapter 2* presenting a systematic review of the existing literature on the EQ-5D and SF-36 across the main mental health conditions and *Chapter 3* presenting new psychometric analyses of existing data sets. For those situations where the EQ-5D and/or SF-6D were found to be acceptable in terms of their psychometric properties, the next stage was to estimate mapping functions between those mental health-specific measures widely used in clinical studies and the two generic measures. This is presented in *Chapter 4*. *Chapters 5* and *6* present qualitative evidence on the domains of quality of life that appear to be important to people with mental health problems in order to shed light on the validity of the content of these generic measures. *Chapter 5* presents a review of previous qualitative research; although this evidence was helpful in starting to identify key themes on the impact of mental health problems on quality of life, it was limited in terms of the groups covered and by the topic guide used by the researchers. Therefore, in-depth interviews were carried out with mental health service users and a thematic analysis undertaken, and this is presented in *Chapter 6* in order to provide the basis for reviewing the appropriateness of content of the generic measures.

Chapter 7 presents a brief overview of the findings of each study before discussing the main findings of this research, and presents the recommendations for further research and implications for policy.

Chapter 2 A systematic review of the validity and responsiveness of the EQ-5D, SF-36, SF-12 and SF-6D in mental health

This chapter examines the validity and responsiveness of two generic preference-based measures of health (the EQ-5D and SF-6D) and two related generic non-preference-based measures (the SF-36 and SF-12) in populations with mental health problems. The assessment is based on a systematic review of studies reporting one or more of these measures alongside various condition-specific indicators of mental health that can be used to assess validity using known-group comparisons and convergence, and responsiveness to changes in health over time. It forms the first study presented in this report.

This review covers five mental health conditions: schizophrenia, bipolar disorder, personality disorders, depression and anxiety. Four separate systematic reviews were undertaken from one common search of the literature, with depression and anxiety reviewed together. This chapter presents the methods and an overview of the findings. The detailed findings with tables of results by study are available in published articles or in discussion paper form.[60–62]

Methods

Selecting review studies

Inclusion and exclusion criteria

Studies were eligible for inclusion if they contained HRQoL data obtained using one or more of the instruments under study (SF-36, SF-12, SF-6D or EQ-5D) within the specified population of adults (aged ≥ 18 years) suffering from one of the five conditions. HRQoL data could be from descriptive systems (i.e. their items and dimensions) or health-state utility values generated by the EQ-5D or SF-6D, or the EQ-VAS. Studies whose primary focus was on individuals with alcohol and/or drug dependency were excluded whether or not those individuals had one of the five conditions. The outcomes had to include data that allowed measurement of the construct validity (i.e. known groups, convergent or discriminant) or the responsiveness of the HRQoL instrument(s). Responsiveness data had to be in the form of effect sizes, standardised response means (SRMs) or correlation with change scores on symptom measures. Studies that only provided data on other psychometric properties, such as reliability, face validity and content validity, were not included.

Identification of studies

A literature search was performed to identify relevant research for all mental health conditions being investigated within the wider review, using a database thesaurus and free text terms. Two sets of search terms were combined: terms for each of the four HRQoL measures and terms for each mental health condition (the search strategy is in *Appendix 2*). Ten databases were searched for published research from inception: Cochrane Database of Systematic Reviews (CDSR), Cochrane Central Register of Controlled Trials (CENTRAL), NHS Economic Evaluation Database (NHS EED), Health Technology Database, Database of Abstracts of Reviews of Effects (DARE), MEDLINE, PreMEDLINE, Cumulative Index to Nursing and Allied Health Literature (CINAHL), EMBASE and Web of Science. Searches were limited to English language only but not by any date restriction. All searches were initially conducted in August 2009, though updates were undertaken for personality and bipolar disorders until March 2011, and for depression and anxiety until December 2010. The reference lists of relevant studies were searched for further papers.

Citations identified by the searching process were screened by one reviewer (DP or TP) using the inclusion criteria. The full texts of papers were retrieved for any titles or abstracts that appeared to satisfy the inclusion criteria, or for which inclusion or exclusion could not be definitely determined. The same inclusion and exclusion criteria were used to assess full-text papers and any queries over inclusion were resolved by discussion and consensus between two reviewers (DP and JB or TP and JB).

Data extraction

Data from all included trials were extracted using a form designed specifically for this review, and piloted on a sample of papers. Data extracted included: country of publication, type of disorder, study sample characteristics (numbers, age, sex), other measures used, mean scores on HRQoL measures, type and method of validity assessment, type and method of responsiveness assessment, and validity and responsiveness data. Extractions were performed by one reviewer (DP or TP). Where duplicate publications reported on similar data, the most complete and recent data were extracted.

Quality assessment

There is no formal method for assessing the quality of these studies (i.e. there are no quality assessment checklists), and thus we used the methods described by Fitzsimmons *et al.*[63] to evaluate HRQoL data in their systematic review on the use and validation of HRQoL instruments with older cancer patients. These included examining whether or not tests of statistical significance were applied, differences between treatment groups reported (where applicable, e.g. in known-group validity), clinical significance discussed and missing data documented. We also report on response and completion rates where these were provided.

Evidence synthesis and meta-analysis

Owing to the large degree of heterogeneity between studies (including types of study designs, HRQoL measure, population characteristics and methods of determining construct validity and responsiveness), it was not appropriate to perform meta-analysis. Analysis was by narrative synthesis and data were tabulated. All analyses were performed based on the HRQoL measure, with data analysis grouped by type of validity (convergent/discriminant or known groups) or responsiveness test used.

Defining validity and responsiveness

Validity and responsiveness were assessed using the definitions presented in *Chapter 1*. For convergent validity, the strength of correlation between the two measures is calculated using statistical tests (e.g. Pearson's product moment correlation or Spearman's rank correlation). We have used the following categories for evidence of correlation: $> 0.6 =$ very strong, ≥ 0.5 to $< 0.6 =$ strong, < 0.5 to $\geq 0.3 =$ moderate, and $< 0.3 =$ weak. Statistical significance is also attached to correlations ($p < 0.05$). Responsiveness can be measured in a number of ways by effect size statistics[64] standardised in different ways, such as dividing through by the SD at baseline or SD of the change in scores over time (i.e. SRMs). Within this review, Cohen's[65] categories for magnitude of effect size were used: $\geq 0.80 =$ large, < 0.80 and $\geq 0.50 =$ moderate, and 0.30 to $< 0.50 =$ small. As pointed out in *Chapter 1*, these tests need to be used with some care as there is no gold standard and the application of these tests sometimes uses indirect indicators of the concept (e.g. symptoms rather than quality of life).

Study characteristics

The initial search for studies for the wider review retrieved 4115 unique citations across the five mental health conditions (*Figure 1*). Of these, 3849 were excluded at the title and abstract stage and 266 were examined in full text. From these, 154 studies were found that met the inclusion criteria. A further 12 studies were identified through reference list checking. Overall, the findings from 91 studies are discussed in this chapter for the five conditions. SF-36 and SF-12 studies were not ultimately included in the depression and anxiety review as a sufficiently large number of studies used the SF-6D to be able to extrapolate to these longer versions of the measure. *Figure 1* shows a flow diagram of study identification and the characteristics of the studies reviewed by condition are presented in *Appendix 2*.
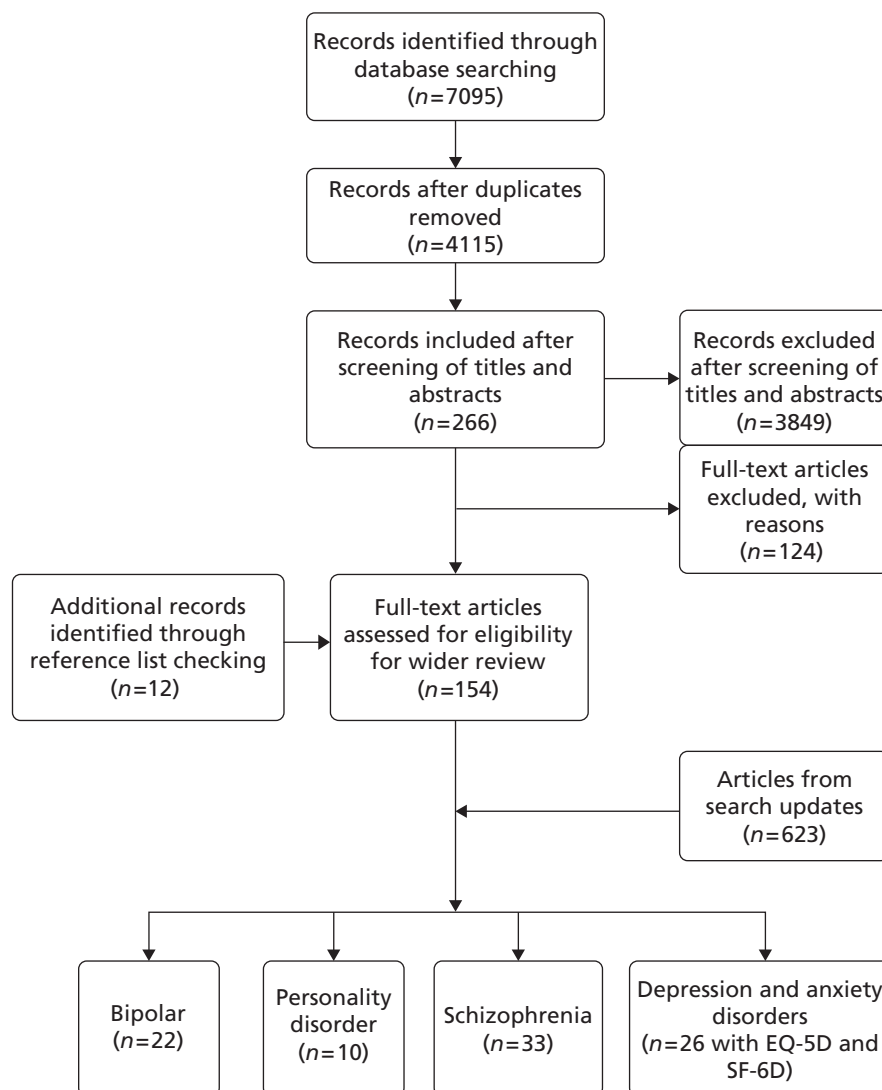
FIGURE 1 Flow diagram of study identification.

Schizophrenia

Thirty-one studies were identified that provided data on the validity and/or responsiveness of the EQ-5D, SF-36, SF-12 or SF-6D within individuals diagnosed with schizophrenia, schizophreniform disorder or schizoaffective disorder (see *Appendix 2*, *Table 32*).[48,66–95] Six studies were undertaken internationally across more than one country.[66–71] The numbers of participants in the studies with schizophrenia or related conditions ranged from 15 to 2657. Participants included males and females with a mean age of participants with a schizophrenia spectrum disorder, reported in 21 of the 33 studies, ranging between 20.3 and 57.9 years.[48,68,70–73,75–80,82,84,86,88–91,94,95]

All studies obtained HRQoL information from patients; seven of these studies compared patient HRQoL values with published general population 'normative' values,[70–76] three compared HRQoL values with normal comparison participants that were recruited to the study[77–79] and two used 'norms' from healthy participants who had taken part in large surveys.[80,96]

Bipolar disorder

Twenty-two studies were identified that provided data on the validity and/or responsiveness of the generic HRQoL measures in bipolar disorder (see *Appendix 2*, *Table 33*).[74,97–117] Nineteen studies[74,97,98,102–117] contained data on the SF-36, one study involved the SF-12[101] and four studies[98–101] contained data on the EQ-5D. No studies were identified that examined the SF-6D in individuals diagnosed with bipolar disorder.

The numbers of participants in the studies with bipolar disorder ranged from 30 to 1999. Participants included men and women. The mean age of participants with a bipolar disorder, reported in 19 of the 22 studies,[97–103,106–117] ranged between 29.3 and 60.2 years.

All studies obtained HRQoL information from patients; six of these studies compared patient HRQoL values with 'norms' derived from published general population 'normative' values,[48,76,80,81,96,118] three compared HRQoL values with normal comparison subjects that were recruited to the study[74,79,82] and one study used 'norms' from healthy subjects taking part in large surveys.[83] Four studies investigated differences in HRQoL between mood groups in bipolar disorder.[75,84–86] Two of the four studies investigating the EQ-5D used general population preferences for EQ-5D health states to generate EQ-5D index values.[87,96]

Personality disorder

In total, there were 10 studies reporting HRQoL data on patients with personality disorder.[96,119–128] Six studies looked at the EQ-5D,[120,124–128] two at the SF-36[119,121] and two at the SF-12 (corresponding to three articles).[96,122,123] No studies were found investigating the validity or responsiveness of the SF-6D in this patient group. Studies were undertaken in four countries. Nine[96,119–128] of the 10 studies presented data for different personality disorders together. One study looked exclusively at individuals with borderline personality disorder. The numbers of individuals included within the studies that were diagnosed or screened as having one or more personality disorders ranged from 48 to 1708. Participants included males and females (proportions can be seen in *Appendix 2*, *Table 34*). The mean age of participants with personality disorders, reported in 9 of the 10 studies, ranged between 29.4 and 45 years.[96,119–121,124–128]

Two studies[120,127] investigated the known-group validity of the EQ-5D, one study[127] investigated the convergent validity of the EQ-5D and four studies[124–126,128] investigated the responsiveness of the EQ-5D. Two studies[119,121] investigated the known-group validity of the SF-36 and two studies[96,122,123] investigated this property in the SF-12. One study[119] investigated the responsiveness and convergent validity of the SF-36.

Depression and anxiety

Owing to the large number of studies reporting SF-36 and SF-12 data in this group, it was decided to focus on EQ-5D and SF-6D data. SF-36 and SF-12 are not preference-based and have been included in the other studies to give an indication of the likely performance of the derivative SF-6D. In all, there were 22 studies[50,129–149] identified with data on the validity and/or responsiveness of the generic HRQoL measures in depression and anxiety for EQ-5D and SF-6D. Fourteen studies[50,129,136–139,142–149] had data on the EQ-5D and seven[130,131,133–135,140,141] contained data on the SF-6D. Studies were undertaken in at least 12 countries (a number covered Europe). The numbers of participants with depression and anxiety in these studies ranged from 44 to 3815. Participants included men and women with a mean age between 39.2 and 49 years of age.

Six studies[139,142,144,147–149] investigated the known-group validity of the EQ-5D across severity groups, five studies[129,136,137,143,145] reported convergent validity of the EQ-5D and 14 studies[50,130–138,140,141,145,146] investigated the responsiveness of the EQ-5D. Two studies[139,142] had known-groups differences in the SF-6D, three[132,134,137] had convergent validity and two[136,145] had responsiveness.

Quality of included studies

Most studies reported tests for statistical significance of the properties measured for difference between groups (e.g. known-group validity) and responsiveness to change over time. A minority of studies considered what constituted a clinically significant difference in HRQoL scores, by either providing a predefined value or discussing whether or not the results were clinically meaningful. There was little discussion or inclusion, however, of clinical significance defined in terms of patient perception, and thus, from the perspective of preference-based measures, the lack of patient preference undermines the concept of clinical significance. Most studies did not report missing HRQoL data. This has implications for the representativeness of these samples due to possible selection bias.

Results

Detailed findings with tables of results on the validity and responsiveness have been reported elsewhere.[60–62] This section summarises the results using a simple classification of the evidence: ✓ indicates results in support of validity or responsiveness and ✗ indicates an inconsistent or non-significant result. The results on validity have been divided into known-group differences across severity groups typically defined using symptoms, known differences against a normal case–control group and convergence with a measure of the condition.

Schizophrenia

The majority of the evidence (25 studies) examined the validity and responsiveness of the SF-36.[66,67,69–85,88,92–95] Although there appears to be strong evidence that the SF-36 is able to distinguish between general population norms and scores of people with schizophrenia (known-group validity), the evidence for convergent validity and responsiveness is less certain (*Table 4*). Similar findings exist for the EQ-5D, with mixed evidence for the properties of convergent validity and responsiveness. Indeed, when strong associations were found between individual EQ-5D health-state dimensions (e.g. anxiety/depression or self-care) and symptom or functioning measures, this did not necessarily translate into comparable changes in overall EQ-5D index scores, i.e. health-state utility values.[48,90] There was some evidence that associations with measures of depression were comparatively stronger than those with symptom measures of schizophrenia [e.g. the Positive and Negative Syndrome Scale (PANSS)].[71,88–90]

When testing associations between measures for convergent validity (or change scores in responsiveness), there are reasons to predict that stronger and more consistent correlations might exist between generic HRQoL measures and functioning [e.g. Global Assessment of Functioning (GAF), Social and Occupational Functioning Assessment Scale (SOFAS)] or mental health/schizophrenia-specific HRQoL [e.g. Quality of Life Scale for Schizophrenia (QLS)] measures than purely symptom-based measures such as the PANSS. By their very nature, symptom measures are measuring different concepts from HRQoL measures, so it might be reasonable to predict that it is less likely that a strong correlation might exist. A re-examination of the evidence, taking into account evidence for the type of measure used to assess convergent validity (symptom vs. functioning or HRQoL measures, subjective vs. objective measures), produced mixed results. Functioning and schizophrenia HRQoL measures did not fare much better than clinical and symptom-based measures, with four studies indicating strong evidence for convergent validity,[82–84,86] and four indicating uncertain or no evidence of such a relationship.[69,81,88,89]

Bipolar disorder

There was positive evidence that the SF-36 is a valid and responsive measure in bipolar disorder when individuals are in a depressed, euthymic or mixed state (*Table 5*). There was little evidence available on the EQ-5D and SF-12 and none for the SF-6D. What evidence there is on the EQ-5D is mixed, with three tests supporting the validity of the EQ-5D in this group and three against it across four studies.[98–101]

TABLE 4 Results for schizophrenia

Criterion	EQ-5D	SF-36	SF-12	SF-6D
KGV: severity	✓	✓	None	None
KGV: case–control	None	✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓	✓	None
Correlation	✓ ✓ ✓ ✗ ✗ ✓ ✗ ✓ ✓ ✗ ✓ ✓ ✓ ✓ ✗ ✗	✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✗ ✗ ✗ ✗ ✗	None	✓
Responsiveness	✓ ✓ ✓ ✗ ✗	✓ ✓ ✓ ✗ ✗ ✗ ✗	None	✗

KGV, known-group validity.
✓ indicates results in support of validity or responsiveness; ✗ indicates an inconsistent or non-significant result.
Each ✓ and ✗ represents an individual study.

TABLE 5 Results for bipolar disorder

Criterion	EQ-5D	SF-36	SF-12	SF-6D
KGV: severity	✓ ✓	✓ ✓ ✓ ✗ ✓ ✓	✓	None
KGV: case–control	✗	✓ ✓ ✓ ✓ ✓ ✓ ✓ ✗	None	None
Correlation	✓ ✗ ✓ ✗ ✓	✓ ✓ ✓ ✓ ✓ ✓ ✗	✓	None
Responsiveness	None	✓ ✓ ✓ ✓ ✗	None	None

KGV, known-group validity.
✓ indicates results in support of validity or responsiveness; ✗ indicates an inconsistent or non-significant result.
Each ✓ and ✗ represents an individual study.

It is unclear if these generic measures are valid in manic or hypomanic individuals. Only 7 out of 22 SF-36 studies included individuals in a manic or hypomanic state,[74,99,101,104,106,110,117] and these suggest it is not a valid and responsive measure within this population. However, where studies examined convergent validity with clinical measures of mania, the numbers of patients in the manic or hypomanic mood state were too small to be meaningful. More generally, there is some concern around how to obtain reliable HRQoL ratings within bipolar disorder individuals in manic or hypomanic states as this relies on self-report.

Depression and anxiety

The SF-6D and EQ-5D demonstrate good construct validity and responsiveness for patients with depression (*Table 6*). They can both distinguish between groups that are known to vary according to severity of depression, and across differences in quality of life of depressed patients. Both measures respond to clinical and quality-of-life improvement and deterioration. Indeed, in many cases they are more responsive than depression-specific measures (this may be due to the integrated nature of mental and physical health problems and potential simultaneous improvement in comorbid conditions).

The performance of the EQ-5D for patients with anxiety is a little more mixed. The measures were found to be more highly correlated with depression scales than with clinical anxiety scales in patients with anxiety.

The relationship between the EQ-5D and the SF-6D reflects that found for other conditions. The EQ-5D shows a lower level of utility at the most severe end for depression, and the SF-6D shows equal or greater detriment at the milder end. The SF-6D identifies utility loss in patients that report full health on the EQ-5D, though patient averages for mild depression and anxiety are still able to show lower than normal population utility using the EQ-5D.

Personality disorder

The EQ-5D appears responsive in individuals with personality disorders (*Table 7*). Data on other properties such as convergent and known-group validity were very limited. There was also little evidence on the SF-36 or SF-12 and none on the SF-6D. Nevertheless, the studies which did exist provided some positive evidence that the measures are valid for use in personality disorders.[100,102,107] An exception was Narud *et al.*,[119] who found that most dimensions on the SF-36 were not able to detect changes in patients in the same way as clinical measures. They concluded that this may be because some SF-36 dimensions are not relevant to HRQoL, so that, even if patients change clinically, this does not translate to a change in HRQoL.

TABLE 6 Results for depression and anxiety

Criterion	EQ-5D	SF-6D
KGV: severity	✓ ✓ ✓ ✓ ✗ ✓ ✓	✓ ✓
KGV: case–control	None	✓ ✓
Correlation	✓ ✓ ✓ ✓ ✓	✓ ✓ ✓
Responsiveness	✓ ✓ ✓ ✓ ✗ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓	✓ ✓

KGV, known-group validity.
✓ indicates results in support of validity or responsiveness; ✗ indicates an inconsistent or non-significant result.
Each ✓ and ✗ represents an individual study.

TABLE 7 Results for personality disorders

Criterion	EQ-5D	SF-36	SF-12	SF-6D
KGV: severity	✓	✗	✓ ✓	None
KGV: case–control	✓	✓ ✓	None	None
Correlation	✓	✓	None	None
Responsiveness	✓ ✓ ✓	✓	None	None

KGV, known-group validity.
✓ indicates results in support of validity or responsiveness; ✗ indicates an inconsistent or non-significant result.
Each ✓ and ✗ represents an individual study.

Discussion

This review is the first to have comprehensively identified studies that report on the construct validity and responsiveness of these four generic HRQoL measures, and to tabulate and give a narrative synthesis of the findings. Overall, the evidence suggests that generic HRQoL measures are appropriate for patients with depression and personality disorder, but it is more mixed in relation to anxiety, bipolar disorder and schizophrenia.

The findings for depression are encouraging, but there is a concern that this may be driving the differences between groups found for other conditions. For anxiety, the ability of generic preference-based measures to distinguish between subgroups of patients with anxiety may be driven by aspects of depression within anxiety disorder and the presence of comorbid depression. There was some evidence that associations with measures of depression were comparatively stronger than those with symptom measures of schizophrenia (e.g. the PANSS).[71,88–90] This may indicate that: (a) the generic HRQoL measures were only able to detect this component of HRQoL, or (b) depression is the only component of HRQoL within these groups of patients that is important within the context of HRQoL measurement. The issue is whether schizophrenia or anxiety has quality-of-life implications not adequately described by the five dimensions of the EQ-5D. It is also difficult to predict how HRQoL is affected by the manic or hypomanic states from the perspective of the individual with bipolar disorder. These non-depression consequences of these conditions are explored later in this report through qualitative interviews with patients.

The review has some limitations, resulting from the need to compromise on some elements of the review process because of the large scope of the project. Although the search for studies was reasonably comprehensive, it was limited to key databases and reference list checking of included studies, and study selection was undertaken by one reviewer. Ideally, further searching could be undertaken in trial registries, conference proceedings and by citation searching to make the search more comprehensive in terms of process. Study quality assessment has been pragmatic and focused on the elements that contribute to HRQoL analysis. The populations included in this review were heterogeneous in terms of the nature of their conditions, particularly for conditions such as schizophrenia and personality disorder where there are numerous subgroups, and not all studies provided detailed or uniform information on these characteristics. Such clinical variables clearly have an impact on HRQoL, and thus these factors will have had an impact upon the results of individual studies.

It is also difficult to draw any firm conclusions on the basis of this review, owing to the limited nature of much of the evidence in terms of the number of studies, the size of some of the studies and the heterogeneity within the conditions. There is very limited evidence of validity or responsiveness for the SF-12 and SF-6D, and though these are derivatives of the SF-36, their more limited item coverage (12 and 11, respectively) means that they may not perform as well. Therefore, further research needs to be directed towards demonstrating these properties for these instruments. To improve the evidence base, the next chapter will conduct further psychometric tests on existing data sets containing the EQ-5D and SF-6D. More evidence is also required on the validity and responsiveness of generic measures for older people with depression, as this group may be different from the younger adults typically found in published trials.

There is another general concern regarding whether or not it is reasonable to assess HRQoL when an individual is in a particular state, such as a manic or hypomanic state, as he or she may view the effect that the state had on his or her HRQoL very differently when not actually in that health state.

The findings are also limited by the measures used to establish validity and responsiveness. It is difficult to determine, in theory, how strongly correlated generic HRQoL measures should be with symptom and/or other clinical measures, and there is little guidance on what constitutes reasonable correlation. Indeed, it is impossible to prove validity of HRQoL instruments, as no 'gold standard' exists. Also, as discussed previously, where health dimensions and changes appear to have been missed by preference-based HRQoL measures, these may not actually be important to patients or valued by the general population. The former will be examined in the qualitative research reported in *Chapters 5* and *6*.

The dominance of physical health in the EQ-5D may explain why it is not sensitive to differences in some mental health populations.[48] Although this does not seem to have been a problem in depression, it may account for the more mixed results in other conditions. There are also concerns that the descriptive systems of the generic measures are too narrow in terms of the dimensions they cover. Some of the questions raised are addressed later in this report using the findings of qualitative interviews of people with mental health problems, who can provide some insight into the shortcomings of the content of the descriptions contained in these generic measures (see *Chapters 5* and *6*).

Research needs to be directed towards developing robust methods of demonstrating validity and responsiveness for generic HRQoL measures. For known-group validity, the evidence discriminating between healthy and non-healthy individuals could be considered fairly crude; large differences should be obviously apparent between such groups. Therefore, research is required to test instruments in terms of the ability to reflect known-group differences using indicators of condition severity that are important to patients. For convergent validity, this might mean consideration of which measures to choose for assessment of strength of correlation, both by considering the type of measure (e.g. symptom functioning or HRQoL) and the nature of the measure (subjective or objective). Studies need to be explicit at their outset about the hypothesised associations when investigating validity and responsiveness. In addition, wherever studies investigate the feasibility of administering generic HRQoL measures alongside construct

validity and responsiveness using quantitative and qualitative methods within this disease area, this will allow a greater overall understanding of which measures are useful within schizophrenia.

Conclusion

Despite the shortcomings identified in the evidence base, this review gives an overall picture of the validity and responsiveness of the EQ-5D and SF measures across mental health populations. It has shown a mixed picture, with the generic measures appearing to perform acceptably well in depression and personality disorder, but less well in anxiety, schizophrenia and bipolar disorder. This has highlighted the need for further quantitative research, and the insights that can be gained from people regarding the content validity of the measures in terms of coverage of the dimensions of their life impacted upon by their mental health problems. The following chapters report both quantitative and qualitative studies that further investigate the validity of these measures.

Chapter 3 Assessing the psychometric performance of the EQ-5D and SF-6D using existing data sets

The review of psychometric evidence reported in the last chapter found mixed evidence on the performance of the EQ-5D and SF-36 (and its derivatives). The aim of the study reported in this chapter was to use existing data to supplement the psychometric evidence on both the EQ-5D and the SF-6D in samples of common mental health conditions, schizophrenia and personality disorder. The validity of the instruments and their responsiveness to change over time were assessed in comparison with widely used condition-specific outcome measures that have been validated for use in the mental health population, using existing data sets sourced from a range of studies. These data sets included 3512 people with common mental health problems (including depression and anxiety), 480 people with a mixed diagnosis of moderately severe common mental health problems and personality disorders leading to self-harm, and 1756 people with schizophrenia and other personality disorders. The analysis adds to the evidence base regarding the psychometric performance of the EQ-5D and SF-6D in mental health populations, providing information on the conditions in which the generic measures can be used and those for which improvements to the existing instruments and/or new measures are required.

Methods

Measures used in data sets

The EQ-5D and SF-6D were compared to one widely used and validated condition-specific measure within each data set. The measures used were restricted to those included in the data sets available. *Table 8* displays the measures used in the analysis. Details of the two generic preference-based measures are presented in *Chapter 1*. The condition-specific measures are described below.

Hospital Anxiety and Depression Scale

The Hospital Anxiety and Depression Scale (HADS)[156] is a 14-item self-report measure that contains two seven-item subscales: depression (HADS-D) and anxiety (HADS-A). The total score for each dimension is 21 (items are scored 0–3) with high scores indicative of increased levels of anxiety and depression (a score of 8+ indicates a possible case, and a score of 11+ indicates a probable case). The overall score [Hospital Anxiety and Depression Scale – Total score (HADS-T)] is also used as a measure of global functioning. The HADS has been widely used across clinical groups and research settings,[157] and there is evidence for its psychometric validity in comparison with other mental health[158,159] and generic measures.[160] In this study, HADS was used to assess the psychometric performance of the EQ-5D in two separate samples of people with anxiety and depression.

Clinical Outcomes in Routine Evaluation – Outcome Measure

The CORE-OM[161–165] is a self-reported measure developed in the UK for routine use in psychological services at assessment and outcome. It comprises 34 items addressing domains of subjective well-being, symptoms (anxiety, depression, physical problems, trauma), functioning (general functioning, close relationships, social relationships) and risk (risk to self, risk to others). Items are scored on a 5-point, 0–4 scale. CORE clinical scores are computed as the mean of all completed items multiplied by 10 (range 0–40). This is so that clinically meaningful differences are represented by whole numbers. The psychometric validity of the CORE-OM in comparison to other widely used measures of mental health has been demonstrated.[166,167] In this study, the CORE-OM was used to assess the psychometric performance of the SF-6D in two separate samples of people with anxiety and depression.

TABLE 8 Summary of data sets and measures used for analysis

Data set	Description	Time points used	Generic measure	Condition-specific measure	Characteristics		
					n (baseline)	Age (years), mean (SD)	Female (%)
Common mental health conditions							
Assessing Health Economics of Antidepressants (AHEAD)[150]	RCT of the cost-effectiveness of antidepressants (three time points). Sample includes common mental health concerns – depression, MADD and phobias	Baseline 6 months 12 months	EQ-5D	HADS	327	43.1 (15.4)	67.0
Psychological Interventions for Postnatal Depression (PoNDER)[151]	Study of two psychologically informed interventions for women with postnatal depression	Baseline 6 months 12 months	SF-6D	CORE-OM	2640	31.5 (5.1)	100.0
IAPT cohort study[152]	Evaluation of the outcomes from two IAPT demonstration sites. Sample reports common mental health condition	Baseline 4 months 8 months	SF-6D	CORE-OM	527	40.9 (14.2)	72.3
Common mental health and personality disorders							
The POPMACT study[153]	RCT of MACT vs. TAU in recurrent self-harm. Sample includes patient with diagnoses of common mental health conditions and personality disorders leading to self-harm	Baseline	EQ-5D	HADS	480	32.0 (11.2)	67.9

Data set	Description	Time points used	Generic measure	Condition-specific measure	Characteristics		
					n (baseline)	Age (years), mean (SD)	Female (%)
Schizophrenia							
Quality of Life following Adherence Therapy (QUATRO)[154]	Multicounty RCT of adherence therapy in patients with a clinical diagnosis of schizophrenia who needed continuing antipsychotic medication for at least 1 year from assessment, and had evidence of clinical instability in the year before assessment	Baseline 12 months	EQ-5D SF-6D	BPRS-E	409	41.5 (11.5)	59.9
Multicentre study of Art Therapy In Schizophrenia: Systematic Evaluation (MATISSE)[155]	RCT of group art therapy for people with a clinical diagnosis of schizophrenia	Baseline 12 months	EQ-5D	PANSS	417	41.0 (11.5)	33.3
Schizophrenia and personality disorders							
Study of Cost-Effectiveness of Personality disorder TREatment (SCEPTRE)[120]	Dutch study of adult patients with personality disorders (UK EQ-5D tariff used for comparability)	Baseline 12 months	EQ-5D	DSM-IV personality disorder category	932	35.1 (9.8)	68.1

BPRS-E, Brief Psychiatric Rating Scale – Expanded; DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders – Fourth Edition*; HADS, Hospital Anxiety and Depression Scale; MACT, manual-assisted cognitive therapy; MADD, mixed anxiety and depressive disorders; RCT, randomised controlled trial; TAU, treatment as usual.

Brief Psychiatric Rating Scale – Expanded

The Brief Psychiatric Rating Scale (BPRS)[168] was developed to assess symptom change in psychiatric inpatients and is one of the most widely used measures to assess psychotic and affective symptoms. The current study used the expanded version, the Brief Psychiatric Rating Scale – Expanded (BPRS-E), which has 24 items developed for use in schizophrenia patients. The BPRS-E is administered using semistructured interviews and each of the 24 items has a possible score of 1 (not present) to 7 (extremely severe). Total scores range from 24 to 168 with higher scores indicating higher severity. The BPRS has been shown to be highly sensitive to change in psychiatric patients.[169] In this study, the BPRS-E was used to assess the psychometric performance of both the EQ-5D and SF-6D in a sample of people with schizophrenia.

Positive and Negative Syndrome Scale

The PANSS was developed to evaluate positive, negative and other symptom dimensions in schizophrenia by combining the 18 items of the BPRS with the 12 items of the Psychopathology Rating Schedule with detailed instructions on completion by interview. It has 30 items which are scored from 1 (absent) to 7 (extreme) and result in 3 subscales: positive, negative and general psychopathology.[170,171] The PANSS is sensitive to changes in both positive and negative symptoms in patients.[169] The PANSS was used to assess the performance of the EQ-5D in a sample of people with schizophrenia.

Data sets

Identification of data sets

The systematic reviews reported in *Chapter 2* examined the performance of generic preference-based measures in depression, anxiety and schizophrenia, and these reviews identified 91 studies. Corresponding authors of these studies were contacted and the data sets were requested for use in this study, and publicly available data sets were sourced. In total, 12 data sets were received (13% of those requested), and these were reviewed for acceptable condition-specific comparison measures or clinical indicators. Three data sets were excluded as they focused on general population samples, and two were excluded as they did not include a comparison measure of interest.

Seven data sets were selected for use in this study. Three included common mental health problem samples [Assessing Health Economics of Antidepressants (AHEAD), Psychological Interventions for Postnatal Depression (PoNDER) and IAPT]; one included mixed common mental health and personality disorder diagnoses (POPMACT); two included schizophrenia samples [Quality of Life following Adherence Therapy (QUATRO) and Multicentre study of Art Therapy In Schizophrenia: Systematic Evaluation (MATISSE)]; and one included schizophrenia and personality disorder [Study of Cost-Effectiveness of Personality disorder TREatment (SCEPTRE)] (see *Table 8*).

Common mental health problems data sets

The AHEAD study[150] was a prospective randomised controlled trial (RCT) of the cost-effectiveness of tricyclic antidepressants, selective serotonin reuptake inhibitors (SSRIs) and lofepramine. The sample ($n = 327$) consisted of patients recruited in a primary care setting who reported a range of common mental health concerns including depression, mixed anxiety and depressive disorders (MADD) and a number of phobias. The EQ-5D and HADS were collected at baseline and at six follow-up time points. In this study, baseline and two follow-up time points (6 months and 12 months) have been considered.

The PoNDER trial,[151] which assessed two psychologically informed interventions for women with postnatal depression, was also selected. The sample used is those patients who completed the study at all time points ($n = 2640$). SF-6D and CORE-OM data were collected at baseline, 6 months, 12 months and 18 months, with the first three time points considered in this study.

The IAPT[152] cohort study ($n = 527$ patients), which examined access to psychological therapies and associated outcomes for patients with mild to moderate depression and anxiety, was also selected. SF-6D and CORE-OM were administered at baseline, 4 months and 8 months.

Mixed common mental health and personality disorder diagnoses data set

The POPMACT study[153,172] was a RCT of manual-assisted cognitive therapy versus treatment as usual in recurrent self-harm. The baseline sample was used in this study and consisted of patients presenting with recurrent deliberate self-harm ($n = 480$).

Schizophrenia data sets

The MATISSE study was a three-arm randomised trial of group art therapy for people with schizophrenia.[155] Study participants ($n = 409$) were recruited from inpatient and community-based mental health and social care services at four centres in England and Northern Ireland. Participants were > 18 years of age with a clinical diagnosis of schizophrenia. Assessment was at baseline and at 12 and 24 months, and EQ-5D and PANSS were collected at all time points. Analysis in this study focuses on baseline and 12 months.

The QUATRO study was a two-arm RCT of adherence therapy in patients with schizophrenia.[154] Participants ($n = 417$) were recruited from four countries (the Netherlands, Germany, England and Italy) and were included if they had a clinical diagnosis of schizophrenia [based on *International Classification of Diseases*, Tenth Edition (ICD-10) criteria], needed continuing antipsychotic medication for at least 1 year from assessment, and had evidence of clinical instability in the year before assessment. EQ-5D and BPRS-E data were collected at baseline and 12 months after randomisation, and these data were used in the analysis.

Schizophrenia and personality disorder data set

Study of Cost-Effectiveness of Personality disorder TREatment study[120] participants ($n = 932$) were recruited from a consecutive series of admissions to six mental health care institutes in the Netherlands offering outpatient, day hospital and/or inpatient psychotherapy for adult patients with personality pathology and/or personality disorders. Patients with clear signs of unreliable or incomplete data and those with serious intellectual impairment were excluded. Overall, 88% had at least one personality disorder based on the Dutch version of the Structure Interview for DSM-IV (*Diagnostic and Statistical Manual of Mental Disorders – Fourth Edition*) Personality. The data set consists of EQ-5D baseline assessments and follow-up at 6, 12, 24 and 36 months. We used the UK EQ-5D tariff to allow for comparability with the other samples. Analysis was limited to baseline and 12-month follow-up.

Demographic characteristics available across the data sets are displayed in *Table 8*. *Table 9* displays the scores for each sample on the measures used in the analysis for those who returned either the generic preference-based measure or the condition-specific instrument. The POPMACT sample has lower EQ-5D index scores than the AHEAD sample, and the mean HADS scores also indicate that the POPMACT sample displays higher levels of anxiety and depression. SF-6D and CORE-OM scores indicate that the IAPT sample displays lower levels of health than the PoNDER sample. For the schizophrenia and personality disorders groups, baseline EQ-5D scores indicate that those in the MATISSE and QUATRO samples have similar health levels, whereas those in the SCEPTRE sample display lower health as measured by the EQ-5D.

Analysis

Feasibility

The feasibility of administering the measures to respondents was assessed in terms of the level of completion of each measure for respondents who participated at baseline. Baseline measures were used as these were available for all of the data sets used. Completion is a highly simplistic measure of feasibility from the patient's perspective, but provides an indication of the acceptability of the instruments and the items included using the level of missing data as a proxy. Completion rates of ≥ 95% are high.

The analysis was carried out for both the common mental health conditions group, where the measures used were completed using self-report either at home or with an interviewer, and also for the schizophrenia and personality disorders samples, where the measures were interviewer-administered but may still lead to missing data due to item unacceptability. It is important to note that although the level of completion acts as a useful proxy for the feasibility of the measure, the results need to be interpreted with caution.

TABLE 9 Descriptive statistics and completion rates

	Time point	*n* (completing measure)	Mean	SD	Completion (%)
Common mental health conditions – EQ-5D					
AHEAD					
EQ-5D	Baseline	320	0.604	0.264	97.86
	6 months	174	0.752	0.257	
	12 months	164	0.777	0.249	
HADS-A	Baseline	324	13.11	3.48	99.08
	6 months	202	8.78	3.54	
	12 months	169	8.30	3.59	
HADS-D	Baseline	324	10.50	3.87	99.08
	6 months	202	4.94	3.92	
	12 months	169	4.34	3.99	
Common mental health conditions – SF-6D					
IAPT					
SF-6D	Baseline	504	0.613	0.13	95.60
	4 months	425	0.645	0.14	
	8 months	390	0.668	0.15	
CORE-OM clinical score	Baseline	494	20.06	7.81	93.74
	4 months	409	16.58	8.53	
	8 months	403	15.09	8.83	
PoNDER					
SF-6D	Baseline	2600	0.669	0.09	97.82
	6 months	2614	0.822	0.14	
	12 months	1697	0.839	0.13	
CORE-OM clinical score	Baseline	2640	5.23	4.97	99.32
	6 months	2641	4.73	4.91	
	12 months	1713	4.52	4.85	
Common mental health and personality disorders					
POPMACT					
EQ-5D	Baseline	476	0.503	0.320	99.20
HADS-A	Baseline	479	14.13	3.94	99.80
HADS-D	Baseline	478	11.22	4.58	99.58
Schizophrenia					
SCEPTRE					
EQ-5D	Baseline	899	0.566	0.284	99.34
	12 months	693	0.741	0.249	

TABLE 9 Descriptive statistics and completion rates (*continued*)

	Time point	*n* (completing measure)	Mean	SD	Completion (%)
MATISSE					
EQ-5D	Baseline	409	0.676	0.271	98.08
	12 months	357	0.678	0.297	
PANSS total score	Baseline	411	79.45	24.19	98.56
	12 months	334	76.15	27.11	
QUATRO					
EQ-5D	Baseline	394	0.679	0.291	96.33
	12 months	367	0.710	0.286	
SF-6D	Baseline	383	0.668	0.125	93.64
	12 months	367	0.682	0.134	
BPRS-E total	Baseline	406	45.17	13.02	99.27
	12 months	371	37.71	10.54	
Schizophrenia and personality disorders SCEPTRE					
EQ-5D	Baseline	899	0.566	0.284	99.34
	12 months	693	0.741	0.249	

This is because it is not always clear how many questionnaires respondents have completed before those included in the study or the impact fatigue may have on the level of missing data, and also because of the nature of data collection in trials where artificially high response rates are achieved.

Validity

Validity assesses how well an instrument measures what it was intended to measure, and is assessed in comparison with other instruments and clinical indicators that have been validated for use in the area (in this case mental health). In this study we will assess the validity of the measures by carrying out of number of tests, including known-group validity and convergent validity.

Convergent validity

The convergence between the generic preference-based measures and the condition-specific instruments was tested using Spearman's correlation coefficients and locally weighted scatterplot smoothing techniques.[173] Good correlations indicate that the preference-based measures can measure mental health-related factors that are assessed by the validated condition-specific instruments. Correlations are assessed as: $\geq 0.7 =$ strong, < 0.7 to $\geq 0.3 =$ moderate and $< 0.3 =$ weak. Locally weighted scatterplot smoothing (LOWESS) is a form of non-parametric regression that attempts to capture general patterns in the relationship between two measures without making assumptions about the actual relationship between the variables, and demonstrates the relationship between the measures across the scoring range. LOWESS plots a line on a scatterplot on the central tendency between the two variables. This allows the visualisation of the relationship between the variables.

In the common mental health conditions group, the convergent validity of the EQ-5D was assessed in comparison with the HADS-T, HADS-A and HADS-D, and the SF-6D was assessed in comparison with the CORE-OM clinical and dimension scores. In the schizophrenia analysis, the EQ-5D was assessed in comparison with the PANSS and the BPRS-E, and the SF-6D was assessed in comparison with the BPRS-E.

Known-group validity

For the common mental health conditions sample, the known-group validity of the EQ-5D was assessed using HADS-A and HADS-D cut-off points indicating probable anxiety or depression (a score of ≥ 11), and this analysis was carried out across both available data sets. For the SF-6D, known-group validity was assessed using CORE-OM clinical cut-off points (where a score of > 10 indicates clinical concerns). For the schizophrenia samples, known-group validity of the EQ-5D and SF-6D was assessed using BPRS-E cut-offs (31 for 'mildly ill', 41 for 'moderately ill', 53 for 'markedly ill' and 70 for 'extremely ill'), developed by Leucht *et al.*,[174] and PANSS[175] cut-offs (58 for 'mildly ill', 75 for 'moderately ill', 95 for 'markedly ill' and 116 for 'severely ill'). In the schizophrenia and personality disorder sample, the known-group validity of the EQ-5D was tested using diagnostic categories based on the Structured Interview for DSM-IV (defined as those with and without a personality disorder diagnosis, and also the number of personality disorders diagnosed).

One-way ANOVA was used to assess the magnitude of differences in the generic preference-based measure scores across the severity groups. Standardised effect sizes across severity subgroups were also calculated as the difference in mean scores between two adjacent subgroups of study participants with different levels of severity divided by the SD of scores for the milder of the two subgroups. Values between 0.2 and 0.5, 0.5 and 0.8, and ≥ 0.8 denote ranges containing small, medium and large effect sizes, respectively, according to Cohen's criteria.[65] Values < 0.2 are usually regarded as not clinically significant. However, care must be taken when comparing these between preference-based measures, because more is not necessarily better in terms of effect sizes.[54] These simply indicate whether or not the generic preference-based measure reflects what appears to be an important difference.

Responsiveness

The responsiveness analysis reported here assessed the sensitivity to change of the EQ-5D and SF-6D in comparison with the condition-specific measures. This included an analysis of both the floor and ceiling effects and the magnitude of the change in score between two study time points. If a large proportion of the sample is at the floor (lowest possible score) or ceiling (highest possible score) this impacts on the ability of the measure to detect deterioration or improvements in health, respectively. Change is assessed before and after an intervention. We accept that this is a crude indicator of change; however, for each study there was evidence of improvement, on average, for patients between baseline and follow-up. Where there has been an overall improvement, this should be reflected in a clinically significant change in the generic preference-based measures. The magnitude of the change reflected in the measures between the time points is assessed using the SRM statistic, which is calculated by dividing the mean change on the measure by the SD of the change,[176] and the effect size, which is calculated by dividing the mean change in score by the SD at baseline. Effect sizes of 0.2 are defined as small, 0.5 as moderate, and 0.8 as large.[65]

Results

Feasibility

Common mental health conditions

The AHEAD[150] data set demonstrates that both the EQ-5D and HADS had completion rates in the high range (97.86–99.08%) at baseline for those participants who returned a questionnaire. Across the IAPT[152] and PoNDER[151] data sets, the completion rates for the SF-6D and CORE-OM for those who returned a questionnaire at baseline were between 93.74% and 97.86%.

Common mental health and personality disorders

The EQ-5D and HADS were fully completed by more than 99% of the POPMACT[153] sample who were interviewed.

Schizophrenia

Across the two data sets,[154,155] the EQ-5D completion rates (for those taking part at baseline where the measures were collected via interview) range between 96.33% and 98.08%. This is slightly higher than the SF-6D which has a completion rate of 93.64%. The condition-specific PANSS and BPRS-E also have high completion rates (98.56% and 99.27%, respectively; see *Table 9*).

Schizophrenia and personality disorders

The EQ-5D had high completion rates in the SCEPTRE[120] sample (99.34%) but the low levels of missing data may reflect those who were excluded for this reason.

Across the samples, the higher levels of completion among those who returned a questionnaire may act as some form of proxy for the feasibility of using the questionnaires in mental health populations. However, there are a number of caveats to this that are described in detail in the discussion.

Convergent validity

Common mental health conditions

The correlation between the EQ-5D and HADS-T, HADS-A and HADS-D indicated a moderate level of convergence (*Table 10*). The SF-6D is correlated with the CORE-OM clinical score and the functioning, well-being and symptoms domain scores in the moderate-to-strong range across both samples (see *Table 10*). Correlation with the risk domain score is moderate for the IAPT sample and low for the PoNDER sample. All correlations are significant ($p < 0.01$).

Figure 2 displays scatterplots of the relationships between the generic and condition-specific measures and the LOWESS fit lines. The lines demonstrate that the relationship between the EQ-5D and HADS differs across the severity scale (the concordance between the measures is better at the less severe end of the scale). The relationship between the SF-6D and CORE-OM is more consistent across the severity scale, and is similar for both the IAPT and PoNDER samples.

TABLE 10 Convergent validity of the EQ-5D and SF-6D in common mental health conditions and common mental health conditions and personality disorder samples

	Common mental health conditions			Mixed
	AHEAD	**IAPT**	**PoNDER**	**POPMACT**
Condition-specific measure	**EQ-5D index**	**SF-6D index**	**SF-6D index**	**EQ-5D index**
HADS-T	−0.38*	–	–	−0.49*
HADS-A	−0.25*	–	–	−0.39*
HADS-D	−0.35*	–	–	−0.45*
CORE-OM				
Clinical score	–	−0.62*	−0.56*	–
Functioning score	–	−0.51*	−0.49*	–
Symptoms score	–	−0.65*	−0.59*	–
Well-being score	–	−0.56*	−0.46*	–
Risk score	–	−0.35*	−0.18*	–

* Significant at 0.01.
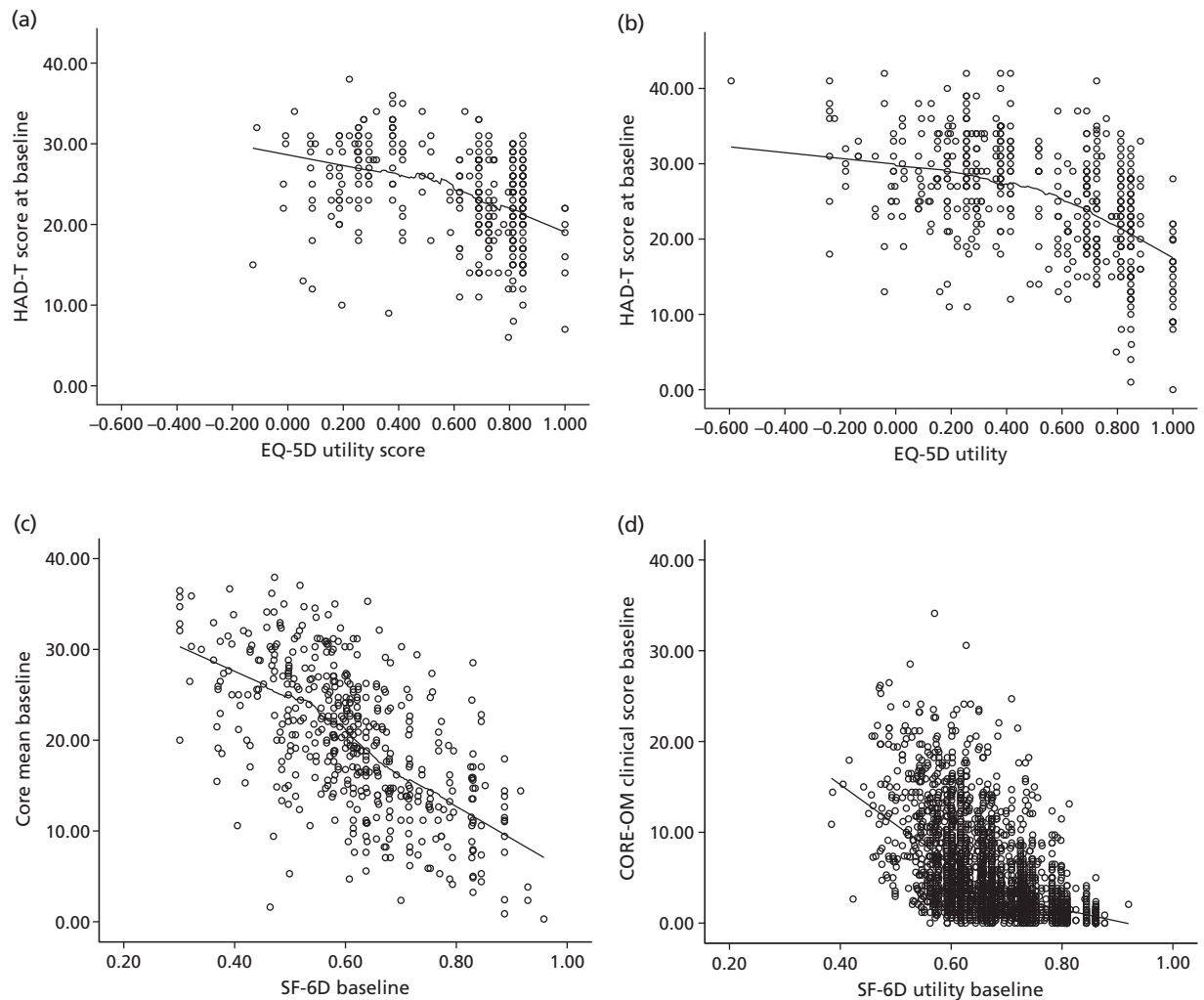
FIGURE 2 Scatterplots and LOWESS lines for the common mental health conditions and mixed diagnosis data sets. (a) AHEAD; (b) POPMACT; (c) IAPT; and (d) PoNDER.

Common mental health conditions and personality disorders

The correlation between the EQ-5D and HADS-T, HADS-A and HADS-D indicate a moderate level of convergence (see *Table 10*). Again, the LOWESS fit line for the POPMACT data indicates that the relationship between the EQ-5D and HADS differs across the severity scale, where the concordance between the measures is higher at the less severe end of the scale.

Schizophrenia

The correlations between the EQ-5D and condition-specific measures vary across the two schizophrenia samples. Correlations with the BPRS-E in the QUATRO sample[154] are moderate for the total score and the depression and positive symptom dimensions, whereas they are weak for the other dimensions (*Table 11*). Correlations with the PANSS in the MATISSE sample[155] are weak, indicating little convergence.

The correlations between the SF-6D and BPRS-E follow a similar pattern to those of the EQ-5D although the correlations are smaller in magnitude, with weak correlations across most of the dimensions apart from depression (see *Table 11*). There is, therefore, poor evidence of convergence for the SF-6D.

The LOWESS lines for the QUATRO sample (which completed both the EQ-5D and SF-6D) demonstrate a tendency for the generic preference-based measure scores to increase as scores on the BPRS decrease (equivalent to less severe problems on both measures). However, the EQ-5D displays a large ceiling effect,

TABLE 11 Convergent validity of the EQ-5D and SF-6D in schizophrenia and personality disorders

Condition-specific measure	QUATRO		MATISSE
	EQ-5D index	SF-6D index	EQ-5D index
BPRS-E total	−0.34*	−0.31*	–
BPRS-E disorganization	−0.22*	−0.17*	–
BPRS-E depression	−0.39*	−0.38*	–
BPRS-E negative symptoms	−0.14*	−0.13*	–
BPRS-E positive symptoms	−0.25*	−0.22*	–
PANSS total	–	–	−0.15*
PANSS positive	–	–	−0.14*
PANSS negative	–	–	−0.03
PANSS general symptoms	–	–	−0.21*

* Significant at 0.01.

meaning that a score of 1 on the EQ-5D is associated with a wide range of BPRS scores. There is a trend towards a linear relationship between the EQ-5D and PANSS, and again there is a large EQ-5D ceiling effect whereby a score of 1 equates to a wide range of PANSS scores (*Figure 3*).

Discriminant validity

Common mental health conditions

EQ-5D index scores are significantly higher in the 'no case' group (a score of 0–8) than in the 'probable case' group (a score of > 8) as measured by both the HADS-A and HADS-D for the AHEAD sample ($p = 0.002$). In both the IAPT and PoNDER samples, the SF-6D index score is significantly higher in the non-clinical population than in the clinical group, as measured by CORE-OM (both $p < 0.001$; *Table 12*).

Common mental health conditions and personality disorders

For the POPMACT sample, there is a similar pattern as for the AHEAD sample. The EQ-5D index scores are significantly higher in the 'no case' group than in the 'probable case' group for both the HADS-A ($p < 0.001$) and HADS-D ($p < 0.001$).

Schizophrenia

EQ-5D scores are significantly higher for those with a lower level of severity as measured by both the BPRS-E ($p < 0.001$) and the PANSS ($p = 0.003$) in the two schizophrenia samples (*Table 13*). The effect sizes across the severity subgroups are moderate in size for the BPRS-E and small for the PANSS, indicating that there is evidence that the EQ-5D can discriminate between severity groups to some extent.

The SF-6D scores significantly discriminate between BPRS-E severity groups ($p < 0.001$), with scores in the most severe group higher than those for the EQ-5D (see *Table 13*). Effect sizes indicate that the difference between the mild and moderate severity groups is in the range defined as small.

Schizophrenia and personality disorders

The EQ-5D scores also vary with the number of diagnoses, with lower scores for those with one or more personality disorders. However, these differences are not statistically significant ($p = 0.202$). There is a significant difference in EQ-5D scores between samples with different types of personality disorder, but this is difficult to interpret.
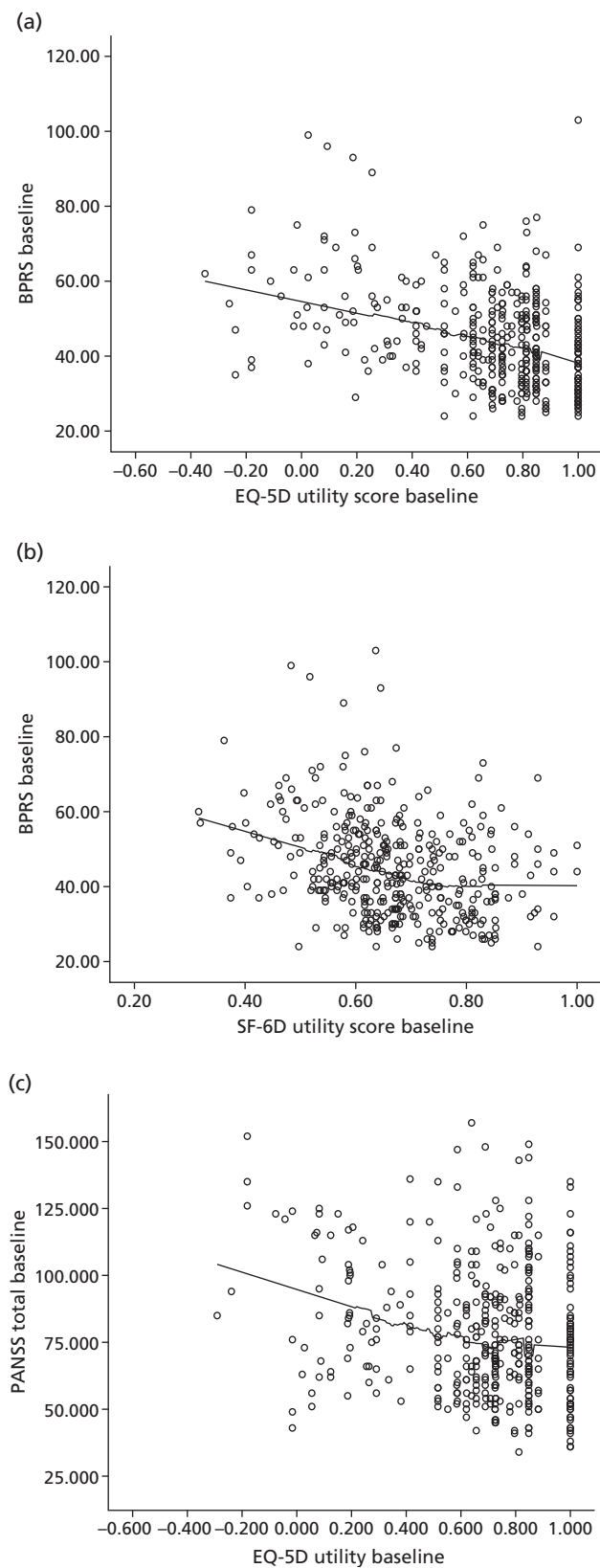
FIGURE 3 Scatterplots and LOWESS lines for the schizophrenia samples. (a) QUATRO (EQ-5D); (b) QUATRO (SF-6D); and (c) MATISSE.

TABLE 12 Discriminant validity of the EQ-5D and SF-6D in common mental health and mixed diagnosis conditions

Data and indicator	Groups	n	Mean (SD)
Common mental health conditions – EQ-5D (AHEAD)			
HADS-A caseness	No case	98	0.671 (0.25)
	Probable case	219	0.573 (0.27)
	p-value		0.002
	ES		0.37
HADS-D caseness	No case	163	0.677 (0.24)
	Probable case	154	0.525 (0.27)
	p-value		0.000
	ES		0.60
Common mental health conditions – SF-6D (IAPT)			
CORE-OM clinical	Non-clinical	53	0.740 (0.11)
	Clinical	422	0.597 (0.12)
	p-value		0.000
	ES		1.24
SF-6D sample (PoNDER)			
CORE-OM clinical	Non-clinical	2241	0.683 (0.08)
	Clinical	399	0.595 (0.07)
	p-value		0.000
	ES		1.16
Common mental health and personality disorders – EQ-5D (POPMACT)			
HADS-A caseness	No case	84	0.718 (0.28)
	Probable case	392	0.457 (0.31)
	p-value		0.000
	ES		0.88
HADS-D caseness	No case	210	0.622 (0.30)
	Probable case	265	0.410 (0.30)
	p-value		0.000
	ES		0.71

ES, effect size.

TABLE 13 Discriminant validity of the EQ-5D and SF-6D in schizophrenia and other personality disorders

Variable	Groups	EQ-5D			SF-6D		
		n	Mean (SD)	ES	*n*	Mean (SD)	ES
QUATRO							
BPRS-E	None/mild (24–31)	56	0.831 (0.16)		56	0.727 (0.09)	
	Moderate (32–41)	112	0.744 (0.23)	0.54[a]	112	0.682 (0.12)	0.47
	Marked (42–53)	119	0.652 (0.28)	0.40[a]	119	0.664 (0.12)	0.16
	Severe (> 53)	88	0.543 (0.36)	0.39	88	0.625 (0.13)	0.31
				p = 0.000			*p* = 0.000
MATISSE							
PANSS	Normal/mild (30–58)	86	0.747 (0.24)				
	Moderate (59–75)	114	0.693 (0.23)	0.23			
	Marked (76–95)	112	0.660 (0.27)	0.14			
	Severe (> 95)	92	0.606 (0.33)	0.20			
				p = 0.003			
SCEPTRE							
Number of diagnoses	No personality disorder	84	0.648 (0.24)				
	One personality disorder	248	0.606 (0.27)	0.18			
	Two personality disorders	95	0.549 (0.30)	0.21			
	Three personality disorders	42	0.493 (0.29)	0.18			
	Four or more personality disorders	41	0.416 (0.27)	0.27			
				p = 0.202			
Diagnosis[b]	None	113	0.657 (0.23)				
	Borderline	41	0.581 (0.29)	0.32[c]			
	Avoidant	69	0.638 (0.25)	0.08[c]			
	Obsessive compulsive	55	0.578 (0.27)	0.34[c]			
	Depressive	60	0.525 (0.28)	0.56[c]			
	Not otherwise specified	142	0.616 (0.29)	0.17[c]			
				p = 0.042			

ES, effect size.
a *p* < 0.05 in test of difference between adjacent severity groups.
b Sample of those with a single diagnosis.
c Effect size between adjacent severity groups.

Responsiveness

Common mental health conditions

At baseline, the EQ-5D and HADS display no evidence of floor or ceiling effects. However, at follow-up there is evidence of a large ceiling effect for the EQ-5D and a moderate ceiling effect for the HADS-D (*Table 14*). The SRM for EQ-5D is in the moderate range and for the HADS is large. This demonstrates that the HADS is more responsive in the AHEAD sample.

TABLE 14 Responsiveness of generic and condition-specific measures

Measure	Percentage at floor		Percentage at ceiling		Mean change (SD)	ES	SRM	*t*-test
	T0	T1	T0	T1				
Common mental health conditions								
EQ-5D – AHEAD (n = 164[a])								
EQ-5D	0.00	0.00	2.19	34.15	0.17 (0.38)	0.64	0.45	–
HADS-T	0.00	0.00	0.00	0.00	−10.74 (8.83)	−1.85	−1.22	–
HADS-A	0.00	0.00	0.00	0.00	−4.81 (4.98)	−0.70	−0.97	–
HADS-D	0.00	0.00	0.62	14.79	−5.93 (5.67)	−0.68	−1.05	–
SF-6D – IAPT (n = 390[a])								
SF-6D	0.00	0.00	0.00	1.54	−0.06 (0.12)	0.46	0.50	–
CORE-OM clinical score	0.00	0.00	0.00	0.00	−4.71 (6.71)	0.60	−0.70	–
Functioning score	0.41	0.00	0.82	1.50	−0.37 (0.75)	0.44	−0.49	–
Symptoms score	1.22	1.24	0.20	0.50	−0.58 (0.84)	0.62	−0.70	–
Well-being score	7.46	2.72	0.81	3.95	−0.57 (0.97)	0.63	−0.59	–
Risk score	0.2	0.00	39.27	54.48	−0.18 (0.55)	0.22	−0.32	–
PoNDER (n = 1697[a])								
SF-6D	0.00	0.00	0.00	18.33	0.17 (0.13)	1.89	1.31	–
CORE-OM clinical score	0.00	0.00	3.48	7.82	−0.58 (4.69)	−1.16	−0.12	–
Functioning score	0.00	0.06	12.35	17.24	−0.04 (0.57)	−0.06	−0.07	–
Symptoms score	0.00	0.00	8.60	16.13	−0.10 (0.57)	−0.17	−0.18	–
Well-being score	0.00	0.06	20.14	29.77	−0.10 (0.76)	−0.13	−0.13	–
Risk score	0.04	0.00	90.23	89.55	−0.01 (0.20)	−0.07	−0.05	–
Schizophrenia								
QUATRO (n = 328[a])								
EQ-5D	0.0	0.0	16.8	20.7	0.035 (0.29)	0.12	0.12	0.026
SF-6D	0.0	0.3	0.6	0.9	0.014 (0.12)	0.12	0.12	0.027
BPRS-E	1.2	4.3	0.0	0.0	−7.60 (13.06)	−0.58	−0.58	0.000
BPRS-E positive	17.1	26.8	0.0	0.0	−3.04 (5.70)	−0.52	−0.53	0.000
BPRS-E negative	21.3	35.1	0.0	0.0	−1.37 (4.06)	−0.33	−0.34	0.000
BPRS-E disorganisation	20.1	36.9	0.0	0.0	−1.62 (4.22)	−0.42	−0.38	0.000
BPRS-E depression	0.0	15.9	0.0	0.0	−1.90 (5.41)	−0.34	−0.35	0.000

TABLE 14 Responsiveness of generic and condition-specific measures (*continued*)

Measure	Percentage at floor		Percentage at ceiling		Mean change (SD)	ES	SRM	*t*-test
	T0	T1	T0	T1				
MATISSE (n = 321[a])								
EQ-5D	0.0	0.0	16.8	20.2	−0.005 (0.29)	−0.02	−0.02	0.767
PANSS	0.0	0.0	0.0	0.0	−3.41 (20.85)	−0.16	−0.14	0.004
PANSS positive	2.5	3.4	0.0	0.0	−0.93 (6.17)	−0.15	−0.15	0.007
PANSS negative	2.2	4.0	0.0	0.3	−0.78 (6.48)	−0.12	−0.11	0.031
PANSS general symptoms	0.3	0.0	0.0	0.0	−1.21 (10.65)	−0.11	−0.10	0.042
Personality disorders								
SCEPTRE (n = 679[a])								
EQ-5D	0.0	0.0	4.0	21.6	0.170 (0.29)	0.61	0.58	0.000

ES, effect size; T0, baseline; T1, follow-up.
a *n* corresponds to the number of patients who completed both measures at both time points.
Note: ES/SRM size – small = > 0.2, ≤ 0.5; moderate = > 0.5, < 0.8; large = ≥ 0.8.

The SF-6D displays some evidence of a ceiling effect for the PoNDER data (see *Table 14*). The SRM statistics for the SF-6D and CORE-OM in the IAPT validation sample are in the moderate range. For the PoNDER sample, the SF-6D SRM is in the large range, in contrast to the CORE-OM dimensions which are in the small range. Therefore, there is evidence that the responsiveness of the SF-6D is in the same range as the CORE-OM for depression, and that it is more responsive in postnatal depression.

Schizophrenia

The EQ-5D displays no evidence of floor effects at baseline but there is evidence of a large ceiling effect at both baseline and follow-up time points in the QUATRO and MATISSE samples (see *Table 14*). Mean change in the QUATRO sample is statistically significant, but the effect sizes and SRMs are < 0.2 (below the clinically significant range). The BPRS-E has larger effect size and SRM statistics, which indicates that the EQ-5D was less responsive in this sample. The SF-6D displays no evidence of floor or ceiling effects in the QUATRO sample. Mean change on the SF-6D is smaller than the EQ-5D and the effect size and SRM statistics were consistently < 0.2.

In the MATISSE sample, mean change for the EQ-5D is not statistically significant, leading to a small effect size and SRM statistics. The PANSS demonstrates statistically significant mean change; however, the effect sizes are in the low range. The small change demonstrated indicates that neither the EQ-5D nor PANSS are responsive in the MATISSE schizophrenia sample.

Schizophrenia and personality disorders

In the SCEPTRE sample, the EQ-5D displays minimal floor and ceiling effects and shows good responsiveness, with moderate effect sizes and SRMs at 12 months.

Discussion

Seven data sets were used to examine the psychometric validity of the EQ-5D and SF-6D across a range of mental health conditions in comparison to existing condition-specific measures. The findings add to the growing literature relating to the validity of the instruments across mental health populations in the light of a range of caveats to the findings (described below). The results suggest that the generic preference-based measures are psychometrically valid for use in common mental health condition and mixed diagnosis groups in comparison with widely used and validated measures of mental health (HADS and CORE-OM). Furthermore, there is some evidence that the measures respond to change in health status over time. In comparison, the evidence for the psychometric validity of the generic preference-based measures in the schizophrenia and personality disorder patient groups is less clear. There is support for construct validity across some related domains and some evidence for the discriminant ability of the measures, but responsiveness to change in health status over time is low.

The psychometric validity of the EQ-5D and SF-6D in common mental health patient samples examined in this study is consistent with previous empirical work reported in *Chapter 2*, which established the validity of the instruments for use in mild depression and anxiety samples. The measures may be valid for use in these populations because both descriptive systems include mental health-specific questions that are relevant to depression and anxiety, and they may therefore have a level of sensitivity to the conditions and some level of association with the widely used comparison measures. We have also established some evidence that the psychometric performance of the EQ-5D is acceptable in a sample with common mental health and personality disorder diagnoses who self-harm, and this group could be seen as a moderately severe sample. There are some differences between the EQ-5D and SF-6D, but direct comparisons are difficult because the analysis of each measure was carried out using different samples with different characteristics. However, the evidence for both measures can be seen as complementary in establishing an overall picture of the psychometric performance of the instruments. The growing evidence base regarding the validity of the instruments means that the EQ-5D and SF-6D can be considered for use in the economic evaluation of interventions for common mental health disorders with some level of confidence.

The review in *Chapter 2* found limited and mixed evidence for the performance of generic preference-based measures in schizophrenia. We have also found evidence for and against the validity of these measures in schizophrenia and personality disorder samples with mixed evidence on the ability of the measures to reflect schizophrenia-specific symptoms. There is evidence that the EQ-5D may be related to some condition-specific domains (for example, depression) but not to others (such as positive symptoms), and again this may be linked to the classification system which directly assesses anxiety and depression but is not sensitive to other schizophrenia-specific domains. The results indicate that the EQ-5D has the ability to discriminate across severity groups in schizophrenia, and also across different personality disorder diagnoses. However, the large ceiling effects at baseline of the EQ-5D in the schizophrenia data sets may impair its ability to detect change over time. Direct comparisons between the EQ-5D and SF-6D were only possible for one of the data sets included in this study. The QUATRO study[154] demonstrates that neither instrument converges with the condition-specific measure (but this may not be expected), and neither instrument responds to change over time at the same level that is reflected in the condition-specific measures. The mixed evidence regarding the schizophrenia and personality disorder samples means that the EQ-5D and SF-6D should be used with caution in these groups, and further research in other samples to investigate psychometric performance in more detail is warranted.

Psychometric analysis of the preference-based measures is one method of assessing validity, and should be considered alongside other types of evidence to establish a detailed picture of the performance of the measures. For example, this work should be considered alongside systematic review evidence presented in *Chapter 2*, and also qualitative work assessing the content validity and acceptability of the instruments with mental health service users, presented in later chapters in this report. This allows for detailed insight into the performance of the instruments and will inform future work to increase the sensitivity and validity of measurement across a range of mental health conditions. A five-level version of the EQ-5D has been

developed,[177] and it is possible that this version may be more sensitive to different severity levels and change across time. Further research could assess the validity of the five-level version in people with mental health conditions.

This study has a number of limitations. We used the level of missing data as a form of proxy for the feasibility of the measure in mental health populations, an approach which can be criticised as the external pressures and expectations felt by respondents in trials to complete the measures is unclear. There are a range of reasons why measures may be incomplete that do not specifically relate to the measure, including fatigue, illness, lack of motivation, or the position of the questionnaires in the study; those appearing later in the assessment protocol may have higher levels of missing data.

The inferences that can be drawn from the results are also limited by the data available to carry out the analysis. The differing levels of performance in terms of construct validity, convergent validity and responsiveness may reflect systematic variance attributable to different types of data used in this study, the different patient populations and different study designs. Furthermore, only one data set included both the EQ-5D and SF-6D. This means that the level of transferability to other mental health samples with similar diagnoses but different characteristics is unclear, and full comparisons between the generic preference-based measures are not possible. However, the results presented here across the different samples should be seen as complementary, and part of the process of building an overall picture of the psychometric validity of the instruments. Further work into the performance of the generic measures could repeat the analysis reported here on different data sets using different condition-specific measures and indicators. It would also be useful to directly compare the EQ-5D and SF-6D in samples where both have been completed.

In summary, we have reported the first work to pool data from a variety of sources to test the psychometric performance of two widely used generic preference-based measures of HRQoL in populations with mental health problems. The study adds to the evidence base regarding where generic preference-based measures can be used in terms of applicable mental health populations. It also highlights possible areas where new preference-based measures, or additions to existing measures, would improve the measurement of HRQoL in mental health.

Chapter 4　Mapping mental health condition-specific measures to generic preference-based measures

The previous chapter presented evidence on the validity of the generic preference-based measures, the EQ-5D and the SF-6D, for common mental health problems such as depression and anxiety. Mental health condition-specific measures (e.g. the HADS and the CORE-OM) tend to be used in clinical studies of mental health interventions rather than generic measures. Scores from these measures cannot be used for cost-per-QALY analysis as they are not preference based. One potential solution to the unavailability of the generic preference-based measure in studies is to map, or cross-walk, the condition-specific measures onto the generic measure in order to obtain the generic scores. This involves estimating the relationship between the non-preference-based condition-specific measure and a generic preference-based measure using statistical association.[178] This requires a degree of overlap between the descriptive systems of the two measures and that the two measures are administered in the same population. The psychometric analysis indicated that there was some correlation between depression and anxiety measures and generic preference-based measures, which means that mapping may offer a solution to providing EQ-5D or SF-6D scores. This chapter aims to develop and test mapping functions from the condition-specific measures for depression and anxiety that would be used to generate scores where generic preference-based measures had not been used.

Methods

Analysis focused on the baseline measures for the depression and anxiety data sets presented in the previous chapter (AHEAD, POPMACT and IAPT) for those with non-missing items, and one additional data set, Ethnic Minority Psychiatric Illness Rates in the Community (EMPIRIC),[179] described below (see *Data*). We focused on the two generic measures (target measures), the EQ-5D and SF-6D, as well as the condition-specific measures used in *Chapter 3*, the HADS and CORE-OM (source measures). We included other condition-specific measures that are commonly used in the assessment of depression and anxiety: the PHQ-9, GAD-7 and GHQ-12, described below (see *Measures*). This section provides an overview of the analysis followed by details of measures and mapping functions developed in each data set.

Analysis

Preliminary analysis was undertaken prior to mapping in order to identify which condition-specific dimensions/items to use (model specification) and which technique to use to model the relationship between these condition dimensions/items and the target generic measures (model type). Correlation analysis exploring the relationship between generic scores/dimensions and condition-specific scores/items was used to identify which dimensions and/or items would be expected to have a strong relationship with the EQ-5D or SF-6D scores/dimensions and hence the appropriate model specification.[180] Correlation analyses were also used to determine whether or not any independent variables were highly correlated and, therefore, not recommended for inclusion in the same regression model, at both the dimension and item levels. We assessed the distribution of the scores of the generic measures to help inform the choice of estimation technique.

Mapping functions can be generated using different specifications based on assumptions regarding the relationship between generic measures and condition-specific measures. The simplest function is the additive model, where our target measure, the EQ-5D or SF-6D score, is regressed on the total score of the condition-specific measure of interest, for example the HADS total score. This is a limiting specification as it assumes that the dimensions/items of the starting measure carry equal weight and that responses on all items lie on the same interval scale.[178] This assumption can be relaxed by using the dimensions or items

from the starting measure instead. Dimension scores are treated as continuous variables whereas items are modelled as dummy variables, i.e. there is a variable for each level in the item (all of the time, some of the time, etc.), which has a value of 1 for those at that level and 0 for everyone else. The assumptions of the additive model can be relaxed by adding interaction terms and squared terms for dimensions.[178] This is often limited to dimension level, especially where there are a large number of items with several levels which can result in a large number of variables. Analysis may also be limited to statistically significant items. Other variables can also be included, for example clinical and sociodemographic information. In this study, we tested additive models with total scores and dimension scores and then tested the inclusion of interaction and squared terms at the dimension level. We also analysed item-level models but did not test for interactions or squared terms for these models. Sociodemographic information (age and sex) was added to the best-fitting models for each condition-specific measure. Comparisons were made between models that had all variables and those that only included variables that were significant at the 10% level.

We started by using ordinary least squares (OLS) regression to model the relationships, as this is the most common model technique.[178] However, OLS fails to take into account the bounded nature of generic preference-based measures which have a maximum value of 1. It also fails to take into account the distribution of some measures such as the EQ-5D, where there may be a large number of respondents at a score of 1 with non-normal (bimodal or trimodal) distributions.[178] The Tobit model can be used to take into account the bounded nature of the preference-based measure values.[180] The two-part model, which models the probability of having a score of 1 and the generic preference-based measure score for those not at 1 in two separate regressions, can also be used where there is a sufficient number of respondents at 1 (full health).[180] Assessment of the range and distribution of utility scores was used to determine whether to use the Tobit or two-part model in the analysis and the models were used where appropriate.

More complex models can be developed whereby, rather than the target measure being the generic measure score, separate regressions are undertaken on the dimensions of the generic preference-based measure, referred to as response-level mapping in the literature.[181] For example, for the EQ-5D, separate regressions are carried out for the five dimensions (mobility, self-care, usual activities, pain and anxiety/depression). Here, the condition-specific measure is used to model the probability of an individual being at a particular level of the generic measure dimension, for example level 2 for mobility, self-care, etc. This results in separate regressions for each dimension of the generic preference-based measure, which are then used to classify individuals in each dimension. Existing tariffs are then used to generate the utility scores. Multinomial or ordered logistic regressions are used to model these probabilities. We tested the application of response-level mapping.

Regression coefficients were assessed based on expected relationship with generic preference-based measure scores and dimensions as well as statistical significance using the Wald test ($p \leq 0.1$). Model goodness of fit was measured using R^2, adjusted R^2, pseudo-R^2 and Bayesian information criterion (BIC). With R^2 values, the higher the value, the better the model, and with BIC, the lower the value, the better the model. The main concern when assessing mapping results is the accuracy of the predictions.[180] Model prediction properties were tested using the root-mean-square error (RMSE), a measure of the difference between the observed and predicted generic scores. Smaller RMSE values indicate better-performing models. We also compared RMSE alongside predicted means across the EQ-5D and SF-6D range as well as using severity measures where group differences were expected. As the aim of the study was to develop mapping functions that could be used in external studies, it was important to assess how well these functions would work in an external data set within a similar patient population. Follow-up samples were available for some of the data sets used. However, patient clinical characteristics may have changed at follow-up because of interventions, and so generic and condition-specific measure scores may not be comparable to baseline values. Validation was therefore undertaken for the best-fitting models by splitting the baseline samples (75%/25%). The larger sample (75%) was used to estimate the model and the smaller sample was used to validate this model.

We undertook standard tests for the modelling techniques that we used. For OLS, we assessed whether or not there was evidence of heteroscedasticity using White's test. Multicollinearity was assessed using the collin user-generated command in Stata 11 (StataCorp LP, College Station, TX, USA) and variables exhibiting collinearity were excluded. Non-linearity was tested using fractional polynomials and squared terms were tested where there was evidence of non-linearity. The *omodel* user-generated Stata command was used to test for the proportional odds assumption when undertaking response-level mapping. Multinomial regressions were used where this assumption was violated in any one of the dimension regressions, and ordered logit regressions were used where this assumption held for all the regressions. Logit regressions were used where there were dependent variables that had been collapsed to two levels because of small proportions of individuals at the lowest levels; for example, very few individuals reporting level 3 in the mobility dimension of the EQ-5D. Model specification was assessed using the link test. We report the performance statistics for the best-fitting models for each condition-specific measure in the main report, and performance statistics for other models are reported in *Appendix 2*. Stata 11 was used in all the analyses.

Measures

Generic preference-based measures
We used the two generic preference-based measures – the EQ-5D and the SF-6D described in *Chapter 1*. The EQ-5D scores range from –0.594 to 1 and it has five dimensions (mobility, self-care, usual activities, pain/discomfort and anxiety/depression), each with three levels. The SF-6D scores range from 0.30 to 1 and it has six dimensions (physical functioning, role limitation, social functioning, pain, mental health and vitality) with four to six levels. We mapped to both the scores and dimensions of the two measures where they were available.

Condition-specific measures
We focused on the common condition-specific measures for depression and anxiety, some of which have already been introduced in the previous chapter, as follows.

Hospital Anxiety and Depression Scale
The HADS[156] has two dimensions: the HADS-D (seven items) and the HADS-A (seven items), with scores from 0 to 21 as well as the total score. The dimension scores are used separately to identify depression and anxiety, with high scores indicative of increased levels of symptoms, and these were used as independent variables to predict EQ-5D and SF-6D scores (NB HADS items were not available in the data sets).

Patient Health Questionnaire
The PHQ-9 is the nine-item depression module of the Patient Health Questionnaire.[182] It is a self-report checklist which covers aspects related to depression. Items are completed on a four-point scale from 0 to 3 ('not at all', 'several days', 'more than half the days', and 'nearly every day') with higher scores indicating depression. An additional item asks respondents to rate the impact of depression on their daily activities. The PHQ-9 has been shown to be valid for use with patients and the general public to assess depression and distinguish between levels of severity.[183,184] The total score and the items (including the additional item) were used to predict SF-6D scores.

Generalised Anxiety Disorder Assessment
The GAD-7 items describe the most prominent diagnostic features of the DSM-IV diagnostic criteria for generalised anxiety disorder.[185] On the GAD-7, subjects are asked how often, during the last 2 weeks, they have been bothered by each of the seven core symptoms of generalised anxiety disorder as well as an additional question on the impact of these symptoms on their activities. Response options are the same as the PHQ-9 with higher scores indicating anxiety. The GAD-7 was shown to be a valid measure for screening for generalised anxiety disorder and assessing severity levels.[185,186] Total scores and all items were used to predict SF-6D scores.

Clinical Outcomes in Routine Evaluation – Outcome Measure

The CORE-OM[165] dimensions of subjective well-being, symptoms (anxiety, depression, physical problems, trauma), functioning (general functioning, close relationships, social relationships) and risk (risk to self, risk to others), as well as the 34 items were used to predict SF-6D scores. Items are scored on a five-point, 0–4 scale ('not at all', 'only occasionally', 'sometimes', 'often', 'all or most of the time') and means are computed for dimension and total scores, with higher scores indicating problems in those dimensions.

General Health Questionnaire

The GHQ-12 is a self-report measure of psychological morbidity, intended to detect 'psychiatric disorders . . . in community settings and non-psychiatric settings'.[187] It is widely used in clinical practice, epidemiological research and psychological research. It comprises six items that are positive descriptions of mood states (e.g. 'felt able to overcome difficulties') and six that are negative descriptions of mood states (e.g. 'felt like a worthless person'). Total scores range from 0 to 12 and item levels range from 1 to 4, with higher scores indicating psychological morbidity. The GHQ-12 has been validated for use in mental health populations.[188] The total GHQ score and the items were used to predict SF-6D scores.

The condition-specific total and dimension scores increase with increasing severity and were therefore expected to be negatively associated with generic measure scores and positively associated with generic measure dimensions.

Data

We focused on two common mental health conditions data sets (AHEAD and IAPT) and one common mental health conditions and personality disorders data set (POPMACT), described in detail in *Chapter 3*. The PoNDER sample was not used as it focused on women with postnatal depression and mapping functions developed from this sample may not be transferable to other, more general samples. We also included an additional sample (EMPIRIC), described below.

Assessing Health Economics of Antidepressants

The AHEAD study,[150] which focused on patients who reported a range of common mental health concerns including depression, MADD and a number of phobias ($n = 286$) was used to map from the HADS to the EQ-5D.

POPMACT

A second data set that contains the EQ-5D and HADS is the POPMACT study,[189] which recruited patients presenting with recurrent deliberate self-harm ($n = 475$). There are differences in AHEAD and POPMACT populations in terms of HADS and EQ-5D scores and separate algorithms were therefore developed for the two samples. HADS depression scale cut-offs identifying probable cases of depression (> 10) were used as the measure of severity in assessing the predictive ability of the mapping functions in the AHEAD and POPMACT samples.

Improving Access to Psychological Therapies

The IAPT cohort study,[152] which focused on patients with mild/moderate depression and/or anxiety ($n = 394$), was used to develop mapping algorithms between the PHQ-9, GAD-7, CORE-OM and SF-6D. The PHQ-9 depression scale cut-offs identifying severity levels of depression were used to assess the predictive performance for the mapping functions developed from these measures.

Ethnic Minority Psychiatric Illness Rates in the Community

The EMPIRIC study,[179] which was used to estimate the prevalence of psychotic symptoms and risk factors for reporting psychotic symptoms in a community-based sample of people from different ethnic groups, was used to map between the GHQ-12 and the SF-6D. The study was a cross-sectional survey of 4281 adults aged 16–74 years living in private households in England. The analysis focuses on those who had any depressive or anxiety episodes ($n = 213$). The Revised Clinical Interview Schedule (CIS-R),[190]

which is used for identifying the presence of common mental health disorders, was used to assess the performance of mapping functions.

Results

Preliminary analysis results

Range and distribution of scores

Descriptive statistics are reported in *Table 15*. Across all four samples, most of the generic scores did not cover the full range of the measures, apart from in the POPMACT sample. None of the SF-6D scores were at 1, and there were only a few individuals (7/286) at 1 in the EQ-5D in the AHEAD sample. Most of the condition-specific measures cover the full range of scores.

The EQ-5D scores showed evidence of bimodal distributions whereas SF-6D scores were approximately normally distributed across the samples (*Figure 4*). Distribution of EQ-5D dimension levels was skewed, with < 5% of individuals in level 3 in all dimensions apart from the anxiety and depression dimension (25%) in the AHEAD sample. In the POPMACT sample, the EQ-5D dimensions had < 1% in level 3 for mobility and self-care. SF-6D dimension levels were evenly distributed in both IAPT and EMPIRIC samples.

Correlation analysis

Preliminary analysis of correlations between the generic preference-based measure and the condition-specific measures as reported in *Chapter 3* indicated that there was a moderate level of convergence between the EQ-5D and the HADS, with negative correlations ranging from 0.25 to 0.49 across the AHEAD and POPMACT samples. The correlations were smaller in the AHEAD sample, which may indicate that the HADS will be a poor predictor of EQ-5D scores in this sample. Further correlation analysis (see *Appendix 3*, *Table 36*) indicated that the HADS dimension scores and total scores were poorly correlated with most dimensions of the EQ-5D, with positive correlation scores ranging from 0.01 to 0.21 in mobility, self-care and pain dimensions across the two samples. Correlations with the usual activities dimensions ranged from 0.17 to 0.40, with high correlations in the POPMACT sample. The anxiety/depression dimension had moderate correlations of 0.39 to 0.50 (see *Appendix 3*, *Table 37*). HADS dimensions were therefore expected to be poor predictors of the EQ-5D dimensions. In contrast, the SF-6D scores had better correlations, ranging from 0.33 to 0.56, indicating that HADS dimensions could be better predictors of SF-6D scores. SF-6D dimensions were not available so correlations were not tested.

Total scores of the PHQ-9, GAD-7 and CORE-OM total and dimension scores were moderately correlated with SF-6D score (0.52 to 0.65), apart from the CORE-OM risk dimension which had a smaller correlation (0.35) (see *Appendix 3, Table 38*). These measures were positively correlated with SF-6D dimensions with the largest correlations in the SF-6D mental health dimension, whereas physical functioning and pain had smaller correlations (< 0.4). Correlations between the SF-6D score and the PHQ-9 items ranged from 0.34 to 0.59; for the GAD-7 they ranged from 0.36 to 0.51. SF-6D dimension correlations with PHQ-9 items were moderate to high in most dimensions apart from the pain dimension; the same was true for the GAD-7 items, but correlations with the physical functioning dimension were also low. CORE-OM items in the IAPT sample had moderate negative correlations with the SF-6D score. Item-level correlations mainly ranged between 0.3 and 0.5. Correlations between SF-6D dimensions and CORE-OM items were also positive and moderate for most of the items. This indicated that the PHQ-9, GAD-7 and CORE-OM could be good predictors of SF-6D scores and dimensions apart from pain and physical functioning. Correlation results for these three measures are reported in the appendices (see *Appendix 3*, *Tables 39–42*).

The GHQ-12 in the EMPIRIC sample had low correlations with the SF-6D at the total score and item level (< 0.3) (see *Appendix 3*, *Table 43*). Correlations with SF-6D dimensions were also small. This indicated that the GHQ-12 would be a poor predictor of SF-6D scores and dimensions.

TABLE 15 EQ-5D, SF-6D and condition-specific measure summaries and descriptive statistics

Variables	Mean	SD	Minimum	Maximum
AHEAD (n = 286)				
EQ-5D	0.62	0.26	−0.13	1
SF-6D	0.62	0.09	0.30	0.80
HADS-D	10.48	3.95	0	20
HADS-A	13.06	3.47	2	20
HADS-T	23.54	5.89	6	38
Age	42.5	15.32	17	91
Female n (%)	192 (67.1)			
POPMACT (n = 475)				
EQ-5D	0.50	0.32	−0.59	1
SF-6D	–	–	–	–
HADS-D	11.22	4.55	0	21
HADS-A	14.15	3.95	0	21
HADS-T	25.37	7.45	0	42
Age	32.0	11.23	16	66
Female n (%)	325 (68.4)			
IAPT (n = 394)				
SF-6D	0.62	0.13	0.30	0.93
CORE-OM T	19.84	7.77	0.88	37.94
CORE-OM W	2.57	0.93	0.00	4.00
CORE-OM S	2.37	0.92	0.08	4.00
CORE-OM F	2.10	0.85	0.00	4.00
CORE-OM R	0.60	0.79	0.00	4.00
PHQ-9	15.59	7.05	0.00	27.00
GAD-7	13.42	5.59	0.00	21.00
Age	40.0	13.9	16	89
Female n (%)	283 (71.8)			
EMPIRIC (n = 213)				
SF-6D	0.59	0.11	0.35	0.88
GHQ-12	16.97	7.33	2	36
Age	39.0	13.97	16	73
Female n (%)	127 (59.6)			

CORE-OM F, CORE-OM functioning; CORE-OM R, CORE-OM risk; CORE-OM S, CORE-OM symptoms; CORE-OM T, CORE-OM total; CORE-OM W, CORE-OM well-being.
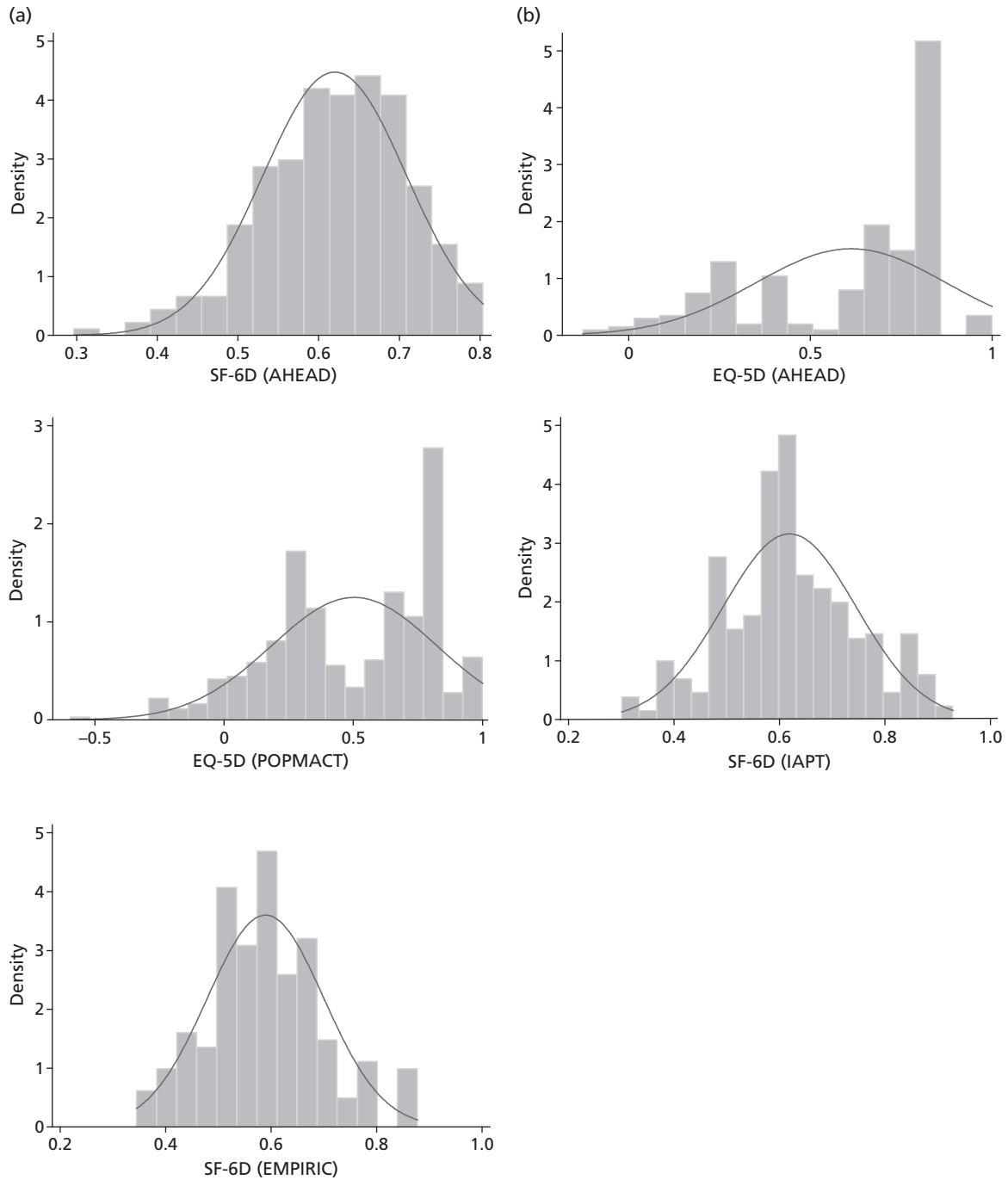
FIGURE 4 SF-6D and EQ-5D histograms. (a) SF-6D; and (b) EQ-5D.

This initial analysis indicated that OLS regression analysis was likely to be the best approach for the SF-6D scores; as there were no individuals with a score of 1, the Tobit model would not improve results and two-part models could not be undertaken. The EQ-5D scores had a small percentage at full health in the AHEAD and POPMACT samples, so Tobit models may result in some differences, and these were tested. However, two-part models cannot be run where there are very few individuals in full health, as was the case in these samples. Multinomial regressions for predicting EQ-5D and SF-6D dimensions could be undertaken and, again, small numbers at lower levels for EQ-5D dimensions and poor correlations meant that these models were not expected to generate large improvements over the OLS model. *Table 16* provides an overview of the data sets and measures and the analysis that was undertaken. This includes the dependent variables based on the available data, the independent variables which are the condition-specific measures, the model type and the model selection information, based on the preliminary analysis results.

TABLE 16 Overview of analysis

Dependent variables	Independent variables	Model selection and specification	Model type	Performance
AHEAD (n = 286)				
EQ-5D index EQ-5D dimensions SF-6D index	HADS-T HADS-A total score HADS-D total score	EQ-5D scores range from –0.126 to 1 (seven individuals at 1). Bimodal distribution EQ-5D dimensions < 5% in level 3 apart from anxiety and depression dimension (25%) SF-6D score range 0.2968 to 0.8037. Normal distribution	OLS Tobit (EQ-5D) Response-level mapping (EQ-5D)	Measure of severity: HADS depression scale cut-offs identifying probable cases of depression
POPMACT (n = 475)				
EQ-5D index EQ-5D dimensions	HADS total score HADS-A total score HADS-D total score	EQ-5D scores range from –0.594 to 1. Bimodal distribution EQ-5D dimensions < 1% in level 3 for mobility and self-care	OLS Tobit Response-level mapping	Measure of severity: HADS depression scale cut-offs identifying probable cases of depression
IAPT (n = 394)				
SF-6D index SF-6D dimensions	PHQ-9 total PHQ-9 items + 1 item daily activities GAD-7 total GAD-7 items + 1 item daily activities CORE-OM mean CORE-OM dimensions CORE-OM items	SF-6D score range from 0.301 to 0.929. Normally distributed SF-6D dimensions mainly evenly distributed across levels	OLS Response-level mapping	Measure of severity: PHQ-9 depression scale cut-offs identifying severity levels of depression
EMPIRIC (n = 213)				
SF-6D index (SF-12 vs.1 UK) SF-6D dimensions	GHQ-12 score GHQ-12 items	SF-6D score range from 0.345 to 0.878. Normally distributed SF-6D dimensions mainly evenly distributed across levels	OLS Response-level mapping	Measure of severity: CIS identifying presence of common mental disorders

CIS, Clinical Interview Schedule.

Mapping results

In this section we report performance statistics of the best-fitting models for each source measure (HADS, PHQ-9, GAD-7, CORE-OM and GHQ-12) and the two target measures (EQ-5D and SF-6D). Performance statistics of the best-fitting models from the other models that were tested, and OLS regression results, are reported in *Appendix 3*.

Hospital Anxiety and Depression Scale and EQ-5D

Ordinary least squares regressions were undertaken for the EQ-5D regressions with HADS dimensions and total scores in the AHEAD and POPMACT samples. The best-fitting models were the total score model in the AHEAD sample and the model with both anxiety and depression dimensions in the POPMACT sample. Coefficients were negative as expected, indicating that as depression and anxiety increase, EQ-5D scores fall. *Figure 5* shows the predicted and observed scores for these two models; these have a weak association, particularly on the ends of the scale. Interaction terms were not statistically significant and fractional polynomials indicated that HADS dimensions and totals have a linear relationship with the EQ-5D.
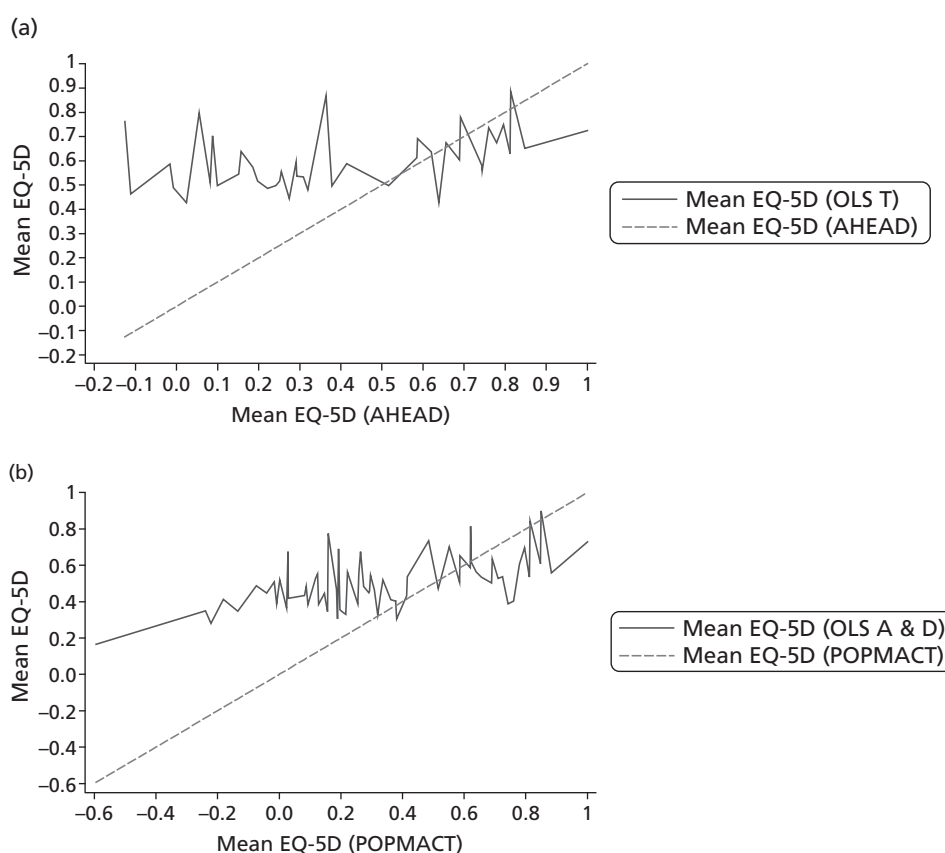
FIGURE 5 Predicted and observed EQ-5D – HADS. (a) HADS total score (AHEAD); and (b) HADS-A and HADS-D (POPMACT).

Summary goodness of fit and predictive measures are reported in *Tables 17* (AHEAD) and *18* (POPMACT). As OLS is used, the predicted mean value for the best-fitting models is the same as the observed values, but the minimum and maximum scores indicate a smaller range of values. This reflects findings shown in *Figure 5* that these models overpredicted at the bottom end of the EQ-5D and underpredicted at the top end. As expected, there was poor association between these predictors and the EQ-5D with small R^2 values, particularly in the AHEAD sample. In the AHEAD sample, age was statistically significant and improved predictive performance (*Table 17*, model 2) based on BIC. However, the link test indicated that this model was misspecified and the White test showed evidence of heteroscedasticity. Predictive performance was poor, with a RMSE of 0.375 for EQ-5D scores < 0.3 and 0.193 for scores > 0.8 in the AHEAD sample. Although not directly comparable, RMSE scores in the POPMACT sample exhibited a similar pattern, with large RMSEs at the severe end (RMSE = 0.299 for EQ-5D values < 0). Predictions were slightly better at the top end of observed EQ-5D scores for the POMACT sample (RMSE = 0.082 to 0.088). Age and sex were not statistically significant in this sample. There was evidence that predicted EQ-5D scores could discriminate between individuals with and without depression in both samples but the RMSEs were still large.

TABLE 17 Summary performance statistics: OLS regressions, EQ-5D onto HADS (AHEAD)

		Predicted EQ-5D			
		Model 1		Model 2	
Independent variables	Observed EQ-5D	HADS-T		HADS-T + age	
Summary statistics					
Mean	0.61	0.61		0.61	
SD	0.262	0.104		0.115	
Minimum	−0.126	0.3571		0.3332	
Maximum	1	0.9208		0.9173	
Regression performance statistics					
R^2		0.156		0.193	
Adjusted R^2		0.153		0.187	
RMSE		0.194		0.188	
BIC		8		1	
EQ-5D range	***n***	**Mean EQ-5D**	**Mean EQ-5D**	**RMSE**	**Mean EQ-5D**	**RMSE**
< 0.3	59	0.1816	0.5565	0.375	0.5495	0.368
0.3 to 0.599	34	0.4186	0.5421	0.129	0.5254	0.124
0.6 to 0.699	52	0.6709	0.6142	0.098	0.6002	0.099
0.7 to 0.799	34	0.7402	0.655	0.098	0.655	0.105
0.8 to 1	107	0.8408	0.6495	0.193	0.6655	0.179
HADS-D	***n***	**Mean EQ-5D**	**Mean EQ-5D**	**RMSE**	**Mean EQ-5D**	**RMSE**
Non-case/possible ≤ 10	147	0.6931	0.6821	0.169	0.6816	0.160
Probable case > 10	139	0.5258	0.5374	0.221	0.5379	0.218

Regression results are reported in *Appendix 3*, *Table 48*.

TABLE 18 Summary performance statistics: OLS regressions (POPMACT – EQ-5D)

| Independent variables | Observed EQ-5D | Predicted EQ-5D |
| | | Model 3 |
		HADS-D and HADS-A
Summary statistics		
Mean	0.5039	0.5039
SD	0.3195	0.1573
Minimum	–0.594	0.1492
Maximum	1	1.016
Regression performance statistics		
R^2		0.242
Adjusted R^2		0.239
RMSE		0.227
BIC		150

EQ-5D range	*n*	Mean EQ-5D	Mean EQ-5D	RMSE
≤ 0	33	–0.1245	0.4021	0.299
0.001 to 0.199	60	0.1295	0.4423	0.117
0.2 to 0.299	64	0.263	0.4324	0.045
0.3 to 0.599	90	0.4242	0.4449	0.017
0.6 to 0.699	57	0.6699	0.532	0.037
0.7 to 0.799	45	0.7374	0.5421	0.062
0.8 to 0.899	103	0.8382	0.586	0.082
0.9 to 1	23	1	0.7295	0.088

HADS-D	*n*	Mean EQ-5D	Mean EQ-5D	RMSE
Non-case/possible ≤ 10	210	0.622	0.6375	0.220
Probable case > 10	265	0.4104	0.3981	0.233

Regression results are reported in *Appendix 3, Table 49*.

Tobit regressions were also undertaken. Summary statistics of the best-fitting model indicated that OLS performed better than Tobit in terms of BIC for both samples, with little difference in terms of overall RMSE (see *Appendix 3*, *Tables 44* and *45*). However, RMSE over the EQ-5D range showed that the Tobit model performed slightly better at the severe end of the EQ-5D scale and worse at the top end. Response-level mapping models were also tested for the two best-fitting OLS models: those with anxiety and depression combined and with the total scores. As already noted, there were very few individuals at level 3 in the five EQ-5D dimensions and dimension levels had to be collapsed in order to run these models. This resulted in loss of information and poor predictive performance, especially at the severe end of the EQ-5D scale, with RMSE > 0.5. The best-fitting response-mapping model performance statistics are also reported in *Appendix 3*, *Tables 44* and *45*.

Hospital Anxiety and Depression Scale and SF-6D

Ordinary least squares regressions to predict the SF-6D from the HADS in the AHEAD sample indicated that the best-fitting model based on BIC was the one with both the HADS-D and HADS-A dimension scores. This model showed evidence of heteroscedasticity but no evidence of misspecification based on the link test. The coefficients were negative, as expected. Plots of the predicted and observed SF-6D scores indicated that the HADS was poor at predicting SF-6D scores, particularly at the severe end of the scale (*Figure 6*). Minimum predicted SF-6D scores for Models 4 and 5 were much higher than the observed SF-6D scores (*Table 19*), although RMSE for scores < 0.5 indicate that this might be an outlier as these range between 0.141 and 0.134. There was also evidence of underpredicting at the top end of the SF-6D scale. The R^2 values were higher than those reported for the EQ-5D regressions, indicating better association. As with predicted EQ-5D scores, predicted SF-6D scores are able to discriminate between those with and without depression. Interaction and squared terms do not add to the model, although age was statistically significant and added to the model (*Table 19*). Tobit models did not provide improvements over OLS and SF-6D dimensions were not available to test response-level mapping models.
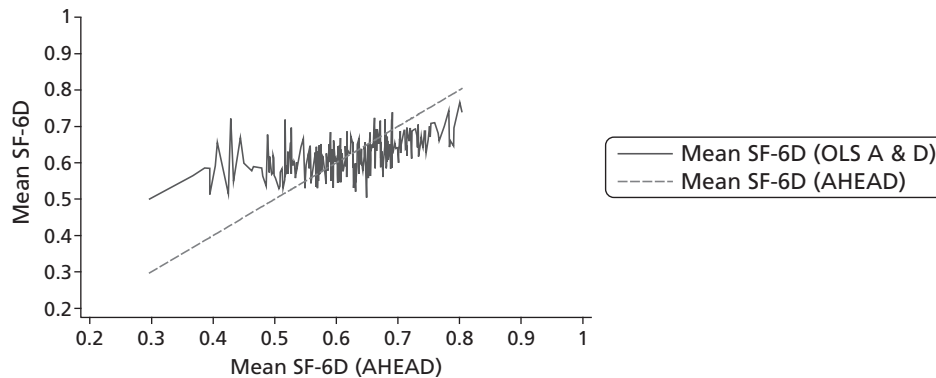
FIGURE 6 Predicted and observed SF-6D – HADS (AHEAD).

TABLE 19 Summary performance statistics: OLS regressions (AHEAD – SF-6D)

		Predicted SF-6D			
		Model 4		Model 5	
Independent variables	Observed SF-6D	HADS-A and HADS-D		HADS-A and HADS-D + age	
Summary statistics					
Mean	0.6118	0.6118		0.6118	
SD	0.089	0.049		0.051	
Minimum	0.2968	0.5003		0.4942	
Maximum	0.8037	0.766		0.7645	
Regression performance statistics					
R^2		0.305		0.326	
Adjusted R^2		0.301		0.319	
RMSE		0.059		0.057	
BIC		–660		–663	
SF-6D range	***n***	**Mean SF-6D**	**Mean SF-6D**	**RMSE**	**Mean SF-6D**	**RMSE**
< 0.5	27	0.4478	0.5885	0.141	0.5816	0.134
0.5–0.599	82	0.5556	0.5963	0.050	0.5956	0.049
0.6–0.699	124	0.6496	0.6214	0.042	0.623	0.042
0.7–0.804	53	0.7379	0.6692	0.069	0.6702	0.068
HADS-D	***n***	**Mean SF-6D**	**Mean SF-6D**	**RMSE**	**Mean SF-6D**	**RMSE**
non-case/possible ≤ 10	147	0.6584	0.6575	0.059	0.6574	0.058
Probable case >10	139	0.5793	0.5803	0.058	0.5804	0.057

Regression results are reported in *Appendix 3*, *Table 50*.

Patient Health Questionnaire and SF-6D

The best-fitting model from the PHQ-9 regressions based on the BIC was the item-level model with significant items only; however, the full item-level model performed better in terms of predictive ability and results from the latter model are therefore reported. The coefficients on most of the items were negative and statistically significant. Items related to suicidal thoughts and 'feeling bad about yourself' were positive, which is contrary to expectations as this indicates that reporting these symptoms is associated with higher SF-6D scores. *Figure 7* shows the predicted and observed SF-6D scores indicating some association. R^2 was high in the item-level models (*Table 20*, model 6). There was evidence of overpredicting severe health problems and underpredicting for those in better health. RMSEs are similar at both ends of the SF-6D scale. Predicted SF-6D scores are able to discriminate across patients with depression with RMSEs < 0.1. As expected, Tobit models did not offer any advantage over the best-fitting OLS model as there were no individuals at the top of the SF-6D scale (see *Appendix 3*, *Table 46*). Response-level mapping models also performed poorly relative to OLS, with large RMSEs at the severe end of the scale (see *Appendix 3*, *Table 47*). This was mainly because dimensions and items had to be collapsed into smaller numbers of levels in order for the models to run.
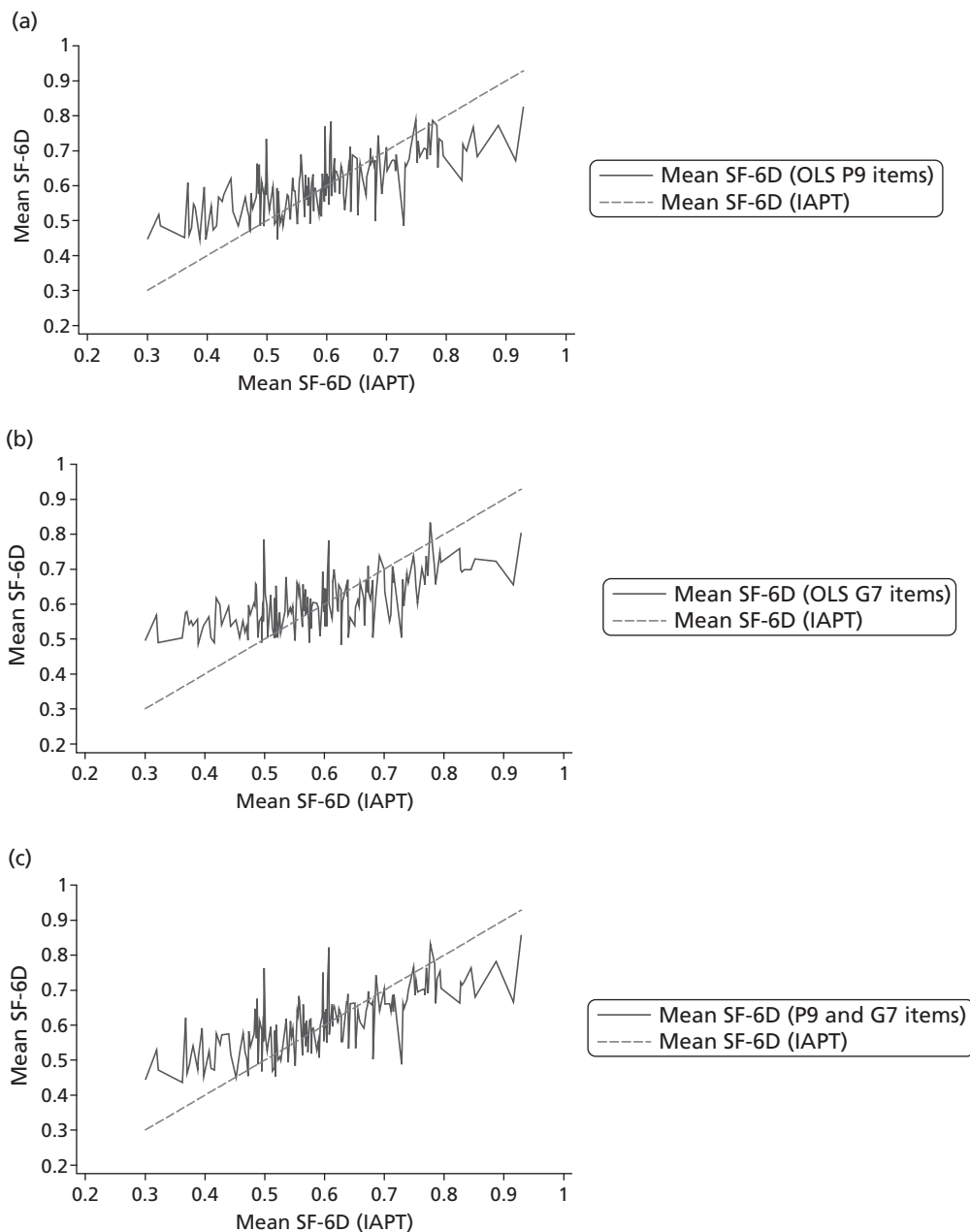
FIGURE 7 Predicted and observed SF-6D – PHQ-9 and GAD-7 (IAPT). (a) PHQ-9 items; (b) GAD-7 items; and (c) PHQ-9 and GAD-7 items.

TABLE 20 Summary performance statistics: OLS regressions (IAPT – SF-6D and PHQ-9, GAD-7)

Independent variables	Observed SF-6D (n = 394)	Predicted SF-6D					
		Model 6 PHQ-9 items		**Model 7** GAD-7 items		**Model 8** PHQ-9 and GAD-7 items	
Summary statistics							
Mean	0.6193	0.6193		0.6193		0.6193	
SD	0.126	0.091		0.078		0.093	
Minimum	0.301	0.4242		0.4834		0.4079	
Maximum	0.929	0.8595		0.8336		0.8719	
Regression performance statistics							
R^2		0.516		0.383		0.538	
Adjusted R^2		0.476		0.342		0.487	
RMSE		0.070		0.079		0.067	
BIC		–613		–554		–578	
SF-6D range	**n**	**Mean SF-6D**	**RMSE**	**Mean SF-6D**	**RMSE**	**Mean SF-6D**	**RMSE**
< 0.4	20	0.3610	0.4929	0.132	0.5319	0.171	0.4944	0.133
0.4–0.499	51	0.4690	0.5592	0.093	0.5699	0.101	0.5537	0.088
0.5–0.599	100	0.5587	0.5692	0.053	0.5871	0.056	0.5682	0.053
0.6–0.699	125	0.6386	0.6354	0.051	0.625	0.054	0.6344	0.048
0.7–0.799	60	0.7444	0.687	0.066	0.6747	0.079	0.6904	0.064
0.8–1	38	0.8552	0.738	0.119	0.7099	0.145	0.7453	0.111
PHQ-9 cut-offs	**n**	**Mean SF-6D**	**RMSE**	**Mean SF-6D**	**RMSE**	**Mean SF-6D**	**RMSE**
None/mild	90	0.7272	0.7273	0.073	0.6934	0.085	0.726	0.069
Moderate	84	0.6501	0.6475	0.071	0.6481	0.079	0.6498	0.068
Severe	220	0.5634	0.5643	0.068	0.578	0.076	0.564	0.066

Regression results are reported in *Appendix 3*, *Tables 51–53*.

Generalised Anxiety Disorder Assessment and SF-6D

The GAD-7 OLS regression results indicated that there was poor association between this measure and the SF-6D. The best-fitting model was the full-item model, but most of the items were not statistically significant, and those that were are mainly negative, as expected. *Figure 8* shows that there was poor association between observed and predicted SF-6D scores. R^2 values were smaller than those of the PHQ-9. Maximum and minimum scores were lower and higher than those of the PHQ-9 (see *Table 20*, model 7). RMSEs were therefore also larger. Despite evidence of poor predictive performance, predicted scores are able to discriminate across depression severity groups. Similar to the PHQ-9 results, Tobit did not offer any advantage over OLS and the response-level mapping performed poorly (see *Appendix 3*, *Tables 46* and *47*).

Patient Health Questionnaire-9 items and Generalised Anxiety Disorder Assessment

We also assessed the use of both PHQ-9 and GAD-7 total scores and items to predict the SF-6D scores. The measures assess depression and anxiety which can occur as comorbidities, and so using both measures may capture more variation than each individual measure. Both measures are used in routine data collection in the IAPT services, so may provide a better specification for predicting SF-6D scores for this routine data. A number of items were excluded from the item-level models as there was a high degree of collinearity. Items that were statistically significant were mainly negative and were from the PHQ-9, with one item from the GAD. The best-fitting model based on BIC was the PHQ-9 significant item model, but the full item model (excluding highly collinear items) performed better in predictive terms and this is reported (model 8). This model shows no evidence of heteroscedasticity or misspecification, although there is some evidence of collinearity. Combining these measures offered slight improvements in terms of predictions over the PHQ-9 predictions [see *Table 20*, model 8 and *Figure 7(c)*].
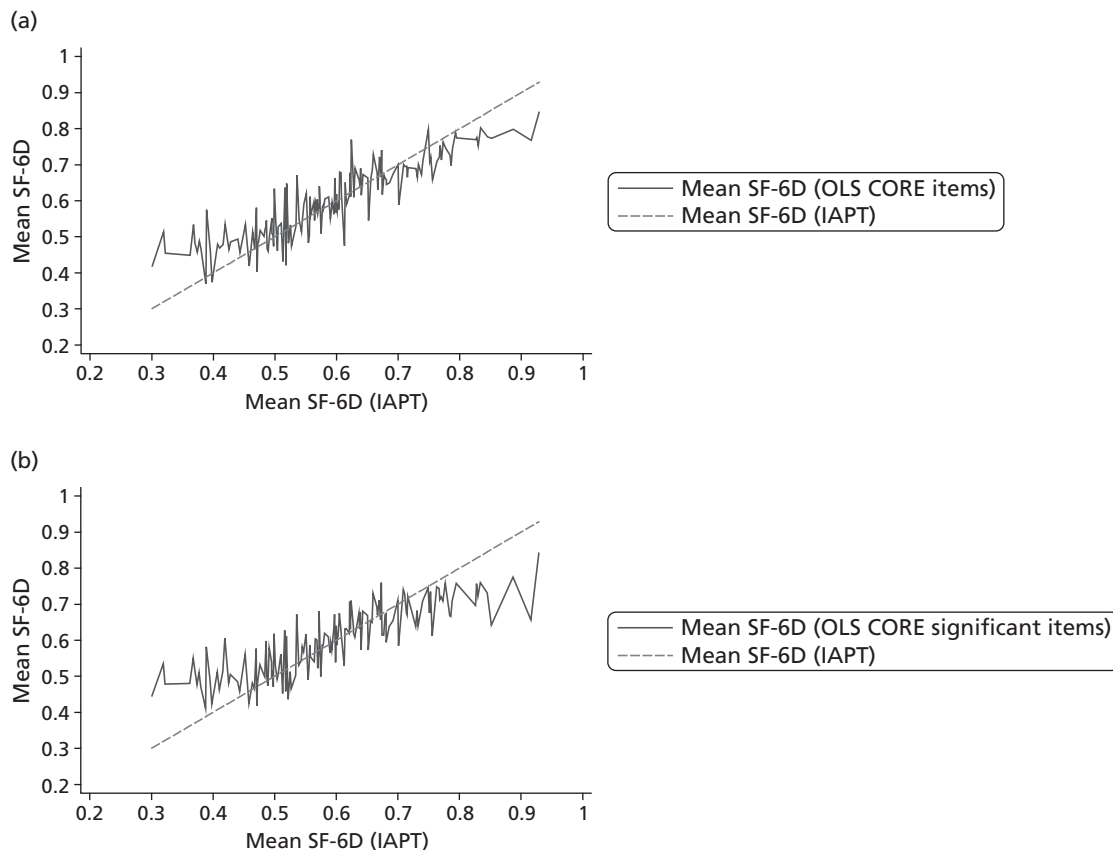
FIGURE 8 Predicted and observed SF-6D – CORE-OM (IAPT). (a) CORE-OM items; and (b) CORE-OM significant items.

Clinical Outcomes in Routine Evaluation – Outcome Measure and SF-6D

The OLS results for the dimension scores of the CORE-OM indicate that only the symptoms dimension was statistically significant. As the CORE dimensions have a short range (0 to 4, but this includes values between whole numbers), a different specification using the dimension scores as levels was tested. This did not improve the results, with weak associations between predicted and observed SF-6D scores. CORE-OM items provided better-fitting models. Items from all four dimensions (well-being, symptoms, functioning and risk) were statistically significant in the item-level regressions. There was evidence of a strong association between the observed and predicted scores using CORE items (see *Figure 8*). R^2 values were high (> 0.5), although adjusted R^2 values were slightly lower (*Table 21*). There was some evidence of misspecification for the CORE item models as well as collinearity. Although the CORE models had better R^2, they overpredicted SF-6D scores more than the PHQ-9 models at the severe end of the health scale as well as underpredicting better health. Predicted SF-6D scores can discriminate between those with and without depression. Tobit and response-level models were tested but these did not offer improvements over OLS (see *Appendix 3*, *Tables 46* and *47*).

TABLE 21 Summary performance statistics: OLS regressions (IAPT – SF-6D and CORE-OM)

		Predicted SF-6D			
		Model 9		**Model 10**	
Independent variables	**Observed SF-6D** (n = 394)	**CORE items**		**CORE significant items**	
Summary statistics					
Mean	0.6193	0.6193		0.6193	
SD	0.126	0.107		0.099	
Minimum	0.301	0.3639		0.4113	
Maximum	0.929	0.8898		0.8528	
Regression performance statistics					
R^2		0.714		0.609	
Adjusted R^2		0.576		0.574	
RMSE		0.054		0.062	
BIC		−235		−686	
SF-6D range	**n**	**Mean SF-6D**	**Mean SF-6D**	**RMSE**	**Mean SF-6D**	**RMSE**
< 0.4	20	0.3610	0.4593	0.103	0.4855	0.125
0.4–0.499	51	0.4690	0.4971	0.046	0.5087	0.053
0.5–0.599	100	0.5587	0.5685	0.047	0.5729	0.050
0.6–0.699	125	0.6386	0.6442	0.047	0.6456	0.051
0.7–0.799	60	0.7444	0.7079	0.053	0.7005	0.057
0.8–1	38	0.8552	0.779	0.080	0.745	0.111
PHQ-9 cut-offs	**n**	**Mean SF-6D**	**Mean SF-6D**	**RMSE**	**Mean SF-6D**	**RMSE**
None/mild	90	0.7272	0.7156	0.061	0.7063	0.071
Moderate	84	0.6501	0.6508	0.053	0.6483	0.064
Severe	220	0.5634	0.5678	0.051	0.5726	0.057

Regression results are reported in *Appendix 3*, *Table 55*.

General Health Questionnaire and SF-6D

The GHQ-12 item-level model performed better than the total score model. The coefficients were mainly negative, as expected, but the plots of observed and predicted SF-6D scores show that there is weak association in these values (*Figure 9*). R^2 and adjusted R^2 values were low as expected. Exclusion of correlated items did not improve the GHQ-12 item-level models. There was no evidence of misspecification, but the White's test indicated the presence of heteroscedasticity. There was evidence of over- and underpredicting of the SF-6D scores, with large differences between observed and predicted SF-6D scores at the minimum and maximum values (*Table 22*). RMSE values ranged from 0.116 to 0.166 at the top and bottom of the SF-6D scale; these values were not as large as those reported by some of the models for other measures. Small RMSEs might be driven by small numbers at the top and bottom of the SF-6D range. The predicted SF-6D scores did not vary significantly based on CIS-R severity levels. Tobit and response-level models were not estimated as they were unlikely to improve on the poorly performing OLS models.

Validation of best-performing models

Models 1 to 8 (HADS, PHQ-9, GAD-7) and model 11 (GHQ-12) were poor predictors of EQ-5D and SF-6D scores and these were therefore not validated. The CORE-OM items models (models 9 and 10) performed reasonably in estimating SF-6D scores within the sample. Validation of these models was therefore undertaken using 75% of the sample to estimate and 25% of the sample to validate the models. Results are reported in *Table 23*. Mean predicted SF-6D values were lower than observed values. As with in-sample predictions, the RMSEs showed evidence of over- and underpredicting severe and better health states. The RMSEs were larger across the EQ-5D range and severity measures, although focusing on significant items improves the model's predictive abilities. This is evident in *Figure 10*, which shows observed and predicted scores with evidence of poor association, particularly at both ends of the SF-6D scale. However, a larger sample may be required to validate this model, as although the sample matches the estimation sample in terms of distribution of scores, there are very few individuals at the extremes of the scale.
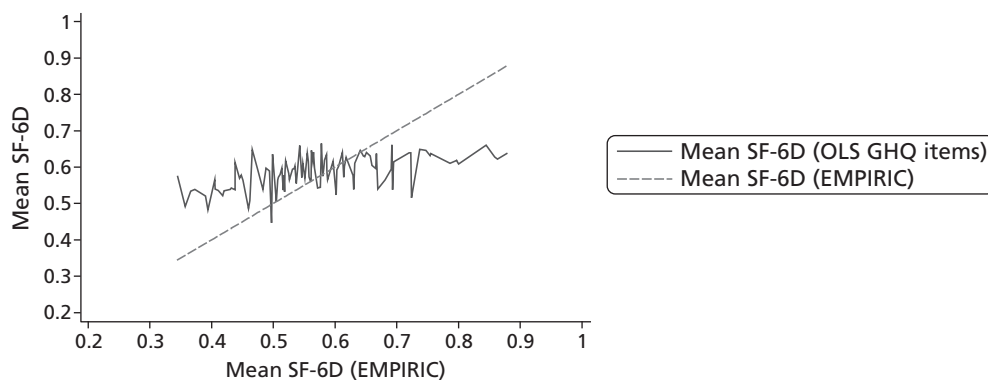
FIGURE 9 Predicted and observed SF-6D – GHQ-12 (EMPIRIC).

TABLE 22 Summary performance statistics: OLS regressions (EMPIRIC – SF-6D)

		Predicted SF-6D	
		Model 11	
Independent variables	**Observed SF-6D ($n = 213$)**	**GHQ items**	
Summary statistics			
Mean	0.5895	0.5895	
SD	0.111	0.051	
Minimum	0.345	0.4329	
Maximum	0.878	0.7231	
Regression performance statistics			
R^2		0.209	
Adjusted R^2		0.047	
RMSE		0.077	
BIC		−185	
SF-6D range	**n**	**Mean SF-6D**	**GHQ items**	**RMSE**
< 0.5	38	0.4341	0.5469	0.116
0.5 to 0.599	82	0.5515	0.5862	0.050
0.6 to 0.699	65	0.6421	0.6039	0.051
0.7 to 0.878	28	0.7897	0.6237	0.166
CIS	**n**	**Mean SF-6D**	**GHQ items**	**RMSE**
12 to 17	43	0.6306	0.6061	0.087
18 to 23	58	0.6231	0.6057	0.082
24 to 29	51	0.5858	0.576	0.072
30 to 35	45	0.5374	0.5779	0.073
36+	16	0.5159	0.562	0.062

CIS, Clinical Interview Schedule.

TABLE 23 Summary performance statistics: OLS regressions (IAPT – CORE-OM items) out-of-sample estimation

Independent variables	Observed SF-6D (*n* = 99)	Predicted SF-6D			
		Model 9 CORE items		Model 10 CORE significant items	
Mean	0.6208	0.6165		0.6159	
SD	0.125	0.117		0.092	
Minimum	0.301	0.3555		0.4387	
Maximum	0.916	0.8858		0.7883	
SF-6D range	***n***	**Mean SF-6D**	**Mean SF-6D**	**RMSE**	**Mean SF-6D**	**RMSE**
< 0.4	5	0.3526	0.5357	0.183	0.5033	0.151
0.4–0.499	13	0.4801	0.5085	0.055	0.5223	0.055
0.5–0.599	25	0.5602	0.5705	0.087	0.5727	0.060
0.6–0.699	31	0.6413	0.6667	0.072	0.6528	0.049
0.7–0.799	15	0.7368	0.6465	0.113	0.6777	0.068
0.8–1	10	0.8518	0.7122	0.144	0.6948	0.157
Total		0.6208	0.6165	0.092	0.6159	0.071
PHQ-9 cut-offs	***n***	**Mean SF-6D**	**Mean SF-6D**	**RMSE**	**Mean SF-6D**	**RMSE**
None/mild	19	0.7144	0.7002	0.091	0.6864	0.065
Moderate	19	0.6594	0.6501	0.097	0.6568	0.088
Severe	61	0.5796	0.58	0.091	0.5812	0.068
Total		0.6208	0.6165	0.092	0.6159	0.071
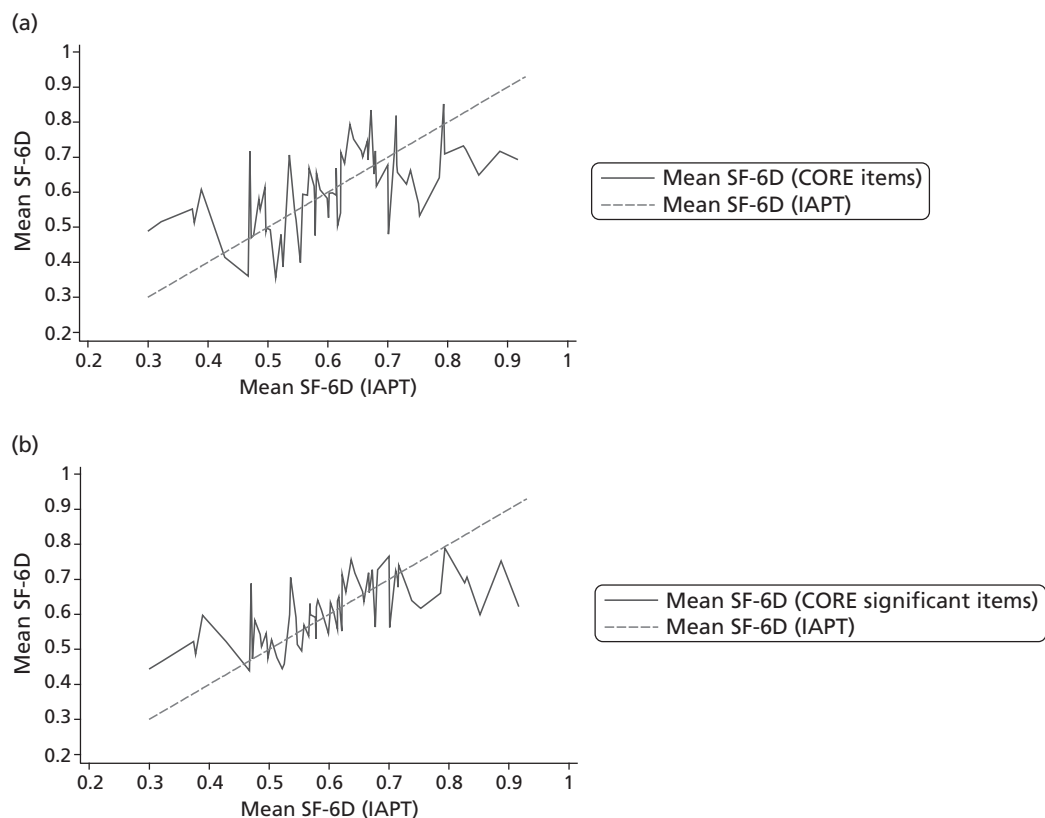
FIGURE 10 Predicted and observed scores: SF-6D and CORE-OM (IAPT) out-of-sample validation. (a) CORE-OM items (*n* = 99); and (b) CORE-OM significant items (*n* = 99).

Discussion

This chapter tested different functions for mapping from commonly used condition-specific measures for depression and anxiety to two generic preference-based measures, the EQ-5D and SF-6D. The results indicate that commonly used mental health measures for depression and anxiety are not good predictors of EQ-5D and SF-6D scores. This is not surprising as most of the mental health measures focus mainly on mental health, and often on very narrow aspects of symptoms of mental health, whereas the generic measures have a wider scope. For example, the HADS, PHQ-9 and GAD-7 focus on depression or anxiety but there are no questions on how this may impact on normal activities in the items included in total scores. The PHQ-9 and GAD-7 include an additional question on normal activities, which does not form part of the total score but is used in the item-level regressions. This was an important predictor of SF-6D scores. The GHQ-12 has a single question on the impact of symptoms on daily activities, but the other 11 questions concern symptoms. Both the EQ-5D and SF-6D have dimensions related to mental health but they also cover other aspects of health, such as physical (mobility) and social functioning and normal activities (role limitation, usual activities, self-care) as well as pain. There is, therefore, poor conceptual overlap between these condition-specific measures and the generic preference-based measures, which can be a problem when mapping from condition-specific measures to generic preference-based measures.[178] The CORE-OM provides a better predictor of SF-6D scores than the other measures. This is mainly because it is not just focused on mental health symptoms, but includes general and social functioning as well as physical problems. Model validation of the CORE item-level models indicates that the model with significant items may be a solution to providing utility values where utility values are not available. However, further validation is required as the small sample size may be driving some of the results. There may also be concerns regarding the transferability of this mapping function as the full range of SF-6D scores was not available in the estimation sample.

Item-level models used to predict the EQ-5D/SF-6D scores fit better than the total or dimension score models. However, the latter are more likely to be used as total and dimension scores are reported in publications. Mortimer *et al.*[191] compared the performance of item-based and dimension-based mapping and concluded that the item-based approaches may be restrictive owing to data requirements. They found that scale/dimension-based approaches offer similar results to item-based approaches where there is conceptual overlap between measures.

Predicting generic preference-based measure dimensions instead of scores did not improve the results. The relatively small sample sizes used in our analysis compared to other mapping studies do not provide enough data to fit these types of regressions, particularly using condition-specific measure items. Generic preference-based measure dimension levels and condition-specific measure item levels had to be collapsed in order to fit the models, and this results in loss of information. Gray *et al.*,[181] who developed the response mapping approach, had large data sets ($n > 10,000$) which allowed them to carry out these regressions without loss of data. Furthermore, the lack of conceptual overlap is magnified when fitting the generic preference-based measure dimensions as condition-specific measures are poor at predicting the majority of these.

Ordinary least squares provides the best model fit in this mapping analysis. The range covered in the observed generic preference-based measures in the samples used did not include many individuals at full health for the EQ-5D and none for the SF-6D. Tobit models which have been used in other mapping studies to take into account the bounded nature of these generic measures did not improve the results. Two-part models used to take into account the large number of individuals who report full health could not be carried out. As with other mapping studies, models overpredicted severe health problems and underpredicted better health. One potential solution to the overpredicting of severe health problems (which tends to be greater than the underprediction of better health) is to use two separate regressions: one for those with severe health problems and another for those in better health.[192] A spline which takes into account differences in the relationship between predictor variables and predicted values as the latter changes may also provide a solution. The EQ-5D shows a bimodal distribution in both AHEAD and

POPMACT samples at the lower end of the scale and so a two-part model may have improved performance; however, given the poor predicting properties of the HADS, this was not tested. Development of mapping functions for measures such as the CORE-OM to the EQ-5D should test such an approach.

A recent review of studies on mapping from condition-specific measures to the EQ-5D[193] found only one published study in mental health, by Serrano-Aguilar *et al.*[194] This study developed a mapping function for the GHQ-12 to the EQ-5D using data from a representative random sample in the Canary Islands (*n* = 3567). Mean EQ-5D scores were high in this study (0.82) and only 19% of the sample showed evidence of psychiatric problems based on the GHQ-12. Validation indicated that the mapping function underpredicted the top end of the scale, with mean differences between observed and predicted scores of 0.25; however, predicted values were able to discriminate between psychiatric morbidity cases. RMSEs were not reported. The lack of other published studies on mapping in mental health may be an indicator of the poor performance of mental health condition-specific measures. However, it may also result from the lack of available data to perform mapping, particularly for the EQ-5D.[195] Future work should consider identifying studies in which both generic preference-based measures and mental health measures for depression and anxiety have been used, in order to perform further testing of mapping algorithms.

Overall, this chapter shows that mapping from mental health condition-specific measures, such as the widely used PHQ-9, GAD and HADS, may not be an appropriate approach to generating EQ-5D and SF-6D scores as these measures focus on specific symptoms and not on the wider impact of mental health conditions. The CORE-OM was better at predicting the SF-6D, but there were still problems in terms of under- and overpredicting. The poor performance of the mapping functions may also be a consequence of the generic measures not being sufficiently sensitive to the range of mental health problems reflected in the condition-specific measures. The way mental health impacts on quality of life is examined in more depth through the use of qualitative techniques in the next two chapters.

Chapter 5 A synthesis of qualitative research on the quality of life of people with mental health problems

The review of the psychometric evidence presented in *Chapter 2* and the analysis of the seven data sets in *Chapter 3* found that the EQ-5D and SF-36 achieved an adequate level of performance in terms of the quantitative psychometric tests of validity and responsiveness in depression, and to some extent in anxiety and personality disorder. The mapping work in the previous chapter raises some additional concerns as to whether or not the EQ-5D in particular is able to fully reflect the specific impact of common mental health problems. Results from the psychometric analyses were mixed in schizophrenia and bipolar disorder, with a suggestion that the EQ-5D and SF-36 may be reflecting depression rather than other consequences of these conditions. It was not possible to determine the reason for the mixed results from these analyses. This was the rationale for looking at qualitative research, as this provides evidence on the content validity of these measures in people with mental health problems.

The overall purpose of the qualitative components of this project is to identify the domains that will be used to assess the content validity of the generic measures presented in *Chapter 7*. The aim of this systematic review of qualitative research was to begin to identify the quality of life domains that are important from the perspective of an individual with mental health problems. It provides an important input into the topic guide used in the interviews reported in *Chapter 6*, and a basis for starting to identify the core themes of quality of life reported by people with mental health problems.

The focus of this review was primary qualitative research studies which, using methods such as interviews and focus groups, explicitly asked adults with mental health problems what they considered to be important to their quality of life or how their quality of life had been affected by their mental health problems. A range of approaches is available for synthesising qualitative research.[196] Paterson *et al.*[197] recommend that the choice is made based on the nature of the research question and design, the prevailing paradigm and the researcher's personal preference. In this review, framework synthesis was used. This is based on the 'framework' approach for the analysis of primary data[198] and is a structured approach to organising and analysing data which permits the expansion and refinement of an a priori framework to incorporate new themes emerging from the data.[196]

This chapter reports a summary of the main findings of the review. A more detailed account is provided in Connell *et al.* 2012.[199]

Methods

Search methods

Systematic reviews of evidence of clinical effectiveness require extensive searching based on a clearly focused search question. Defining a focused question was not possible nor appropriate here, given the exploratory and inductive nature of the review process. An iterative approach to searching was used, in order to accommodate within the search process new themes emerging throughout the review. Extensive searching was undertaken, using a number of search techniques. Database searches were undertaken between October 2009 and April 2010 and included MEDLINE, Applied Social Sciences Index and Abstracts (ASSIA), Cumulative Index to Nursing and Allied Health Literature (CINAHL), PsycINFO and Web of Science. Search techniques included keyword searching, taking advice from experts, hand searching, citation searching of relevant references and internet searching. Four iterations of searching were undertaken (see *Appendix 4, Table 56*). The searches were not restricted by date, language or country. Key search terms used were mental health, mental illness, mental disorder, quality of life, well-being, life

satisfaction, life functioning, life change, recovery, subjective experience, lived experience, lifestyle, coping, adaptation, qualitative and qualitative research. For a full list of search terms and details of the search iterations see *Appendix 4*.

Inclusion and exclusion criteria

Quality of life

The review started from a premise of not imposing a preconceived definition of 'quality of life'. Some studies identified had an explicit aim to explore quality of life whereas others did not. Complexities arose in deciding which studies were about the same substantive concept of quality of life and which were tapping into a separate but overlapping concept. Sandelowski[200] states that 'often research purposes and questions are so broadly stated it is only by looking at the kinds of findings produced that topical similarity can be determined'. Taking this approach, it was evident that many studies had similar findings even though quality of life was not the subject of investigation. The concepts examined in these studies were recovery, lived experience, subjective experience, psychosocial issues, health needs, and strategies for living. Inclusion of these studies had the potential to introduce themes that were not central to the concept of quality of life but were rather allied to a separate but related concept. A pragmatic decision was made to examine the research aims and interview questions of those studies that did not directly investigate the concept of quality of life, and to include those that asked broad, open-ended questions about how participants' mental health affected their lives, or what was important to or would improve their lives, as well as those studies that equated their findings with quality of life in some way. Studies were excluded if they deliberately started with a premise assuming the importance of any particular domain to quality of life, or if they were structured around a preconceived list of domains.

Qualitative research

We included primary qualitative research studies that used qualitative interview or focus group data to collect the views of individuals with mental health problems. We excluded studies that used content analysis and presented results as a frequency list with no supporting participant quotes. Some studies sought the views of a range of people with mental health problems, carers and professionals; in such cases, we only included those studies in which the views of people with mental health problems could be separately identified.

Mental health

We included research on all mood disorders (e.g. depression, bipolar disorder and mania), neurosis and stress-related disorders [e.g. anxiety, phobias and post-traumatic stress disorder (PTSD)], personality disorders and schizophrenia, and schizotypal and delusional disorders. Included studies had to state that participants had mental health problems, as identified through either diagnosis or attendance at an establishment for people with mental health problems. Studies in which mental health problems were secondary to a physical health problem were excluded.

Quality

The use of quality assessment in reviews of qualitative research is contested. Quality assessment is usually used in framework synthesis but this may be associated with its use alongside systematic reviews of effectiveness.[196] In this review, articles were not quality assessed and systematically excluded on this basis. However, it was of paramount importance that any included study elicited the perspective of individuals with mental health problems, and where this appeared not to be the case the study was excluded. Consequently, studies were excluded when it was strongly suspected that the views of the researcher, or the method of analysis, had overly influenced the findings. These articles were examined and discussed at length by the research team before being excluded.

Language

Although the searches were restricted to English-language articles, non-English-language articles were excluded because of the potential for misinterpretation. Five potentially relevant articles were excluded on the grounds of language. *Figure 11* shows a PRISMA flow diagram of the searches undertaken.

Data extraction and analysis

The following details of the studies were extracted: mental health problem studied, author affiliation, time and location of study, number and demographic details of participants, research aims and questions, recruitment and sampling methods, method of data collection and analysis, and themes within the findings and discussion sections of the papers.

Framework analysis[198] was used to allow the identification of common and variable patterns of themes within and across different studies. The first stage of framework analysis – familiarisation – was undertaken by reading all included papers. The second stage involved examining the findings from these papers to identify *initial* themes for a thematic framework. These 10 initial descriptive themes were either identified as main themes from more than one study or arose consistently across studies, and were as follows: activity, relationships, the self, the future/aspirations, symptoms/well-being/emotions, spirituality, control/coping, insight/education, health-care services/interventions, and resources/basic needs. The third stage – data organisation – involved charting data from the findings and discussion sections of papers that corresponded to each theme. Text was transferred verbatim to ensure contextual accuracy. It was common for text to be identified as supporting more than one theme, for example a quote describing how work was good for the individual's self-esteem would be placed in the thematic categories 'activity' and 'self'. At the next stage each initial theme was examined, and further subthemes identified and documented within the framework chart. To assist with the final stage of framework analysis – mapping – the subthemes were listed and examined for their *conceptual* similarities and differences. To aid this process, we searched the
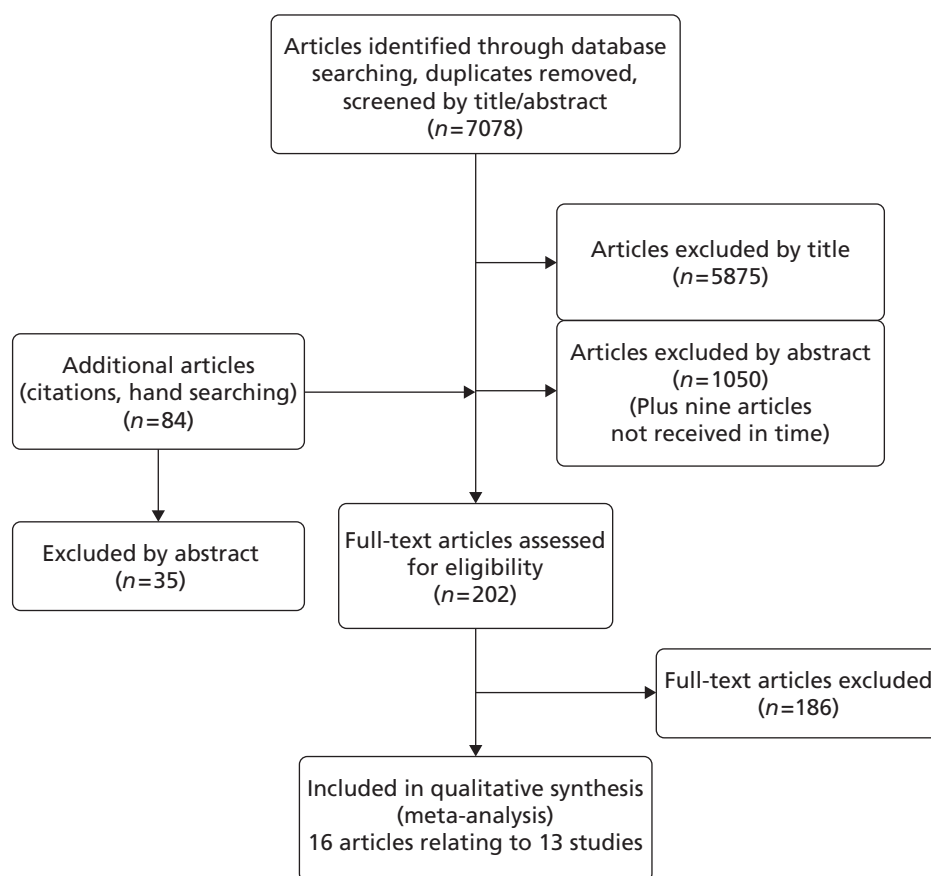
FIGURE 11 PRISMA flow diagram of searched articles.

wider literature to find papers which would help us to understand the data and make connections between subthemes, and would assist in the development of our final themes. For example, 'belonging' was an emerging theme, and we identified Hagerty *et al.*'s[201] research which explored and defined this concept. We then returned to our thematic framework to re-examine our data in the light of the wider literature. Other influential literature at this stage was on the theory of 'doing, being, becoming',[202] ill-being compared with well-being and intrinsic and extrinsic quality of life,[203,204] and demoralisation.[205]

Validation and trustworthiness

Validation procedures were incorporated into the review at all stages. Two researchers (JC and MLJ) independently identified articles from the first search iteration, and compared results to clarify the inclusion and exclusion criteria. Potential full articles were identified from further searches by the primary researcher and independently checked by the second researcher. The included articles were examined independently by both researchers to identify the main themes for the initial framework. Disagreements at all stages were resolved by discussion. Additionally, a multidisciplinary team of researchers met regularly in addition to meetings with clinicians and a user representative to discuss and challenge the inclusion and exclusion criteria, thematic framework, and conceptual interpretations and conclusions.

Findings

Description of included studies

Thirteen studies were identified from 16 articles,[206–220] one of which was unpublished (E Chambers and S Cook, Sheffield Hallam University, 2012; hereafter cited as Chambers 2012). Two of these journal articles[208,209] had fuller reports available.[Chambers 2012,206] The fuller reports only have been referenced in the findings. A further study indicated that not all emerging themes were presented in the paper and had a supplementary paper dedicated to the impact of bipolar disorder on work functioning, which was included in our analysis.[207] The studies were published between 1994 and 2010 in a number of countries: Canada[209,210,212,215–218] UK,[208,211,217] Sweden,[214,219] USA[220] Austrailia[216] and New Zealand.[213] The professional affiliations of the first author were occupational therapy,[208,209,212,215,217] nursing,[210,213,214,219] psychology,[211,216] psychiatry[218] and social work.[219]

The mental health disorder most frequently represented was schizophrenia (or other psychotic disorder); this was the only population researched in three studies[211,212,215] and the majority population in a further two studies.[Chambers 2012,220] Three studies included individuals with bipolar disorder only[216,218,219] and one included those with panic disorder only.[213] Other studies[206,210] had a mixed population including the above disorders plus persons with personality disorder, severe depression and anxiety disorders. Two studies did not specify the disorder; they included persons described as having 'enduring mental health problems'[217] and 'psychiatric disability'.[214]

Two studies had a primarily positive orientation in that they asked 'what is required for a good quality of life?',[218,219] and four showed a negative orientation through asking 'how has your mental health affected your quality of life?'.[210,212,213,216] The remainder considered both 'what had helped and hindered quality of life'. Four studies had a conceptual orientation;[210,212,213,219] the remainder presented their findings descriptively. Further details of the studies[Chambers 2012,206,207,210–220] can be found in *Appendix 4* (see *Table 57*).

We identified six major themes from the initial descriptive themes: well-being and ill-being; control, autonomy and choice; self-perception; belonging; activity; and hope and hopelessness. These themes are now explored in detail.

Well-being and ill-being

Well-being has long been regarded as an important dimension of HRQL scales (see *Chapter 1*). The emotional component of subjective well-being consists of high levels of positive affect (experiencing pleasant emotions and moods), and a lack of low levels of negative affect (experiencing few unpleasant

emotions and moods).[204] Within our papers, symptoms of mental illness and aspects of emotional well-being were intertwined, with an emphasis on the negative rather than the positive. This suggested that ill-being, which is more akin to distress and the symptoms of mental illness, is an important aspect of quality of life for those with severe mental health problems. The most evident 'ill-being' themes were general feelings of distress, experience of psychosis/mania, depressed mood, fear and anxiety, and problems with energy and motivation and fear and anxiety.

Distress from symptoms

Distress, or the subjective experience of the symptoms of mental illness, was evident in the majority of studies[147,206,211–213,216,217,220] and a major theme in four.[206,211,216,220] The subjective experience of mental illness was described as wretched,[217] a burden, debilitating, painful,[220] tormenting[216] and having a tyrannical power over life.[206] Preoccupation with the symptoms of mental health problems interfered greatly with the most basic tasks of everyday living,[Chambers 2012,206,212,220] making it difficult to deal with anything but the present moment.[220] Instead, life was consumed with coping on a daily basis and living 'one day at a time' – sometimes on a moment-to-moment basis.[206,212,215]

Experience of psychosis/mania

Distressing symptoms reported included hallucinations and delusions (particularly hearing voices, thought disturbances and paranoia),[Chambers 2012,206,211] reality disorientation,[206] mania and hypomania,[207] feelings of discomfort, weirdness or oddness,[206] and irritability or agitation.[211] These symptoms could interfere directly with day-to-day living by having an effect on behaviour control,[Chambers 2012,207,211,216] concentration, memory or decision-making[Chambers 2012,211,212] and sense of self-identity.[206,212,218]

> *When I hear voices erm, that stops me from doing a day to day existence, I'm preoccupied with the voices[.]*

> *. . . the voices, how they've affected my life, erm, er just day to day living basically . . . Erm just er, getting out, getting out and doing things er . . . go to the shops, erm, erm, cooking, anything, anything like that[.]*

> *I daren't go out now, thoughts in my head, make me think bad things, I get paranoid when there's crowds of people[.]*

> *Chambers, 2012*

Depressive mood

Depression was a diagnosis in a proportion of participants in two of the studies[206,210] and bipolar disorder was the primary diagnosis in three further studies.[216,218,219] Negative affect, in the more severe form of depression including feeling suicidal (as opposed to simply being sad and unhappy), was also identified in studies in which the primary diagnosis was psychosis related.[Chambers 2012,211,215,217,220] It was also the symptoms of depression in bipolar patients that were reported as being particularly distressing,[216] together with the unpredictability and instability of mood.[207,216,218]

Energy and motivation

Depression was often expressed as associated with a lack of energy and/or motivation. Although energy and motivation might be regarded as two distinct concepts (physical and psychological), they were closely associated and for the most part reported together within the primary research. Energy, or lack of it, was a major theme in one study[219] and all but three of the primary research articles[213,214,216] described the debilitating effects of lack of energy. The three studies in which energy and motivation were not evident focused on the nursing implications of panic disorder,[213] the psychosocial issues related to bipolar disorder[216] and the positive determinants of health.[214] Participants reported feeling generally drained of energy[Chambers 2012,206,207,210,211,220] associated with a lack of motivation, enthusiasm, or interest in things.[Chambers 2012,206,210,211,215,217,220] The side effects of medication[Chambers 2012] or problems with sleep[Chambers 2012,211] were reported as having a causal effect. Whereas lack of energy was the dominant

theme, hypomanic states in bipolar disorder were associated with increased energy and enthusiasm but were often short-lived, with a return to a usual depressed state.[207]

Fear and anxiety

Two studies reported that 'fear' was a theme that was represented throughout its interview data.[206,211] Fear, anxiety or worry was present in some form in all of the studies. The subjective experiences of the symptoms were reported as being very frightening.[Chambers 2012,206,212,213,216,220] This tended to be identified in the studies on schizophrenia, bipolar disorder and panic disorder. As a consequence, individuals lived in fear of relapse or a return to hospital.[Chambers 2012,206,211] There were associated financial worries which had implications for planning for the future and making commitments.[211,215,218]

Anxiety in social situations was especially evident and took various forms, including anxiety about leaving the house, crowds and public places,[Chambers 2012] concerns for their own safety[206,217] and that of others,[Chambers 2012] worrying about what others thought of them and how they appeared,[Chambers 2012,206] and fears of rejection.[220] Worries concerning relapse or aggravation of symptoms and social anxiety often resulted in the avoidance of any activity or situation which might be perceived as stressful,[Chambers 2012,206,211,214] thus limiting the possibilities of improving other aspects of quality of life. Individuals reported avoiding situations they had previously enjoyed because of fear of how they would appear or that the stress associated with those situations would mean deterioration in mental health:

> *I've cut down on the sort of positions I get myself in . . . because of bad experiences in the past . . . you just try less things with the fear that you're going to get very ill again and go to hospital[.]*
>
> *Gee* et al.[211]

Well-being

There tended to be an emphasis on the absence of ill-being rather than the existence of well-being. However, positive themes were identified that included overall sense of well-being,[212,214,219] feeling healthy,[212] peaceful, calm and relaxed,[Chambers 2012,211,214,219] stable,[211,216] safe[214,219] and free from worry and demands.[214,219]

Physical well-being

Physical health was not a strong theme within the reviewed studies (though, as reported in the next chapter, it was a stronger theme in the interviews). The compounding effects of physical problems were indicated in two studies[Chambers 2012,217] and physical health was listed as the second most important aspect of quality of life, alongside mental health, in another.[218] A healthy lifestyle was considered beneficial; this included exercise, avoiding drugs and generally taking care of oneself.[Chambers 2012,206,214]

Control, autonomy and choice

The importance of aspects of choice and control to quality of life was identified in eight studies[Chambers 2012,206,211,212,215,216,218,219] and was a main theme in three of these.[211,215,216] It was often discussed in the context of the availability of external resources which enabled choice and control, including medication and treatment, support, information and finances.

Symptom control

One of the most evident aspects of control was the relief and management of the most distressing or pervading aspects of mental illness.[Chambers 2012,206,207,212–216,219] This was usually described as being achieved through medication.[Chambers 2012,206,207,214–216,218,219] Having this control meant that individuals could move beyond 'the all-encompassing world of their illness'[206] and instead attend to other important areas of their lives.[206,212] However, medication could also have a detrimental effect on quality of life through side effects,[Chambers 2012,206,211,215] feelings of dependency[215,216] and fear of the consequences of not taking it.[211] It was therefore necessary to find the right medication to balance symptom management and side effects[Chambers 2012,206,213,214,216,219] as a means to a sense of well-being.[206]

I think for me, apparently the most important one is just managing the illness . . . different medications, side-effects, knowing what they are . . . for me there's been limited discomfort.

Corring[206]

The concept of control was particularly important for those with bipolar disorder, and was related to an inability to control or pre-empt the onset of mood episodes or their behaviour[207,216,218] and to a need for stability.[216]

Being informed and having an understanding of and insight into the illness was considered to be important.[Chambers 2012,206,213,215,219] To achieve this it was important to have an accurate diagnosis[213,214] and this also meant that people could receive effective medication,[214] knew what to expect for the future[206,214,219] and could develop strategies to manage their illness and deal with it better.[206,215] This was regarded as a first step on the way to recovery[213] and improving quality of life.[219]

Autonomy

There was a complex relationship between support, dependency and independence. Both support[Chambers 2012,206,212,214,215,219,220] and independence[213,214,218] were regarded as being important for quality of life. Support helped people manage their illness, access resources and increase their self-confidence.[214] However, it could also result in feelings of dependency,[Chambers 2012,218] with a resulting loss of a sense of control and self-esteem.[218] Hence there was a dilemma between wanting help and support and at the same time resenting it.[210] On the other hand, *choosing* to be dependent could enhance power and control.[219] Personal autonomy, finding the optimum balance between support and independence, was therefore important to quality of life.[208,214] There was also an aspiration for financial independence and autonomy from the health service, with a desire expressed by individuals to manage their own illness without relying on medication.[207,218]

Personal strength, determination and self-sufficiency were also regarded as important.[Chambers 2012,207,213,214] This meant people were able to make use of available resources and develop self-help and personal coping strategies,[Chambers 2012,206,213,218] which in turn promoted independence and a sense of control.[206,213]

Choice

The concept of choice was most associated with the availability of financial resources[Chambers 2012,206,211,214,215,218] and with limited employment opportunities.[Chambers 2012,206,207,211,215,220] Having sufficient financial resources meant people could more readily have a healthy lifestyle,[214] engage in activities that promoted well-being,[Chambers 2012,206,211,214] attain an optimum balance between dependency and independence,[Chambers 2012,218] have a choice in their surroundings[Chambers 2012,206,215] and be able to plan for the future.[211] Also of value was being able to choose whether or not to take part in things (particularly social activities),[206,215] flexible work conditions,[207] when and with whom to disclose mental illness,[215] and choices associated with mental health services, workers and interventions.[Chambers 2012]

Self-perception

A number of aspects of self associated with quality of life were identified: 'self-efficacy', having a belief and confidence in your own abilities; 'self-identity', having a perception of self and knowing who you are; and 'self-esteem', having a sense of self-worth and self-respect. These were linked to a further theme of self-acceptance. These 'self' concepts were closely associated and used interchangeably within the studies reviewed. Aspects of the self and self-perception were a major theme in three studies[206,213,216] and were present in some form within all of the other studies except one,[210] which had an abstract analytical style and only had undertones suggesting low self-esteem/image.

Self-identity

Problems related to having a sense of self and 'knowing who you are' appeared particularly to be related to bipolar disorder, schizophrenia and panic disorder. The studies described a loss of a sense of self, identity and personality.[212,216,218,219] This loss of a sense of self necessitated a renegotiation[212] or

reclaiming[213] of self, based on self-acceptance, self-knowledge and understanding,[212,213,218,219] and relationships with reliable others.[219] Spirituality also had a role in achieving a sense of self.[206]

Self-efficacy
This concept was expressed in the reviewed studies primarily as a lack of self-confidence, but also as feelings of inadequacy, uselessness, failure, an inability to cope and helplessness.[Chambers 2012,206,211–214,216,218] Mental health problems were associated with a lack of confidence.[Chambers 2012,207,212,213,216] This lack of confidence limited day-to-day functioning and activities[Chambers 2012] and access to helpful resources,[Chambers 2012] and affected choice and opportunities in employment[Chambers 2012,206,207] and relationships.[Chambers 2012,206] However, bipolar disorder could be associated with an increase in self-confidence.[207]

Self-esteem
The theme of self-esteem included the concepts of self-image, worth, value and shame, and a view of the self as 'defective'.[Chambers 2012,206,211,216–218,220] It was closely associated with loss of self-identity[207,218] and confidence.[216] Occupational activity was considered particularly important for self-esteem and status,[206,207,217] as was the satisfaction gained from helping others.[206] However, the difficulties encountered in obtaining employment often resulted in a lowering of self-esteem.[211,220]

Self-acceptance
Self-esteem was usually reported as a sense of loss of self-worth and value. A related concept was that of self-acceptance,[206,213,218,220] acceptance of the self as a person with an illness[213] or the belief that the illness did not represent everything that they were.[206,218]

Self-stigma
The theme of 'the self' was closely related to the next theme of 'belonging', particularly through the concepts of 'stigma' and 'feeling normal'. For example, feeling normal had two inter-related dimensions: not feeling normal within oneself,[212,215,216] and not being treated as normal and not fitting in with the rest of society.[215] This inter-relationship is evident in the concept of self-stigmatisation, an internalisation of the negative views of others, perhaps the most potent of stigmas as it has a profound effect on the inner sense of self.[206]

> I stigmatize myself. I just have a very low self-image. I'm kind of hard on myself for not conducting myself the way I should be . . . not being as productive as I could be. It's a reflection from general community's perceptions of what this illness is all about.
>
> *Corring[206]*

Belonging
The concept of belonging has been defined as the experience of integration and personal involvement in a system or environment at differing interpersonal levels. It can have two dimensions: 'valued involvement' – the experience of feeling valued, needed and accepted; and 'fit' – the person's perception that his or her characteristics articulate with, or complement, the system or environment.[201] Of the primary research studies included in the review, one identified 'connecting and belonging' as being important to quality of life.[215] Others identified closely related main themes: being part of a social context,[214] rejection and isolation from the community,[216] a need for acceptance by others,[220] social support,[218] relationships,[206] barriers placed on relationships,[211] labelling and attitudes from others,[211] stigma,[206,211,218,220] alienation[220] and detachment and isolation.[211]

Relationships
Relationships were clearly central to the concept of 'belonging'. These relationships included both close connections with family and friends and more casual relations with the local community, in the workplace, with service providers or with society at large. The complex nature of relationships and the positive and/or negative effects on quality of life were evident in all the primary studies.

The provision of support was a particularly strong theme, in three studies.[214,218,219] Both practical[Chambers 2012,206,213,214,218,219] and emotional[Chambers 2012,206,213–215,218,219] care and support from family and friends[Chambers 2012,206,212,214,215,218–220] and peers and work colleagues[207,216,219] were identified as important to quality of life, as was support received from professionals.[Chambers 2012,206,213,214,216,219] When families and professionals were unsupportive, quality of life was described as declining.[Chambers 2012,206,219]

> If I didn't have my parents and my family and friends . . . I wouldn't be going nowhere . . . I'd be stuck.
>
> *Gould* et al.[212]

Good and reliable relationships also satisfied the need for love, care and affection,[Chambers 2012,206,214,215,218,220] facilitated the experience of joy, fun and happiness,[210,214] and offered company, shared interests and camaraderie,[206,210–212,214,215,217] as well as someone to talk to/share problems with.[Chambers 2012,206,210,214,219] Such relationships also enabled individuals to feel needed/helpful to others,[206,210,214,219] to be accepted and understood,[Chambers 2012,206,207,214–216,218–220] and to feel that there were people in whom they had trust and confidence[Chambers 2012,214,219] and who provided motivation and encouragement.[214] Given their importance, the well-being of others was also important to the quality of life of the study participants.[214]

Although relationships which satisfied the need to belong were important, difficulties forming and maintaining these relationships were evident.[Chambers 2012,206,210,211,215–217,220] These difficulties included problems and tensions within supportive long-term relationships.[Chambers 2012,206,210,216,218] Problems with relationships represented a complex multidirectional interaction between the person and society at varying interpersonal levels. This interaction involved the effect of the person's illness when relating to others, other people's subsequent reactions and attitudes to them, and the effect of those reactions and attitudes in further exacerbating symptoms of anxiety and depression and affecting the person's perception of themselves. Examples of the barriers experienced in connecting and relating to people included cognitive and thought disorders resulting in problems with concentration and attention,[206,211,220] problems controlling behaviour[211,216,218] including acting out,[211,218] irritability, volatile or inappropriate behaviour,[218] or grandiosity or self-inflation,[218] and feelings of anxiety when talking to or being around people, including problems with trust and paranoia.[Chambers 2012,206,207]

Stigma

Stigma has been defined as 'any condition, attribute, trait or behaviour that symbolically identifies the bearer as culturally unacceptable or inferior'.[221] Stigmatisation was a major theme in four of the studies[206,211,218,220] and evident in three others.[Chambers 2012,215,216] The experience and perception of negative reactions on the part of family, friends, service providers, employers and society at large was shown to have a detrimental effect on quality of life. Individuals felt that they were perceived as lesser human beings who were discriminated against and treated accordingly[206] and that they were feared, avoided, or not accepted, which in turn led to feelings of rejection, marginalisation or being written off.[206,211,216,218,220] As a result, disclosure of mental illness was problematic and often avoided, and this had consequences for employment and close relationships.[206,215,218] Stigma had a detrimental effect on most aspects of life, including relationships,[Chambers 2012,206,211] employment and career,[Chambers 2012,206,211,218] going out and pursuing leisure activities,[Chambers 2012,211] obtaining services[206] and planning for the future.[206] Stigma was considered to be more predominant in bipolar than unipolar depression.[218]

Feeling normal

A major barrier to achieving a sense of belonging was that informants were not perceived by others – and often did not perceive themselves – as 'normal'.[215] Whereas feeling normal was something they held in high regard, instead they were aware of being perceived differently and, consequently, treated differently.[220] Feelings that they were different, and attempts to appear normal, do normal things, or be accepted as normal, formed a theme that permeated many of the studies reviewed,[206,211,212,215,216,220] being a major theme of three.[206,212,215] This is consistent with the dimension of 'fit' within the concept of

'belonging' – the person's perception that his or her characteristics articulate with, or complement, the system or environment.[201]

Loneliness/isolation/alienation

Feelings of isolation, loneliness, and particularly the concept of alienation can be regarded as the antithesis of a sense of belonging. Although highlighted as a main theme in one study only,[220] these feelings were evident within the themes of relationships and stigmatisation in all studies except one.[212] The symptoms of mental illness, the barriers these caused in the formation of relationships, and the stigma and consequential effects on the self, together with feelings of being different and not accepted, resulted in a pervasive sense of loneliness and isolation. People chose isolation, or avoided relationships, as a way of protecting themselves against rejection and dealing with the fears of how they appeared and what others thought of them.[206,211,220] The effects of being consistently treated as undesirable or different became internalised and further influenced their sense of self.[206] Isolation was further compounded by the feeling that they were the only person suffering in this way.[206] Hence, isolation was not just feeling as though they did not have any friends but became a painful feeling of despair that affected all aspects of life.

> *I think one of the things about schizophrenia, I don't know whether it's schizophrenia or whether it's, it happens in other mental illnesses too, is this terrible, terrible kind of inner isolation feeling, like you're the only person . . . who is going through what you are going through and you, and you're completely alone . . . it's just a terrible, painful sense of utter loneliness and isolation[.]*
>
> Reproduced with permission from Corring DJ. 'Being Normal': Quality of Life Domains for Persons with a Mental Illness. Dissertation Abstracts International Section B: Sciences and Engineering 2006;**67**.[206]

For quality of life, people wanted a reciprocal relationship with others[214,218] which involved understanding and acceptance.[Chambers 2012,206,214–216,218,220] This could be achieved through 'supportive own', those who share their illness and experiences,[206,207,214,215,218,219] or through belonging to a religious community.[Chambers 2012] However, it was also possible to have a sense of belonging to a social network that was ultimately not beneficial to quality of life,[218] and difficulties disentangling 'real' spiritual experience from hyper-religiosity when hypomanic could make belonging to a religious community problematic.[218]

Activity

By 'doing', a person achieves a sense of self and mastery, and successfully participates in the external world.[202,222] The importance of activity in some form to quality of life was expressed in all of the studies except the one which examined panic disorder from a nursing perspective.[213] There was a difference in emphasis between studies: some focused specifically on the benefit of employment[207,211,216,220] and others on activity or occupation in its broader sense, including both employment and leisure activity.[Chambers 2012,206,212,214,215,217,219] Whatever the type of activity, it was stressed that it should be meaningful or fulfilling,[Chambers 2012,206,212,214,215,217,219] enjoyable,[212,217,219] and suited to need and capabilities.[212,214]

The benefit of activity is that it can provide the means for many of the factors important to quality of life discussed above. It is through activity that the opportunity arises to interact with others and hence develop a sense of belonging.[215] Activity can also improve mood,[Chambers 2012,206,212,214,215] increase energy and/or motivation,[206,215,219] relieve stress[Chambers 2012] and boredom[215,217] and provide a distraction from problems.[Chambers 2012,214,215] It also helps self-esteem and self-confidence, engenders a positive self-identity[Chambers 2012,206,207,211,215,217,220] and enables people to take control of their lives.[215]

One further factor not already discussed is how activity provides order, routine, and structure.[211,214,215,218,219] Routine and structure can be achieved through employment,[211,218] childcare[211,218] or activity in general, be it work or leisure.[214,215,218,219] However, one study highlighted how too much structure could be problematic and that what was important was flexibility and choice.[218] Having a physiological routine – particularly regular sleep, meals and exercise – was considered important for general

well-being.[218,219] Using activity to structure time increased motivation, provided a diversion from problems and avoidance of negative moods:[215]

> *The actual work, whatever it is, is good for the mind and soul . . . you forget yourself. You forget your own problems when you are working[.]*

> *In the morning I have to do something. Some job or something I should do. Otherwise, I become bored and then become depressed because I don't have anything to do . . .*
> *Laliberte-Rudman* et al.[215]

Although activity was almost universally considered to be beneficial, taking part could be difficult if the activity was too demanding and not suited to needs.[207,214,219] The symptoms of mental illness could make difficult even the most rudimentary of activities, such as self-care, cooking and shopping,[Chambers 2012,217] and taking up employment was especially problematic.[212] Even potentially enjoyable leisure activities were avoided because of concern regarding other people's reactions,[Chambers 2012,211,217] problems relating to people,[217] and the associated fear and stress resulting in a deterioration in health.[Chambers 2012,211] Lack of money also put a restriction on enjoyable pastimes.[211] For those in employment, interpersonal relationships at work were particularly affected due to social withdrawal and irritability, or interfering, inappropriate or volatile behaviour during hypomania, although work productivity could increase during hypomania.[207]

Hope and hopelessness

The importance to quality of life of personal achievement or having dreams and goals was evident in six of the studies;[206,210,212,214,215,218] the importance of activity and/or life in general being fulfilling and having some meaning and purpose was also evident.[206,214,215,218,220] Both were necessary to instigate change, make plans, and to move forward. Again, the difficulty of achieving this was stressed.[206,212,215]

Losses experienced in the past affected the view of the future with a perception of reduced opportunities and choices[216] and diminished hopes and dreams,[210,212] particularly in the fields of employment[207,211,220] and relationships.[210,220] Loss and the effect of past experiences was a theme in eight of the studies,[Chambers 2012,206,207,210–212,216,220] and a major theme in three of these.[210,212,216] These losses included the loss of life roles generally and, more specifically, the loss of work and career opportunities, relationships and the parental role, skills and ability, time, financial losses, and, ultimately, the loss of a sense of self and identity. Losses which had occurred in the past were perceived as a burden,[206] with a pervasive sense of 'something missing'[220] which had long-lasting effects and made life a constant struggle.[210,216,220] Participants compared their own lives negatively with those of others,[210,216,220] or with their own lives before illness struck,[210,212] and all of this brought about feelings of failure and of being cheated, and a sense of unfairness.[207,216,220]

Past losses, including the loss of meaning and purpose in life, and a sense of helplessness and inability to cope, all brought about a feeling of hopelessness, necessitating a renegotiation and a lowering of aspirations and priorities.[206,210,212] The concepts of 'hope' and 'hopelessness' permeated the review studies[206,210,211,216,219,220] and formed a major theme for two[210,216] Hopelessness was an expression of the view that life would never change for the better, and brought about a pervasive feeling of distress.[220] Conversely, hope provided a catalyst for change and a better life.[219]

> *Well, my whole life feels problematic, I feel as if I'm not going anywhere . . . I know it sounds negative and I'm not really negative like this all the time, but you know, I find it hard, projecting myself into the future, and leading a happy life . . .*

> *I don't have hope that I'll ever have a nice boyfriend, I don't have any hope that I'll get married, I don't have any hope that I'll work a full week – week after week after week. I don't really have hope for stability . . .*
> *Lim* et al.[216]

Discussion

We identified six major themes associated with quality of life for those with mental health problems from this review of the literature: well-being and ill-being; control, autonomy and choice; self-perception; belonging; activity; and hope and hopelessness.

Measuring quality of life for people with mental health problems has been of growing interest because of concerns about the emphasis placed by mental health services on reducing symptoms. This has also been true of physical health services. Yet our review identified the importance of distress and symptom control from the perspective of people with mental health problems. The distress caused by the symptoms is integral to the quality of life of people with mental health problems, and in some instances seeing beyond this distress is difficult.

One of the strongest themes revealed by the review was a sense of belonging, achieved principally by good quality relationships and lack of stigma. It has been stated that people are fundamentally motivated by a need to belong,[223] and that belonging is the missing conceptual link in understanding mental health and mental illness.[224] Our review also indicates that negative social relationships are detrimental to quality of life. This is supported by research that shows that whereas a large social network and satisfaction with social relations are associated with a better quality of life,[225] negative social interactions and stigma are related to a worse quality of life.[226] Social exchange theory emphasises that social interaction entails both rewards and costs, and that negative social outcomes can have a greater impact on well-being than positive outcomes.[227] There is also evidence that loneliness is caused more by a lack of intimate connections than by a lack of social contact.[228] Hence, the important factor is the sense of belonging, rather than social contact. So, although there is a strong argument that those people who experience supportive, caring, loving relationships and have a sense of belonging have a better quality of life, it is less clear which is the more detrimental – to experience and risk the negative impact of uncaring and disrespectful relationships, rejection and stigma, or to protect oneself through self-isolation.

As good and poor relationships can have a positive or negative impact, so activity can both help and hinder quality of life. For some, the severity of symptoms can mean that basic self-care and day-to-day functioning are difficult. Activity beyond perceived capabilities can also result in feelings of anxiety, which in turn can lead to deterioration in other mental health symptoms such as hearing voices and paranoia. This results in avoidance of any potentially stressful situations. This finding is supported by the findings of research into the occupational activity of those with severe mental illness, which indicated that, though employment was valued, people made choices constrained by fear of relapse, and entered, avoided and shaped their social and occupational activity to remain well.[229] It was found that doing too much could exacerbate symptoms, yet doing too little could also cause illness and, therefore, people with severe mental illness sought out daily occupations with structure, flexibility and easily-met demands over which they had control.[229] Therefore, to achieve well-being and quality of life, people need to find a balance and be enabled towards that to which they are best fitted.[202]

Although avoidance of social and occupational activity may reduce anxiety and the occurrence of other related symptoms, at the same time it can compromise other aspects of quality of life. The consequent reduction in choice and opportunity has a detrimental effect on self-esteem and confidence. However, self-worth is gained through positive social feedback and successfully engaging in activity. Lack of self-esteem has also been shown to increase the risk of psychiatric disorders, the development of delusions and the maintenance of psychotic symptoms.[230] The perception of self is therefore both a cause and a consequence of mental health, and can be regarded as being pivotal to quality of life.

In relation to the finding of the importance of hope and hopelessness to quality of life, parallels can be seen between the results of this review and the concept of demoralisation,[205,231] whereby a persistent inability to cope with internally or externally induced stresses results in feelings of helplessness, incompetence, and loss of mastery and control, leading to diminished self-esteem, hopelessness and

demoralisation. This in turn adds to the distress of symptoms and further reduces a person's capacity to cope. The demoralised person clings to a small number of habitual activities, avoids novelty and challenge, and fears making long-term plans.[205,231] This feeling of demoralisation further impacts upon ill-being and, if untreated, leads to chronic distress and possible suicide.[205,232]

Strengths and limitations

Complexities arose when setting boundaries around the concept of quality of life. It was evident that there was a considerable overlap in findings, with studies examining 'recovery', 'lived/subjective experience', 'psychosocial issues', 'health needs' and 'strategies for living'. After much discussion and deliberation, these studies were excluded from the review. Since completion of our analysis, a systematic review of the concept of 'personal recovery' has been undertaken,[233] a concept previously defined as 'a way of living a satisfying, hopeful, and contributing life even with limitations caused by illness'.[234] Interestingly, the authors identified five recovery processes comprising 'connectedness', 'hope and optimism about the future', 'identity', 'meaning in life' and 'empowerment', which are very similar to our own final themes. They do not include 'well-being' and this may be due to the rejection of an emphasis on symptoms within the recovery movement and hence a separation of 'clinical' recovery and 'personal' recovery. This suggests that the concepts of 'personal recovery' and 'quality of life' are very closely related. This is important to understand as the concept of 'recovery' is gaining prominence as a guiding principle for mental health services.[235]

There were also difficulties setting boundaries around themes because of the strong inter-relationship between the different domains which make up quality of life. To avoid repetition, subthemes have been placed within the main theme with which they were considered to be most strongly associated, but aspects of these themes could be placed in other themes. For example, 'feeling normal' has been included under the main theme of 'belonging' but could also be regarded as an element of 'ill-being/well-being' and 'the self'.

Reliance on secondary data can limit the ability to make connections between subthemes, as can a reliance on that which researchers choose to report within a paper from their primary data. This is another reason for conducting our own primary research reported in the next chapter.

The primary studies included those with severe mental health problems only, with a majority covering schizophrenia or psychotic disorders. Where there was a mixed population, studies rarely indicated any differences between people with different diagnoses. The findings may therefore have biases towards those with psychotic rather than affective disorders. The interviews reported in the next chapter recruited more people with affective disorders to counteract this bias.

Implications for the wider project

This review has identified six potential domains of quality of life. These formed the basis of a topic guide for the interviews of people with mental health problems reported in the next chapter. They also provided an initial set of themes for the framework used to analyse the data collected from the interviews.

Measuring quality of life for people with mental health problems has been of growing interest because of concerns about the emphasis placed by mental health services on reducing symptoms. Yet this review has highlighted that the distress caused by symptoms is integral to the quality of life of people with mental health problems, and in some instances seeing beyond this distress is difficult. However, the review offers a partial picture because of the focus of studies on people with psychotic rather than affective disorders. The qualitative interview study reported in the next chapter addresses this limitation by including people with affective disorders. There is also concern that the review may be partial due to reliance on secondary data. The qualitative interview study in the next chapter addresses this possible limitation by gathering primary data on quality of life from the same samples as those included in the review.

Chapter 6 A qualitative analysis of interviews with mental health service users

This chapter builds on the systematic review of qualitative research into the meaning of quality of life for people with mental health problems reported in *Chapter 5*. The review identified six major themes: well-being and ill-being; control, autonomy and choice; self-perception; belonging; activity; and hope and hopelessness. One limitation of the review was that available studies focused on the quality of life of people with severe and enduring mental health problems, particularly schizophrenia. To complement the review we undertook primary research with participants from mental health services serving those with severe and enduring mental health problems and mild to moderate common mental health problems. This allowed us to explore the extent to which the review addressed important aspects of quality of life for those with severe mental health problems, given that most concerns have been expressed about the utility of preference-based measures in this group, and also to explore the views of people with less severe problems, which represent a gap in the current evidence base. This chapter presents the additional insights from the interviews and presents an assessment of the content validity of the EQ-5D and SF-36 against the themes identified. This assessment of content validity is extended to a new generic measure, the ICECAP-A, as this was developed from interviews with members of the general public.

Method

We undertook a qualitative study of face-to-face semistructured interviews with current users of mental health services.

Recruitment

Participants were recruited from three mental health service providers in a city in the north of England. One service (IAPT) provided psychological therapies for those suffering from mild to moderate depression and anxiety. The other two were community mental health teams (CMHTs) providing services for those with more severe problems, one working with individuals with severe and complex non-psychotic disorders (e.g. severe depression, PTSD, personality disorder) and the other with those with psychotic disorders (e.g. schizophrenia, bipolar disorder). Recruitment was undertaken by service providers who applied wide inclusion criteria in order to capture as broad a range of mental health problems as possible. Exclusions included those individuals facing acute episodes of their mental health condition, those not well enough to take part, and those who could not speak English or give consent, as well as cases where there was a known recent forensic history or any doubt relating to the safety of the researcher undertaking the interview in the person's own home. After a brief explanation of the research by the mental health provider, those persons who expressed an interest in taking part were given a recruitment pack consisting of a letter from the university-based researcher and a flyer giving concise information about the research, plus an 'expression of interest' form to be returned to the researcher in an included SAE. The researcher subsequently contacted the person to give further details, answer any questions and arrange the interview if the person was still happy to go ahead (none refused at this stage). A confirmation letter was sent, together with detailed information about the research (including confidentiality, anonymity and right to withdraw) and a consent form for completion prior to the interview. A choice of venue was offered, which could be the person's own home, a room at a university in the north of England or a centre frequently used by mental health services in an easily accessible part of the city. A £10 shopping voucher was offered for participation in the research. Approval for the above procedures was given by the local Research Ethics Committee.

The services recruited 21 people to take part in the research and 17 were subsequently interviewed, nine from the IAPT service and eight from the CMHTs. Four potential participants were recruited by the service providers but did not attend interview, two could not be contacted and did not respond to messages left

on their telephones (one IAPT, one CMHT), one arranged an interview but cancelled because they felt too ill to attend (CMHT) and another did not attend an arranged interview (CMHT). A further two participants with psychotic disorders were recruited by one of the participants subsequent to her own interview.

Interviews

All 19 participants were interviewed by a member of the project team (JC), a mental health researcher with a background in behavioural sciences and outcome measure development. The interviews were semistructured with the use of a topic guide (see *Appendix 5*) to ensure that a common set of questions was asked. The topic guide was based on the synthesis of a qualitative review presented in *Chapter 2* and, as the ultimate aim of the research was to establish the suitability or otherwise of the most commonly used generic quality of life measures, it included the item content of the EQ-5D and the SF-36. Three pilot interviews had taken place with people with different diagnoses (schizophrenia, bipolar disorder and depression) to test the topic guide. The guide was not changed substantially apart from a decision being made to include questions relating to the items of the EQ-5D and SF-36 as part of the interview rather than administering the actual measures. The first part of the interview aimed to elicit what was important to quality of life from the perspective of the individual, without any prompts. The interview started by asking for background information, and information on current activities and the participant's health, using responsive questions to probe these aspects of his or her life and to determine their relative importance. Participants were asked general open-ended questions about what affected their quality of life both from a positive and negative perspective, what they enjoyed and why, what they would most like to change, what helped, and what was stopping them doing what they wanted to do. Once their own perceptions had been exhausted, the interviewer introduced concepts that the research review had raised or that were included in the EQ-5D or SF-36. The interviews took place when the researchers were part way through analysing the data from the review, when the themes were descriptive in nature rather than conceptual. These were raised only if they had not already been discussed in the interview, and included questions about the relative importance or effect on their quality of life of relationships, support, stigma, work, leisure activities, mental health symptoms and their relative effects, medication and side effects, physical health/pain, energy/motivation, self-esteem/confidence, mental health services/workers, and finances.

All the interviews were tape recorded apart from one, where notes were taken at the request of the interviewee. One further interview was recorded but accidently deleted; notes for this interview were made 3 days after the interview took place. The interviews lasted between 25 minutes and 1 hour 50 minutes, averaging 1 hour 16 minutes. Seven of the interviews took place in the participant's own home. These tended to be people with more severe problems, five of whom were recruited by the CMHTs. The remaining two, from the IAPT service, had difficulties leaving their homes because of anxiety. The remainder of the interviews were conducted in a room at the university.

Participants

Interviewees included 11 men and eight women. These comprised one participant in his 20s, six in their 30s, eight in their 40s, three in their 50s and one in his 60s. One participant was of African Asian descent and another African Caribbean; the remainder were white European. A broad range of mental health problems were represented including schizophrenia, schizoaffective disorder, personality disorder, PTSD, mild to severe depression, anxiety, agoraphobia, eating disorder and anger. All, with the exception of one, were not in paid employment at the time of the interview although most had worked at some time in the past or were currently employed in a voluntary capacity. Five were married or with a partner; the remainder were not currently in a relationship. Further contextual information about the participants can be found in *Table 24*.

TABLE 24 Interviewee characteristics

Participant	Gender	Age range	Relationship status	Employment	Service from which participant was recruited	Problem/diagnosis from information given by participant during interview	Context
1	F	40–50	Married	House person	CMHT1	Depression/eating disorder	Long-term problems – abused as a child
2	M	20–30	Married	Returning to employment	IAPT	Anxiety	Traumatic event
3	M	40–50	Separated	Benefits Work history	IAPT	Depression/anger	Relationship break-up – long-term anger-related problems
4	M	40–50	Single	Benefits Work history	CMHT1	Depression/anxiety	Long-term problems – abused as a child
5[a]	F	50–60	Married	Unemployed Work history	CMHT1	Depression/anxiety	Financial debt – previous history of depression
6	M	60–70	Single	Benefits Voluntary work Work history	Other[b]	Schizophrenia/depression	Psychosis-related problems controlled by medication and own coping strategies – some symptoms remain in milder form. Previously depressed – 'solved' by long-term medication
7	M	40–50	Married	Benefits Work history	CMHT1	Depression	Multiple stressors – previous history of depression
8	F	40–50	Widowed	Benefits Work history	CMHT1	PTSD/depression/anxiety/agoraphobia	Severe traumatic event, bereavement, abused as a child
9	F	50–60	Divorced	Benefits Work history	IAPT	Depression	Physical health problems and relationship break-up
10	M	40–50	Divorced	Benefits Work history	IAPT	Anxiety/agoraphobia	Stress-related panic attack
11	F	30–40	Separated	Benefits Work history	IAPT	Depression/anxiety	Work-related stress and trauma, redundancy
12	F	30–40	Single	Benefits Work history	IAPT	Depression	Gender reassignment

continued

TABLE 24 Interviewee characteristics (*continued*)

Participant	Gender	Age range	Relationship status	Employment	Service from which participant was recruited	Problem/diagnosis from information given by participant during interview	Context
13	F	30–40	Single	Benefits Voluntary work No work history	CMHT1	Depression/personality disorder/social anxiety	Long-term problems – no attributable cause disclosed
14	M	30–40	Single	Benefits No work history	CMHT2	Schizoaffective disorder	Depression with periodic psychotic episodes – recovering from a recent severe episode. Previous diagnosis of bipolar disorder and schizophrenia
15	M	50–60	Single	Benefits Voluntary work	IAPT	Depression	Long-term difficulties exacerbated by bereavement and housing problems
16	M	30–40	Single	Benefits No work history	CMHT2	Schizophrenia/depression	Psychosis-related problems not controlled by medication. Depression due to schizophrenia illness and current situation
17	M	50–60	Single	Benefits Work history	Other[b]	Schizophrenia/mild depression	Episodic when stressed or does not take medication – no current psychosis-related symptoms
18	M	60–70	Married	Retired Work history	IAPT	Depression	Long-term grief due to bereavement – consequential financial and relationship problems
19[a]	M	40–50	Separated	Returned to work	IAPT	Depression	Relationship break-up

F, female; M, male.
a Interview notes used in analysis.
b Recruited via participant.

Analysis

The interview recordings were transcribed verbatim. Notes were used for two interviews. The interview data were analysed thematically using framework analysis.[198] Framework analysis was used to allow the identification of common and variable patterns of themes within and across different groups, in particular relating to problem severity and type. This analytical method also allowed for comparisons to be made with thematic findings from the review of qualitative research previously undertaken. The themes identified in the review made up the initial themes of the framework. Next, each transcript was read for refamiliarisation. Interview text was then charted onto the framework for the most part using the predefined themes of the review. This was either charted verbatim or paraphrased depending upon the relevance and focus of the response to the research question. Nothing was considered irrelevant and all data were coded in some form. All charted text was indexed with the interview timings that had been noted in the transcript so that actual dialogue could be returned to ensure contextual accuracy. It was not unusual for text to be coded under more than one theme, often one descriptive and another conceptual. Cross-references between themes were noted.

Findings

Subjective well-being and ill-being

Within the review (see *Chapters 5* and Connell *et al.*[199]) the focus of subjective well-being was on the negative rather than the positive aspects of well-being; that is, what takes quality away from life rather than what adds quality to life. Thus the focus was on overall feelings of distress related to the symptoms of mental illness. These included the feelings of depression, a lack of energy, motivation and enthusiasm, fear and anxiety (particularly in social situations, and fear of relapse) and the distressing symptoms of psychosis including hearing voices, hallucinations and paranoia. These aspects of ill-being were also described in our interviews, particularly by those with severe and enduring mental health problems. Thus, when asked what would improve their lives, participants tended to cite the absence of these negative issues rather than the presence of the positive aspects of subjective well-being such as feeling happy. When asked what made them happy the response was often 'nothing'. The expectation of ever feeling happy for this population was low.

> *I'm wanting to change – the biggest issue with my life, but there is a waiting list. There's a new therapy coming out called EMDR [eye movement desensitisation and reprocessing]. I got put down for it once but I was having suicidal thoughts, erm I've not attempted this year so I have been accepted . . . I will still get flash backs but I won't get the trauma afterwards and that's what I want to get rid of, I don't want to be traumatised every day, my coping strategies don't work for me . . . I would just want limited life back, instead of the full quality of life[.]*
>
> *Participant 8 (CMHT, PTSD)*

> Interviewer: *Can you think of anything that might make you happy, or happier?*
>
> Interviewee: *No! Perhaps if I was a different person that would make me happy, I don't know what else to say . . .*
>
> *Participant 4 (CMHT, anxiety/depression)*

In contrast, interviewees with less severe mental health problems did speak of the experience of happiness and enjoyment, although it was often tempered. They spoke of enjoying current hobbies and pastimes, or happiness as a feeling they wished to regain.

> Interviewer: *What makes you happy?*
>
> Interviewee: *Making sure that my wife is comfortable and happy, my son, my garden is just – simple things really, there is no major thing that I can think of . . . my grand-daughter makes me happy[.]*
>
> *Participant 18 (IAPT, depression)*

There were aspects of well-being that were identified in the review, but were much more prominent in the interviews. The first was the desire to feel calm, relaxed and peaceful, which was described across all mental health problem types and levels of severity. For one interviewee the need to feel calm and relaxed was in contrast to feelings of anger, which was an emotion affecting quality of life that was not included in the review.

> I can remember when I used to be relaxed and chilled and not feeling like I do now, but I've felt like this for such a long time and anxious for such a long time, but I can remember what it was like not to feel like this, it's a marvellous feeling[.]
>
> *Participant 10 (IAPT, anxiety/panic attacks)*

> . . . it's hard but I have just got to lose this anger, if not I am just not going to get any better, I'm just going to carry on and I am going to end up, like I said before, I will end up either crippled, dead or locked up and I don't want that, I just want to be able do what I want to do in the daytime and get home at night and just relax[.]
>
> *Participant 3 (IAPT, depression/anger)*

Related to a desire to feel calm and relaxed was a need to feel safe. However, the strategies used to ensure these feelings of safety could inhibit other aspects of quality of life.

> Where you are familiar with, where you feel alright like in [name of student area] full of students, they're harmless, as long as you don't bother them, that's it, they think you are one of them because you're young as well . . . everything matters, whether the people are safe whether the area is nice area, whether you know it properly, everything matters, yeah, every little bit contributes[.]
>
> *Participant 16 (CMHT, schizophrenia)*

> It is the safety issue that is most important, I have to keep myself safe . . . I can't like go out now in case I bump into somebody in [name of district] they go home and they say 'oh I met a young lady called so and so, in a different area' and they know me and my life might not be worth living . . . I am now trying to get out more, even just walk round the shops or summat, but it is hard[.]
>
> *Participant 8 (CMHT, PTSD)*

Lack of concentration was a negative aspect of subjective well-being that was more prominent in the interviews. Those with severe problems, or experiencing feelings of depression at their worst, reported that they were no longer able to enjoy even the simplest of activities such as reading or watching TV because of difficulties concentrating. This also meant that they were unable to find a distraction from their problems. Those with less severe problems reported a perception that their concentration was getting worse, but not to the extent that it took away from enjoyment of hobbies or pastimes.

> I used to play golf quite a lot, but I've not played for quite a while . . . I just don't get the same enjoyment out of it like I used to do. It used to be a nice walk and concentrating on – I think that's the other thing, I can't concentrate on it anymore. It takes a lot of concentration, does golf . . . I can't hit t'ball like I used to, just because I've got depression, it sounds weird, that, but I can't play like I used to do[.]
>
> *Participant 7 (CMHT, depression)*

Missing from the review was any indication of the aspects of subjective well-being affecting quality of life the most. Those who gave an indication at interview of having comorbid problems were asked this question and the overwhelming response was 'the depression'. When asked why, they stated that it was due to its all-consuming nature over which they felt they had no control. They could obtain some relief from anxiety or the voices, but depression was an ever-present darkness that they found much more difficult to cope with.

The depression is worse than the voices, because the voices sometimes, when you are with people, like now with you, I'm not hearing them, if you're doing serious matters, but I don't know how long that be, with me not hearing voices, but the depression is always there . . . it's like the English tradition, you heat milk – in the oven, in a pan, or in the microwave, and when you heat it properly you get something covering the upper layer of the milk, that's the depression, covering all the good moods and everything, and interests . . . and it covers all your body, it affects your motivation, your energy, and you don't want to do nothing and you end up sad . . . so like it covers up all your happiness and everything, to me, in my opinion[.]

Participant 16 (CMHT, schizophrenia)

. . . well both things stop me doing things but the depression is a lot harder to cope with, the anxiety you can take medication so that helps erm it might get severe like if I have panic attacks but panic attacks only last for a few minutes whereas depression just seems to go on and on and feels a lot harder to cope with[.]

Participant 13 (CMHT, depression)

Physical health

Physical health was not evident as a theme within many of the studies in the review. In contrast, the majority of our interviewees reported a physical health problem and this was mentioned by interviewees with all diagnoses and severities of problems covered in the sample. Physical and mental health interacted in various and complex ways.

Some reported that their mental health problems, or medication, caused or exacerbated their physical health difficulties.

It feels physical as well as mental . . . my body aches and like I think I just become really tense and that is what makes my body ache and I feel like erm I feel like my chest is being crushed and erm I can't breathe and things like that and erm I just want to be asleep all the time to escape but I can't sleep[.]

Participant 13 (CMHT, depression/affective personality disorder)

For others it was their physical health problems that contributed to deterioration in their mental health.

I've got diabetes, I've got high blood pressure, I've got arthritis, I've got angina, I'm on too many medications. I don't know what they are all for and it's just this last 6 months it has just been getting me down[.]

Participant 3 (IAPT, depression)

Participants were asked whether their physical or mental health problems affected their quality of life the most. Some found this difficult to answer, or described how the two interacted with each other. The presence of both seemed to make life particularly difficult to cope with.

. . . the two now are sort of like together on a par, I can't cope with both of them at all . . . it is bad enough trying to cope with what's going off in my head, but with my body in pain every day I can't even cope, I can't cope with that as well[.]

Participant 12 (IAPT, depression)

Self-perception

The concepts of self-efficacy, self-identity, self-stigma and self-esteem identified within the review were all revealed as having an effect on many aspects of our interviewees' lives. This was often communicated after prompting rather than being something immediately obvious to them and explicitly stated. Consistent with the review was an acknowledgement that a lack of self-worth and confidence was preventing them from doing things that might potentially improve their quality of life, yet it was difficult to overcome.

. . . it is the area of, um, you know, not willing to take, to step out and get out of my comfort zone is what I'm the most challenged on. So I do generally feel, er, inadequate in a way because I don't, I won't, should I say, I ought to, sort of, you know, take control, risk, if you like, and actually get on with it rather than, you know . . . like I say, well in terms of finding a partner, it's a confidence thing, you know, I see too much danger and, sort of, think perhaps that if you find, it would be difficult to find somebody who genuinely loved me for who I am [laughs], so, um, and I know it's not just in – it's in other life choices as well and er – but I'm not saying be reckless and gamble, recklessly do things but, on the other hand, um, sometimes chance favours those who will be prepared to take a chance . . .

Participant 17 (other, schizophrenia)

A negative perception of the self that affected quality of life was expressed by interviewees with all diagnoses and severities of illness. However, those who appeared most distressed at the time of the interview reported how a lack of self-esteem and confidence currently affected their lives in a detrimental way. Those with less severe problems talked about how their confidence was improving or that they were trying to address this and consequently improve their lives.

Yes it has improved a lot, if you had asked me to come here 7 week ago there was no way I would be able to come but now I have come on the bus. I feel a lot better in myself, if you had asked me 7 weeks ago I would have had to say come to my house, there was no way I could have come out. It was the thought of going out in the open I know it is horrible saying it but it knocked me for six kind of thing.

Participant 2 (IAPT, anxiety)

A further difference identified within our interviews was that for some a lack of confidence and self-esteem was something they felt had affected them all their lives and had contributed to their mental health problems, whereas for others this was a consequence of their illness.

I have got no confidence to do the things I used to do and self-esteem, no, just lack it. You see before I was a right confident person, until what happened to me happened, so I don't know whether my self-confidence will come back, 'cos before I was very confident, I had to do hand overs at work and like talk to GPs and things like that, and it didn't bother me, and now I don't even like talking to people some days [laughs], I don't know what it is, I just put it down to my illness[.]

Participant 8 (CMHT, PTSD)

The review identified how those with bipolar disorder, schizophrenia and panic disorder in particular described difficulties with having a coherent sense of self. This was also evident in our interviews; however, one interviewee diagnosed with schizoaffective disorder described how this could have a positive as well as a negative effect.

. . . it's almost like this different personality of voices bouncing round your head, kind of, you know what you'd call your train of thought, your own, kind of, inner dialogue, that kind of, it's like you've got that, but yet there's other ones trying to, kind of, take over and become that train of thought and they've got their own voices and their own personalities and their own characteristics and you're like struggling to, kind of, stop them hijacking your brain, in a way, and so you kind of, you're having conversations in your head with them and, kind of, talk, and sometimes it's really funny and good, it's like having two mates round and you're just arguing and talking and at other times, you kind of, you lose track of who you are, if you know what I mean[.]

Participant 14 (CMHT, schizoaffective disorder)

Autonomy, control and choice

Consistent with the review, our interviews revealed the importance to quality of life of the related concepts of autonomy, control and choice. The review highlighted the complex juxtaposition between support and independence: the dilemma between needing support and valuing independence while not wishing to be too dependent. This same quandary was communicated by our interviewees. They spoke of an aspiration for independence but with an acknowledgement by some that support, and sometimes dependence, were necessary during periods of illness when they were not in a position to help themselves. The need and desire for support and independence therefore changed over time.

> *. . . when you're depressed, you've got, kind of, you're very reluctant or unable to help yourself, and even if there's loads on offer, you won't find it, you kind of, you need to keep in touch with people, like, so me seeing my family or as many friends as I can and like the social worker, is vital, otherwise you could, you know, get really really down[.]*
>
> *Participant 14 (CMHT, schizoaffective disorder)*

In our interviews, the value of support for quality of life was expressed primarily by those with severe mental health problems, whereas those with less severe problems were more likely to be actively working towards greater independence. Of our interviewees, the person for whom independence appeared to be of the greatest importance and dependency the most troublesome had physical as well as mental health problems which affected their mobility. Being independent was important for dignity, pride and privacy, whereas being dependent resulted in feelings of guilt and being a burden which could put a strain on a good relationship.

> Interviewee: *Just to sort of do something on my own without like oh well I've got to get out or whatever, will you come to bus stop with me, or got to ring a taxi, just to be able to say 'right I'm off' not disappear but just say 'right, I'll see yer' and just go without having 'where you going, what time you coming back' . . . I am dependent on people.*
>
> Interviewer: *How does that make you feel?*
>
> Interviewee: *Fed up sometimes, I mean don't get me wrong, I would go out with my daughter or my friend but that would be by choice. You know like 'oh I am meeting' her rather than [daughter saying] 'oh well I've got to meet her cos I've got to go shopping'. That sounds awful, but that's the – do you know what I mean. You know without her thinking 'oh I can't do this cos I've got to make sure me mam's all right'[.]*
>
> *Participant 9 (IAPT, depression)*

Also consistent with the review was a need for control, particularly symptom control, for those with psychosis and anxiety-related problems. For these interviewees a fear of losing control was evident. However, this prevented them from doing some of the things they would like and increased dependency upon others. This showed that although control was desired, and increased feelings of well-being, too much control could have a detrimental effect on quality of life. Our interviewees were very much aware of this but found it difficult to change and find the balance between the two. A consequence of this was that they felt bad about themselves because they were not living up to their own and others' expectations.

> *I seem a bit of a control freak, I want everything to be worked out before I decide to do a certain thing, you know, I want everything to be fairly straightforward and I mean, you can't, in a way, you can't live life like that, and yet I still want to live life like that, do you know what I mean? . . . it's about the stress, erm having faith or taking this, stepping out of your comfort zone, whatever you want to call it, yeah[.]*
>
> *Participant 17 (other, schizophrenia)*

Having choices and opportunities was also important to quality of life. As with the review, choice was particularly associated with having sufficient finances, which in turn was often linked with the availability of suitable employment. However, the opportunity of employment, or at least employment that was not demeaning to their values and expectations, was something they felt was denied, often due to perceived and experienced discrimination. Not having money meant they did not have the opportunity to pursue those things that could improve their lives in ways that were related to leisure activities, their environment or their physical and mental health care. Our interviews also showed how having choices and opportunities was cumulative, with one leading to another.

> I ended up going to the halfway house [gives name] at halfway homes they had lots of opportunities . . . they were brilliant because they there was a pot of money every month that they could spend I mean I learnt to drive through that . . . it would encourage you and enable you to do these life things so I did driving, somebody else did fishing . . . I don't think they have a training budget now which is a bit of shame because it opened my life to lots of things. I went on courses . . . I got the money for the driving, knowing that I could drive triggered off other things for me and I did get benefits at that point which allowed me a little bit more to do these things . . . So that's how my life got built by having guidance and opportunities really . . . and I think that's a big thing isn't it, there's not all the opportunities there and stuff[.]
>
> *Participant 1 (CMHT, depression/eating disorder)*

Because it was difficult to find suitable work, and because of the stresses that employment entailed, some participants chose the safety of remaining on benefits. Some preferred to do voluntary work instead, which provided them with the benefits linked to employment, apart from financial, but with more flexibility and control to protect themselves from becoming ill.

Relationships and belonging

Within the review, 'belonging' was defined as having two dimensions: 'valued involvement', the experience of feeling valued, needed and accepted, and 'fit', the person's perception that his or her characteristics complemented the environment they lived in.[201] The focus was on how well people perceived they fitted with society. The importance of a sense of belonging was also present in our interviews. In contrast to the review, our interviewees also felt that society did not fit with them. This was expressed as a dilemma about whether or not they wished to be part of a society that they felt had different values to their own. This sense of being 'alien' to society was something expressed by those with severe depression and schizotypal disorders. Those with mild to moderate anxiety/depression expressed feelings of being isolated from society but not that they did not belong to it.

> I have feelings of err not belonging to the human race, like I feel very – it's not an outcast, I just don't feel a connection. I don't know how else to describe that, it's like being an alien, that is the only way I can describe it, and I know that that sounds weird but that is the only way that I can describe the feeling of it . . . I don't cope with most people [sighs] I – [pause] my values and norms are very important to me and I know that everybody has not got the same ones[.]
>
> *Participant 4 (CMHT, depression/anxiety)*

Similar to the review, there was an expression of a desire to do 'normal' things or to return to a normal self, particularly for those who had experienced a traumatic event. However, in contrast to the review, our interviewees communicated a sense of being different, rather than abnormal.

*You know, you will not make any sense to 'em with voices or whatever you want to call them, bouncing around your head, because then that fires my, you know, it fires my imagination, but I'm talking to people that aren't, kind of, a bit oddball or a bit, kind of, you know, a lot of people are really, kind of, straight and they're not strange enough, you know, they don't – I look at things and, kind of, from a certain angle, a lot of people don't, and likewise for them, they'll just think I'm f*****g mad or, you know, dark or depressing, which I can be, but, it's like gallows humour, you know[.]*

Participant 14 (CMHT, schizoaffective disorder)

Stigma was a strong theme in the review and though present in our interviews, was disclosed primarily by those with problems related to gender identity and schizophrenia, particularly related to employment opportunities. Stigmatisation was less of a problem for those with anxiety and depression.

Rather, a need for a better understanding of the impact of mental health problems was expressed by interviewees with all diagnoses and severities of illness. This lack of understanding resulted in them not receiving the support they needed and a feeling of not being able to talk to people about their problems.

Interviewer: *Thinking of quality of life – what would have improved things for you at the time.*

Interviewee: *Well, more understanding from people in authority, like the council, by helping people that have got problems that need the help . . . I find a lot of people don't understand depression and they don't understand the way it affects people and, you know, I've had many people say to me, 'oh you're stupid, you just pull yourself together', but they don't understand the way it affects people, you know.*

Participant 15 (IAPT, depression)

The review identified the importance of good quality relationships. These satisfied a need for support, love, care and affection, and a need to have someone they could trust and confide in, who accepts and understands them, and with whom they could experience fun and happiness. These themes were also identified in our interviews. Similarly, our interviews revealed that, particularly for those with severe and enduring mental health problems, there were difficulties forming and maintaining close, family and social relationships. The relationships that could fulfil these desired needs felt out of reach, which had a reciprocal effect on subjective well-being and self-perception. These interviewees tended to live alone with care and support provided by their mental health worker only.

I alienate myself because I don't share. I have to take responsibility for that, but I can't – without sharing they don't understand why I don't share, it's a bit of a vicious circle. But to have that trust to share, is quite difficult for me, but, of course it is, it is not going to be easy, it takes a long time for me to be friends or something with anybody really . . . [name] from [name] she helps me a lot she is from an organisation to support people with mental health issues to do everyday things to encourage you to go out and things like that, and without [name] I think I would go insane. I only see her once a week or once a fortnight but she does understand me[.]

Participant 4 (CMHT, depression/anxiety)

In contrast, interviewees who tended to have less severe or relatively short-term problems had difficulties associated with current relationships that affected their quality of life, but these difficulties were less profound. They did not express any difficulties in their abilities to form new relationships and spoke of the support received from other family and friends.

Not evident in the review was the way in which people made negative and positive comparisons with others in order to evaluate their own quality of life. In our interviews, envious comparisons were made with people in the interviewee's immediate family or social circle. On the other hand, positive comparisons tended to be made with remote referents, people who were less fortunate than themselves with whom

they might rarely come into contact. For some interviewees there was an internal struggle between the two.

> *I think 'oh everybody's at me and they're all paired off and I haven't got my mother and dad and everybody's, all the world's got a mother and dad' even though they haven't and then I say 'well my sister hasn't got a mother and dad' she's only same as me but she's got a good husband, and he is a good husband. And then I say I've got a sister, I've got a brother-in-law as well . . . and I thought you know 'pull your bloody self together, at least you've got legs that work to moan at, there's some people sat in a wheelchair' . . . and so I know I'm lucky it's just, [sighs] oh I don't know[.]*
>
> *Participant 9 (IAPT, depression)*

Activity and functioning

The findings related to activity from both the review and our interviews were very similar – that leisure and work activity had a positive effect on subjective and psychological well-being, was enjoyable, engendered a sense of belonging through social interaction, and generated feelings of self-worth, pride and a sense of achievement. Importantly, it also provided the income often needed to take part in activities that are pleasurable and aid recovery.

> *I went on a year's course at engineering and I was absolutely scared about going on that, but I did it and I did it, you know, quite well . . . and at that time, I was really happy in my life and I thought well I've done something, I've achieved something here doing this . . . I think [I was happy] because I had the drive and a purpose of getting up and going out every morning and doing what, you know, normal people do, sort of thing, you know, I got into a routine which was very good, so I was happy and more stable . . . I just felt a little bit more worthwhile, you know. I don't like to feel useless . . . I've missed the banter I think as well, as part of being a team, or even on your own or whatever, I miss the camaraderie[.]*
>
> *Participant 10 (IAPT, anxiety/panic attacks)*

Again, similar to the review, our interviews also revealed that when no enjoyment could be found from activity, it still had the benefit of providing a structure and routine to the day, relieved boredom, kept the mind active and – as expressed by many as being of particular importance – provided a distraction from problems.

> *[After feeling suicidal] my interest in things started coming back it has made a lot of difference because it meant that I was able to distract myself to a certain extent . . . so I was grateful that that came back a bit, rather than just being completely immersed in this world of depression, at least I got a few minutes out of, you know, it was just a few minutes at a time, but it gradually built up to the longer periods where I could go back on the computer again or I could say watch EastEnders, whereas I just completely lost interest, so that was one thing that improved[.]*
>
> *Participant 13 (CMHT, depression)*

Also consistent with the review was the finding that in certain circumstances activity could be detrimental to quality of life. This occurred when the activity, usually employment, felt beyond their capabilities. Reasons given for this were stress and anxiety related to pressure of work, problems concentrating, difficulties with social interaction and the exacerbation of other mental health symptoms.

> *. . . they systematically bullied me, unbeknown to me, I didn't even realise because I was that stressed and sort of, out of it, my work life balance doesn't work very well, I don't tend to put myself first, and therefore I get iller and iller whilst enabling other people at the end of it I went off sick because I wasn't sleeping at all, I developed symptoms of erm – it's what I called jelly brain, and it's when you turn your head it feels like your brain is slowly turning, even though your head has already turned, it is a very strange sensation, and it's a sleep deprivation sign. Erm, but obviously I was still awake because of all the anxiety, it was the anxiety that usually kept me going.*
>
> *Participant 4 (CMHT, depression)*

Our interviews highlighted an additional issue around the dilemma people faced when choosing between short- and long-term quality of life: they could foresee the potential long-term benefits of taking part or resuming activities they previously enjoyed, but often found the prospect of doing so too daunting in the short term. The fear that their mental health and well-being might deteriorate further seemed to take priority over the potential for more long-term life benefits – 'I'd like to but I can't'.

> . . . what the psychiatrists advocate is a relatively low stress environment, that's how I keep my mental health on an even – you know, my mental health healthy as it were . . . part of me thinks that perhaps if I could do a job where even if it was just cleaning, and if I could, in a way, confront stress to a certain extent, be more, you know . . . I think to a certain extent, it would probably be a bit more healthy to have a bit more in my life, if you know what I mean[.]
>
> *Participant 1 (other, schizophrenia)*

Our interviews also revealed some differences associated with the severity of mental health problems. Those with severe problems and who appeared most distressed at the time of interview reported finding little enjoyment in any activity, often due to an inability to concentrate. Difficulties with basic functioning and lack of self-care were reported as demonstrating a lack of interest in life itself, as was taking part in particularly risky activity, which was communicated by one participant as an illustration that she did not care whether she lived or died.

> I don't look after myself, I'm not interested in myself . . . I'm not! I'm just not interested in myself I don't care if I am not here today gone tomorrow, I am not bothered[.]
>
> *Participant 8 (CMHT, PTSD)*

In contrast, those with less severe problems spoke of leisure activities and hobbies they enjoyed when asked what improved their quality of life. These were often lone activities as, although social activities were described as potentially the most enjoyable, those activities that involved social contact appeared to be the first to diminish as problems worsened.

Hope and hopelessness

The review showed how important a positive view of the future was to quality of life. This involved having goals and aspirations, and being involved in activities that were fulfilling and had meaning and purpose. These were necessary in order to instigate change and have hope for a better future. This was also evident in our interviews, with some further insight into how difficult this was for some, particularly those with severe and enduring mental health problems.

> I think whereas I was feeling completely hopeless that things would never change and I would be stuck like this for ever, and just the desperation for things to end or things to change, she just gave me that little bit of hope that things would change . . . eventually I came round to their way of thinking and it gave me hope that at least I could hang on for just a bit longer to see what happens and to see if things would change[.]
>
> *Participant 13 (CMHT, depression/affective personality disorder)*

There were noticeable differences between our interviewees in how they viewed the future. This could be related to their expected level of achievement before they became ill and whether or not they felt that was something they could return to, given time.

The most negative outlook and evidently poorest quality of life was reported by those for whom previous attempts at positive action to change their situation had failed, and coping mechanisms they previously

drew upon no longer had an effect. As a result, they could not conceive of their situation changing in the future.

> *. . . why carry on, what with my head f****d, and the pain, oh do you know, I just, I really have to ask myself, really why do I carry on . . . It is really difficult to justify yourself at night, at 3 o'clock in the morning, why are you putting up with all this, and why you just don't commit suicide to be honest because, I don't know, it's just really quite exhausting, I am going to the gym to help with my back and stuff like that, but it hasn't, but my blood pressure has come down. Hurrah! [cynically] after 3 bloody years[.]*
>
> *Participant 4 (CMHT, depression/anxiety)*

In the review, the emphasis was on the negative concept of hopelessness as related above. A more positive outlook tended to be expressed by our interviewees referred via the primary care services who had less severe and/or relatively short-lived mental health problems. In contrast to a perception that things were unlikely to change, they were more likely to talk about having goals and plans in place which they were actively working towards, however difficult. They were also more likely to compare themselves positively with others. Religion also provided solace for some. However, this optimism may in some respect have been due to these participants being recruited through a service specialising in cognitive behavioural therapy interventions which focus on goal setting and positive thinking techniques.

> *I think that I have turned the corner I've settled down even though I probably might be on my own now, but I'll have to wait and see about that . . . alright I have got problems, but there are people worse off than I am, they have got more problems than me and if some people can do it I think I should be able to do it myself[.]*
>
> *Participant 3 (IAPT, depression/anger)*

Content validity of EQ-5D, SF-36 and ICECAP-A

Content validity is the extent to which the items of the EQ-5D and SF-36 (and the SF-6D) reflect the content of the domains of interest identified by the qualitative research. This research has provided a depth of understanding that is rarely possible with conventional psychometric techniques. It has identified seven themes that were important to people with mental health problems. In this section we return to the main aim of this project, which is to examine the validity of the EQ-5D and SF-36. This is achieved by using the seven themes identified in the qualitative research to assess the extent to which these two measures comprehensively cover them.

A key criticism of the EQ-5D and SF-36 is that they have been designed by researchers with little or no input from people with the relevant health problems. As a result, it has been suggested that these measures are too focused on physical health-related problems and so exclude potentially important mental health-related quality of life.[11,12] Patient-reported measures are being increasingly developed from interviews with the relevant populations and this is in line with guidance in the academic literature[56] and some policy-makers.[58] A recent example of this is a new generic instrument called ICECAP-A[40] which, the authors claim, can measure capabilities. A key feature of this instrument is that the content is based on the results of a qualitative analysis of semistructured interviews with members of the general public. The resultant classification system is less focused on physical health and functioning, and instead takes a higher-level and broader view of the constituents of quality of life. It covers five dimensions: feeling settled and secure; love, friendship and support; being independent; achievement and progress; and enjoyment and pleasure. The ICECAP-A was not available at the development stage of this project and there are no relevant data for undertaking quantitative psychometric analysis for this group of people. However, it became apparent during the research that the ICECAP-A has similarities to the emergent themes and provides a broader descriptive system that is potentially more relevant to mental health populations.

It was therefore decided to include it in this assessment of content validity. The review of the content of these three instruments against the seven themes is summarised in *Tables 25–31*.

EQ-5D

The content of the EQ-5D was presented in *Table 1*. As shown in *Tables 25–31* there is only a modest amount of overlap between the EQ-5D and the seven themes.

TABLE 25 Summary of subjective well-being/ill-being and a content validation of the EQ-5D, SF-36 and ICECAP-A

Subtheme	Summary description	EQ-5D	SF-36	ICECAP-A
Distress – overall theme	A key dimension in QoL literature covering positive and negative aspects of well-being. People with severe mental health problems tend to focus on negative feelings and the symptoms of their condition. Symptoms are really a part of subjective well-being, particularly for depression and anxiety For those with the severest mental health problems, feelings of distress associated with depression, experience of psychosis and mania, and anxiety had a major impact on quality of life. The day-to-day experience of symptoms and coping with them was dominant. Described by one as 'wretched', 'debilitating', 'painful', 'tormenting' and having a 'tyrannical power of their life' (Connell *et al.*[199]). Often interactions between different types of stress – with one making the others worse			
Depressive mood	For those with comorbidities, this was often the most important symptom and was described in interviews as being all-consuming, constant (and least amenable to medication), outside of their control, and associated with poor concentration (on things they used to enjoy), low energy and poor motivation Positive well-being: happiness and enjoyment – though not common in severe mental health problems	Anxiety and depression (negative only)	Mental health (positive and negative)	
Fear and anxiety related	Fear or panic can be the primary condition and can be caused by social situations. Respondents sought to reduce anxiety by avoiding stressful situations. Those with mild to moderate depression tended to talk about worry Positive well-being: feeling peaceful, calm, relaxed and safe	Anxiety and depression (negative only)	Vitality	Enjoyment and pleasure (positive only)
Psychosis related	Distress caused by critical voices, like living a nightmare that can be difficult to differentiate from reality		Mental health (positive and negative)	Feeling settled and secure (positive only)

QoL, quality of life.
Note: some symptoms related to specific conditions, whereas depression and anxiety occur across conditions.

TABLE 26 Summary of activity and functioning and a content validation of the EQ-5D, SF-36 and ICECAP-A

Subtheme	Summary description	EQ-5D	SF-36	ICECAP-A
Positive	Activities like work, hobbies or social interaction can be associated with providing structure, energy, motivation and meaning. A person achieves a sense of self and control, develops a sense of belonging and overall improves well-being. A means of achieving other aspects of QoL	Usual activity Self-care	Physical functioning (includes self-care) Role limitations due to emotional problems	
Negative	Can be stressful if too demanding and beyond their capabilities. Fear of stress can result in avoiding activities that might be enjoyable and beneficial in the longer term People with severe mental health problems (like severe depression) may lack energy and motivation for even basic functioning such as self-care		Physical functioning (includes self-care)	

QoL, quality of life.

TABLE 27 Summary of social well-being – relationships/belonging and a content validation of the EQ-5D, SF-36 and ICECAP-A

Subtheme	Summary description	EQ-5D	SF-36	ICECAP-A
Relationships: close friends and family	These relationships include those with close ones, family, friends and care worker Positive consequence: provided practical and emotional support, satisfied need for love and affection, promoted confidence and facilitated enjoyment. Those with more serious problems did not feel close and supportive relationships were not possible for them Negative consequence: can be source of rejection and tension, and low self-esteem. Can be part of a multidirectional interaction, with the reactions and attitudes of others exacerbating the problem. So they may seek to avoid people to whom they were previously close	Usual activities – this is not but specific	Social functioning – but of relationships not quality	Love, friendship and support (positive only)
Social relationships	Participants talked about social life and sociability being important to improving their QoL and may be an alternative to close relationships. For some, withdrawal made life easier in short term. Work colleagues, volunteering or groups with similar problems can provide this Negative: not seen as normal by themselves or others. Feelings of loneliness/isolation and alienation	Usual activities	Social functioning – but of relationships not quality	Love, friendship and support

TABLE 27 Summary of social well-being – relationships/belonging and a content validation of the EQ-5D, SF-36 and ICECAP-A (*continued*)

Subtheme	Summary description	EQ-5D	SF-36	ICECAP-A
Reactions of others – understanding, acceptance and stigma	Being treated in a non-judgemental way. Understanding is an important theme in the interviews	None	None	None
	Negative reactions and lack of understanding, including rejection and stigmatisation, take QoL away. Stigma was reported mainly by people with schizophrenia (i.e. identifies a person as culturally unacceptable or inferior), and is associated with being avoided or not accepted (e.g. by family), resulting in feeling of rejection, being marginalised or written off. Less of a problem with depression			
Sense of belonging	This is feeling a part of society as a whole. Came out as important from the review and was an overall experience of integration, involvement and fitting in. It was found to be less important in the interviews			
	Some did not feel a part of society and felt rejected			

QoL, quality of life.

TABLE 28 Summary of self-perception and a content validation of the EQ-5D, SF-36 and ICECAP-A

Subtheme	Summary description	EQ-5D	SF-36	ICECAP-A
Self-identity	The review identified this as a problem of 'knowing who you are'. It appeared to be mainly related to bipolar disorder, schizophrenia and panic disorder. Required renegotiation or reclaiming of self			
Self-efficacy, self-esteem and self-acceptance	Terms used interchangeably to describe a lack of self-worth and confidence. The review also found feelings of uselessness and inability to cope			Achievement and progress ('I can achieve and progress in all aspects of my life') – seems to be related to it but only positive
	Participants in the interviews recognised it as having a detrimental impact on QoL. Common across conditions and impacts on social- and work-related activities. Those with milder problems saw it as an area they can improve over time			
Self-stigma	Poor self-image resulting from negative connotations associated with the illness. Impacts negatively on life due to the individual's own and other people's perceptions			

QoL: quality of life.

TABLE 29 Summary of control, autonomy and choice and a content validation of the EQ-5D, SF-36 and ICECAP-A

Subtheme	Summary description	EQ-5D	SF-36	ICECAP-A
Dependence and independence	Support was valued and seen as important for QoL. Indeed, those in more severe conditions aspired to more support and accepted they could not be independent			Being independent ('I am able to be independent . . .') – but does not deal with need for support
	However, support can make people feel dependent and this can be associated with loss of control and esteem. Concerns about loss of independence expressed by those with less severe conditions and/or physical comorbidity			
	In review, discussed importance of personal autonomy to find the right balance between support and independence. Autonomy does not imply independence			
Control	Self-control: mainly related to relief and management of symptoms, usually through medication. Anxiety and panic associated with fear of loss of control, and depression with lack of control			Feeling settled and secure
	Environment control: there was a desire to control the environment in order to feel safer. Can involve withdrawing into a smaller world – into a comfort zone			
Choice	Having choice and opportunity is important for QoL. Participants usually associate this with money and access to resources which enable someone to do the things they want to do (capabilities theory). To get money, need opportunities to do paid work; these are limited by perceived discrimination and stress from work (so some did voluntary work)			

QoL, quality of life.

TABLE 30 Summary of hope and hopelessness and a content validation of the EQ-5D, SF-36 and ICECAP-A

Summary description	EQ-5D	SF-36	ICECAP-A
Having dreams and goals, and being involved in activities that gave meaning and purpose in life, was necessary to have hope for the future		General health perception item on future health expectations	Achievement and progress ('I can achieve and progress in all aspects of my life')
Sense of hopelessness that life would never change for the better resulting in a lowering of aspirations			
Those with more positive outlook had goals and plans and were working towards them			

TABLE 31 Summary of physical health and a content validation of the EQ-5D, SF-36 and ICECAP-A

Summary description	EQ-5D	SF-36	ICECAP-A
This can be a physical comorbidity or an experience associated with mental health problem	Mobility, self-care, usual activities and pain or discomfort	Physical functioning, role limitation due to physical problems, vitality, bodily pain	

Well-being and ill-being

The anxiety and depression dimensions of the EQ-5D may reflect the important aspects of ill-being of depressive mood and fear and anxiety, and this is reflected in the psychometric evidence (see *Chapter 2*). However, the qualitative evidence would suggest that these two elements are different in terms of the impact they have on the lives of sufferers: namely, that depression was identified as being an order of magnitude worse due to its all-consuming and long-lasting nature. In contrast, experiences of fear and anxiety tended to be shorter term. It is therefore important to separate these into two dimensions. Depressive mood is associated with low energy, poor concentration and poor motivation, which are not really directly covered by this dimension. Symptoms of psychosis that have important and distinctive implications for quality of life are not measured. Finally, the EQ-5D is concerned with negative aspects of these themes and so excludes the positive counterparts such as happiness and enjoyment or feeling calm and peaceful.

Activity and functioning

Usual activities is a rather crude and generic dimension that covers aspects of activity in a limited way. People limited in their activities by their mental health condition may report a problem on this dimension. However, this does not allow for the negative side of activity highlighted as an issue in the qualitative work. People with mental health problems often find that activity can have a negative as well as a positive impact.

Social well-being – relationships and belonging

Usual activities may again be seen as a catch-all dimension that covers some aspects of this theme. However, it provides no information on the consequences. Relationships can be those with close friends and family through to social relationships with colleagues, volunteers and neighbours. They can be supportive and enjoyable, or they can be stressful or a source of rejection and stigmatisation. The concept of usual activities also fails to capture a sense of belonging to society.

Physical health

The EQ-5D has content to cover key aspects of physical functioning related to mobility, self-care, usual activities and pain. These four dimensions have been shown to be valid in many common physical health problems, but research has found that some consequences of physical ill health are not reflected in the EQ-5D, such as those resulting from poor vision.[236]

This leaves the themes of *control, autonomy and choice*, *self-perception* and *hope/hopelessness*, which are only addressed through mobility and self-care that indirectly contribute to a person's ability to be independent. However, this does not include the more psychological components of this theme.

SF-36

The SF-36 covers more emotional aspects of health than the EQ-5D by including mental health, role limitations due to emotional problems, vitality and aspects of general health (see *Table 2*). This is reflected in the results in *Tables 25–31*.

Subjective well-being/ill-being

The mental health dimension of the SF-36 has five items covering depression (e.g. feeling downhearted and low), anxiety (e.g. being a nervous person) and positive affect ('have you been a happy person' and 'feeling calm and peaceful'). The vitality dimension covers positive (e.g. 'feeling full of life') and negative

energy (e.g. 'did you feel worn out'). On the face of it, the content of this instrument would seem to cover most aspects of well-being and ill-being. One omission is concentration, though this is partly covered in role limitation due to emotional problems (i.e. 'did not do work or other activities as carefully as usual'). Another is psychosis-related symptoms (e.g. hearing critical voices), though the extent to which this should be in a measure of quality of life is debatable.

Activity and functioning
Role limitation due to emotional problems in the SF-36 would appear to cover this theme (i.e. 'cut down on the amount of time you spent on work or other activities', 'accomplished less than you would like', 'didn't do work or other activities as carefully as usual'). However, this dimension does not consider the levels of stress from activities found to be important to people with mental health problems.

Social well-being – relationships/belonging
There are two social functioning items in the SF-36: 'has your health limited your social activities (like visiting friends or close relatives)' and 'to what extent has your physical health or emotional problems interfered with your normal social activities with families, friends, neighbours or groups'. The focus on the way in which health impacts on social activities, through the general terms of 'limited' and 'interfered', should cover some aspects of this theme, but does not capture all issues discussed by participants, such as deep feelings of alienation or stigma, nor the responses of others. It also fails to incorporate the sense of belonging. In other words, this SF-36 dimension is less able to reflect the psychological consequences of social activities.

Hope and hopelessness
This theme overlaps with the depressive mood subtheme under the well-being/ill-being theme, and to that extent feelings of hopelessness may be partly covered by items like 'feeling down in the dumps all the time'. The notion of hopelessness is also partly reflected in the general health perception item of 'I expect my health to get worse'. However, the overall feeling of hope and having goals to aspire to is not captured by the SF-36.

The themes of *self-perception* and *control, autonomy and choice* are not covered by the content of the SF-36. As for the EQ-5D, the items in physical functioning will be related to independence.

SF-6D
The preference-based SF-6D is derived from the SF-36 and so shares much of the same content. However, there are some important differences owing to the fact that it contains only a subset of 11 of the 36 items (see *Table 3*). The SF-6D does not cover general health perception and this means that it does not have any item for tapping the *hope and hopelessness* theme. It combines role limitation due to emotional problems with that due to physical problems. In addition, the SF-6D does not cover positive affect or negative energy (i.e. tiredness).

ICECAP-A
ICECAP-A is a classification like the EQ-5D with a single item per dimension. The five dimensions are feeling settled and secure; love, friendship and support; being independent; achievement and progress; and enjoyment and pleasure. The content of this instrument would seem from the summary in *Tables 25–31* to be more closely aligned with the themes identified from qualitative work, and to some extent this is reassuring as they are based on a similar qualitative approach. The main difference is that the ICECAP-A was based on interviews with members of the general public rather than those with mental health problems.

Subjective well-being/ill-being
The main observation for this theme and the others is that the ICECAP-A only considers the positive aspects of the themes. Two dimensions are related to this theme, one for depressive mood (i.e. 'I can have XX of enjoyment and pleasure') and the other for fear and anxiety ('I am able to feel settled and secure in

all areas of my life'). Although it is important to cover positive aspects, the findings from qualitative research in people with mental health problems is that those with more severe problems do not talk about enjoyment and pleasure, but rather focus on the negative aspects of well-being. The best they hoped for was the absence of depression. Only in those with mild to moderate problems were these positive aspects raised in the interviews. Well-being is part of a spectrum and so to exclude ill-being is likely to result in floor effects (i.e. respondents clustering at the lowest level). ICECAP-A also does not assess the positive or negative components of energy.

Activity and functioning
None of the dimensions of ICECAP-A seem to cover this theme.

Social well-being – relationships/belonging
The dimension of love, friendship and support captures this theme better than either of the generic measures, the EQ-5D and SF-6D. However, again ICECAP-A fails to cover the negative aspects of relationships including rejection and stigmatisation and feelings of loneliness and alienation. For people with mental health problems, these consequences are an important part of their experience. It also does not capture the sense of belonging and being part of society.

Self-perception
ICECAP-A does not cover most of the issues raised under this theme, such as self-identity, self-acceptance and self-stigma. It may capture aspects of the positive side to self-efficacy and self-esteem in the achievement and progress dimension (i.e. 'I can achieve and progress in all aspects of my life'), though this dimension fits better with hope and hopelessness.

Control, autonomy and choice
This is another theme where ICECAP-A provides a better match than the generic measures of health through the dimension of 'being independent' ('I am able to be XXX independent'). The wording is based around the ability to be independent and so allows for respondents to say they could be independent but choose not to be. However, it does not include aspects of control nor choice more generally.

Hope and hopelessness
The notion of hope is captured by the dimension 'achievement and progress' ('I can achieve and progress in all aspects of my life'). Interviewees reported that having dreams and goals gave meaning and purpose in life. The worst level in this dimension is 'I cannot achieve and progress in any aspect of my life', but this may not capture the very low levels of hopelessness reported in the interviews, with some respondents indicating that life would never change for the better. To confirm this would require psychometric evidence to establish the extent of any floor effects and whether or not it is able to discriminate those with the most severe mental health problems.

Physical health
By way of contrast with the generic measures of health, ICECAP-A does not contain items that directly assess aspects of physical health. This was raised in the interviews by participants who reported mental health problems to be associated with physical manifestations, and there were interactions with physical comorbidities. It could be claimed that the impact of these physical health problems is captured by the higher-level domains of the ICECAP-A instrument. However, this requires further psychometric evidence.

Discussion

The primary research reported in this chapter builds upon a review of qualitative studies reported in *Chapter 5*. The aim of the research was to further examine the views of people with mental health problems on what is important to quality of life. We undertook a qualitative study of face-to-face semistructured interviews with current users of mental health services. Participants were primarily recruited

from three mental health services in the north of England. One (IAPT) provided psychological therapies for people suffering from mild to moderate depression and anxiety. The other two were CMHTs providing services for those with more severe problems, one working with individuals with severe and complex non-psychotic disorders (e.g. severe depression, PTSD, personality disorder) and the other with those with psychotic disorders (e.g. schizophrenia, bipolar disorder). This expanded the scope of diagnosis and severity of illness covered in the review to include people with affective and anxiety disorders as well as people with psychosis-related disorders. The seven themes to emerge from this qualitative work have been used to assess the content validity of the EQ-5D and SF-36. Below we summarise the findings from the interviews and compare and contrast them with those from the review. We address the limitations of this part of the study and summarise the implications for the generic measures examined in this report.

Summary of findings

Our review and synthesis of qualitative research studies undertaken with people with mental health problems, reported in the previous chapter, identified six domains of quality of life: well-being and ill-being; control, autonomy and choice; self-perception; belonging; activity; and hope and hopelessness. Despite widening the types and severities of mental health problems studied, we found our interview data fitted well with the themes from the review; any differences tended to be within the themes and related to the degree of impact of the themes on different levels of severity, chronicity and diagnosis. With some exceptions, those with severe chronic difficulties were more likely to talk about losses and what took quality away from life, whereas those with moderate or relatively short-lived problems spoke of the things that added quality to life. The EQ-5D and SF-36 were found to have only a modest overlap with the seven themes identified from the research, and though the ICECAP-A had better overlap, it only considers the positive end of the spectrum.

Importance of physical health

The interviews identified more of the impact of physical health conditions which interacted with mental health problems and made people's life situations much worse and difficult to cope with. This may have been less of an issue in the review as it excluded studies in which participants had major physical health comorbidities. As a result, physical health makes an important additional major theme, as opposed to a minor subtheme in the review. It also means that any measure of HRQoL for use in people with mental health problems should include domains on physical health.

The spectrum from ill-being to well-being

Consistent with the review, we found that the distress related to symptoms is integral to quality of life for those suffering from mental health problems. This indicates that symptomology and quality of life are not separate entities, and consideration should be given to the inclusion of symptoms in quality of life assessment. Our interviews revealed differences in the emphasis placed on distress between those with severe and enduring, and those with mild to moderate, problems. The former tended to be at the negative end of the well-being/ill-being spectrum, where quality of life was perceived as an absence of distressing feelings and emotions. This priority of a reduction in the symptoms of mental health is consistent with other studies[237] where the desired outcome of people with functional mental illness was to feel less anxious or depressed. In contrast, those with moderate and relatively short-lived problems did speak of happiness and enjoyment with a greater expectation of achieving this. This is consistent with evidence indicating that positive and negative affects play an independent role in health outcomes, and that quality-of-life assessment should be bipolar, consisting of the independent dimensions of negative and positive well-being.[238,239] The shift in focus of mental health services away from mental illness and towards mental health over recent years is commendable. However, people still enter these services with suicidal feelings, and a change from feeling suicidal to non-suicidal is an important change in quality of life that needs to be considered in outcome measurement.

Additionally, our interviewees said that depression had a greater impact on quality of life than other mental health problems. This suggests that anxiety and depression should be treated as separate items rather than combined as they are currently in the EQ-5D and SF-36. In their research, Headey et al.[240]

concluded that four dimensions of well-being should be distinguished: life satisfaction, positive affect, anxiety and depression. They found that life satisfaction was negatively correlated with the distress dimension of depression, but there was no association between life satisfaction and anxiety; people could be both satisfied with life and anxious.

Subjective versus psychological well-being

There are strong arguments that hedonic well-being in terms of pleasure, enjoyment and satisfaction is insufficient for a good quality of life.[29] Also crucial is human potential, or eudaemonic well-being,[241] and self-actualisation,[242] which considers how satisfied people are with their lives. This differentiation also has parallels in the study of subjective and psychological well-being.[243] The concept of psychological well-being extends subjective well-being beyond happiness and positive affect,[243,244] suggesting a multidimensional model of psychological well-being consisting of six dimensions: self-acceptance, environmental mastery, autonomy, positive relations with others, purpose in life and personal growth. The initial three concepts are not dissimilar to our themes of self-perception, autonomy, control and choice, relationships and belonging. Furthermore, meaning and purpose is one of the main components of our theme of activity. One difference is that this literature is primarily concerned with what makes a good life and what adds quality to life, whereas for many of our interviewees, quality of life was expressed in terms of what took quality away from life.

An easier, restricted life or a stressful, fuller life?

For those with more severe and enduring problems such as long-term depression, PTSD and schizophrenia, there was an acknowledgement that a lack of self-worth and confidence was preventing them from doing things that might improve their quality of life, and that this was difficult to overcome. For some, life was easier if they were able, or allowed, to stay within their own comfort zone. This might mean choosing to remain on state benefits rather than risk the stresses of employment, despite an acknowledgement of the benefits employment could bring to quality of life, such as self-esteem, social engagement and improved finances. This was one of the common threads throughout our research on the views of people with mental health problems, which we alluded to in the review and which became more explicit in the interviews. There was a conflict and dilemma between short-term well-being and other aspects of quality of life. Our interviewees expressed how they engaged in negative coping strategies, such as avoiding things that made them feel anxious, but at the same time these strategies stopped them doing those things that might ultimately enhance their quality of life. This is consistent with research indicating that certain aspects of positive functioning, such as the realisation of one's goals and purposes, require effort and discipline that may well be at odds with short-term happiness.[245] This seemed to be all the more so for those with serious mental health problems, for whom added stress could exacerbate or induce other mental health symptoms (e.g. hearing voices). It was difficult for some of our interviewees to decide which was more important to quality of life – a restricted life, free from anxiety and stress, or a fuller life which involved negotiating barriers, anxiety and stress and the risk of exacerbating other aspects of their mental illness.

Capabilities approach

Both the review and interviews revealed how a lack of choice and opportunity prevented people with mental health problems from reaching their potential, particularly in the field of employment. The concept of choice is an important aspect of the capabilities approach to quality of life. A person has a set of 'functionings' which refer to what a person can 'do' or can 'be', the relative importance of which are assessed by the person him or herself. Whether or not a person is capable and can benefit from these 'functionings' is determined by the choices and opportunities available to him or her. Many of our interviewees were well educated and once had aspirations of successful careers. However, opportunities were often denied, either due to the limitations imposed upon them as a result of their mental health difficulties or due to both actual and perceived discrimination, particularly in the workplace. So where some were once considering careers as bankers, lawyers or writers (as stated by our interviewees), they were now deliberating whether or not they would be capable of a full-time job stacking shelves or cleaning. Not being able to live up to their own and others' expectations could have a cumulative effect

on subjective and psychological well-being, which in turn made it harder to cope and change their situation. Added to this is a lack of understanding of their problems, both by the people close to them and society at large, and a consequential lack of support and other resources. This was a problem for some of our interviewees more than others, and particularly for those with psychosis-related disorders and severe depression. Those with less severe problems had current difficulties fulfilling their potential but tended to feel it would be something they could potentially regain once they recovered.

The importance of relationships

Understanding and support is an important component of good-quality relationships. The importance of good relationships, and also the effects of poor relationships, was a theme that ran throughout our interviews and was expressed by people with different types and severities of illness. Those with severe and enduring problems were more likely to speak of the damage that had been a result of poor relationships in the past and which, in turn, affected their potential for fulfilling relationships in the future. Avoidance of damaging relationships, usually within the family, improved quality of life. Those with less severe problems had difficulties with specific relationships but those problems were less profound, and they were more likely to speak of the benefits of the positive relationships in their lives. This was reflected in a sense of being alienated from society for those with severe difficulties, which was not expressed by those invited to our research via the primary care service, perhaps because it was something that was taken for granted. The importance of relationships is reflected in the inclusion of this theme in the majority of quality-of-life models, including needs based, psychological, social health and cohesion, environmental and objective social indicators.[239]

Is quality of life different from recovery?

As already described (see *Chapter 5*, *Discussion*), our domains closely match those of a recent review of the 'personal recovery' literature which was released after we had completed the analysis of the review studies.[233] Personal recovery could therefore be regarded as a further model of quality of life in addition to those outlined in *Chapter 1*. Importantly, both our research and that into personal recovery are based on the voices of people with mental health problems.

How does this help us to understand the quantitative findings?

The psychometric evidence presented a mixed picture, with the EQ-5D and SF-36 apparently achieving a satisfactory level of performance in depression, and to a lesser extent in anxiety and personality disorder. However, the evidence was clear for schizophrenia and bipolar disorder, and there was a suggestion that, even in these conditions, these crude quantitative tests were actually picking up differences in depression. The range of themes identified in the synthesis reported in the last chapter and the interviews reported here provides some insight into the problem. Essentially, both generic measures, and particularly the EQ-5D, are focused on physical health and functioning and seem to miss large areas of mental health-related quality of life mentioned by the interviewees. What the participants spoke about was closer to the quality-of-life models described in *Chapter 1* than the narrow dimensions that form the content of these generic measures. The extent of this mismatch of content has been detailed above. This would seem to account for the mixed results, as large elements of the lives of these people that are affected by their mental health problems are not assessed.

Strengths and limitations

One of the primary aims of the interview research was to ascertain what was important to quality of life for people with a wider range of mental health disorders and levels of severity than was available within the review. To a large extent we achieved this, with interviews being undertaken with people recruited from primary and secondary services catering to the needs of people with severe and enduring problems and mild to moderate anxiety and depression. However, despite attempts, we did not receive any referrals from these services of women with psychosis-related disorders, or people with obsessive compulsive disorder. Our findings may not, therefore, be inclusive of their views. There was also an over-representation of referrals from one member of the CMHT who was a nurse. Although we are unaware of any biases that may have been introduced as a result of this, it is possible that the perspective of those being seen by a

nurse in the team could be different from those whose key worker was a psychologist, occupational therapist, etc. Finally, we had hoped that practitioners would introduce the possibility of taking part in the research to all their current clients and would let them decide whether they felt able, or wished, to be interviewed. However, the number of referrals received strongly suggests that service providers acted as gatekeepers and the voices of all types of service users were not necessarily included. The implication is that further interview work could be undertaken to address these gaps.

Conclusions

The review and synthesis of qualitative research undertaken with people with mental health problems, and the interviews reported in this chapter, have together identified seven domains of quality of life: well-being and ill-being; control, autonomy and choice; self-perception; relationships and belonging; activity; hope and hopelessness; and physical health. It is important that these domains are addressed in some way in quality-of-life measures. People enter mental health services with varying degrees of distress and chronicity, and as our research shows, recovery can be a long and difficult process. However, people may perceive large improvements in their quality of life without registering on the positive end of any quality-of-life scale, although undoubtedly this should be the aim. It is therefore necessary that the full spectrum of negative through to positive aspects of each domain is included within quality-of-life measures. The EQ-5D and the SF-36 and its derivatives were found not to cover large elements of the seven themes identified by the research, and this calls into question the content validity of these measures for large parts of the population with mental health problems.

Chapter 7 Discussion and conclusion

We have presented the results of a detailed examination of the appropriateness of the EQ-5D and the SF-36 (and its derivatives) in people with mental health problems. We have used mixed methods to evaluate the key properties of validity and responsiveness. Studies employing quantitative methods, including systematic reviews of the literature and psychometric analyses of existing data sets, have been used to examine the construct validity (testing for known-group validity and convergent validity) and the responsiveness to change in mental health status of these measures. Qualitative evidence on content validity has been obtained from a systematic review of the literature and analyses of interviews with people with a range of mental health problems. This chapter begins by presenting a brief overview of the results of each study. It then discusses the main findings of this research and presents the implications for policy and recommendations for further research.

Summary of main studies

Psychometric evidence

Systematic reviews were undertaken of the psychometric literature in five mental health conditions: depression, anxiety, bipolar disorder, personality disorder and schizophrenia. Overall, the results from 91 studies identified by an exhaustive search of the literature were used to assess the construct validity and responsiveness of the EQ-5D and SF-36.

Generic measures were found to adequately reflect differences between groups or changes over time in populations diagnosed with depression. In populations with anxiety, the evidence was less clear as the differences between known groups may have been driven by comorbid depression rather than anxiety disorders themselves. For personality disorder, most studies supported the construct validity of the EQ-5D, but there was insufficient evidence on the SF-36. Within schizophrenia the evidence demonstrated known-group differences, but this was mostly limited to differences between individuals diagnosed with schizophrenia and the general population. Contradictory evidence was found in studies using clinical measures of symptom severity, where the generic measures failed to reflect differences or correlate with the clinical indicators. In bipolar disorder, generic measures reflected known differences in clinical measures of depression but not mania.

The amount of evidence found in the literature was limited in terms of size and coverage, so it was decided to expand the evidence base by undertaking further psychometric analysis of a number of existing data sets. This provided more evidence in depression, anxiety, personality disorder and schizophrenia and used more patient-based assessments in the validation. These analyses broadly supported the findings of the reviews, with the EQ-5D and the SF-6D being found to be valid in samples with mild and moderate depression and anxiety. For schizophrenia, the findings were less clear, with the EQ-5D and SF-6D not being responsive to change in comparison with the condition-specific measures.

The tests of construct validity and responsiveness tend to be rather crude, as they depend on the validity of the construct used to assess the criteria. Tests of known-group validity, for example, usually either rely on comparison with the general population or use clinical assessments of severity that may not reflect meaningful differences in quality of life from the perspective of the population with mental health problems. Such evidence cannot prove or disprove the validity of a measure; at best it can provide support for appropriateness of the measure. The findings from our review and the further analyses seem to highlight potential concerns in anxiety, bipolar disorder and schizophrenia (particularly with regard to responsiveness to treatments). It is important to judge these findings alongside qualitative evidence on the content validity of the measure.

The psychometric evidence presented an interesting and mixed picture in terms of the performance of the EQ-5D and SF-36 in populations with mental health problems. However, this quantitative evidence is not able to offer an explanation for these mixed results. For this reason, qualitative research was undertaken on the impact of mental health problems on quality of life.

Qualitative evidence

The review of previous research found 13 studies that had interviewed people with mental health problems about the way their condition impacts on their lives. It was difficult to be sure that all informative studies had been identified given differences in terminology. The findings from those studies located were synthesised using the 'framework' approach.[198] We identified six domains from this review: well-being and ill-being; control, autonomy and choice; self-perception; belonging; activity; and hope and hopelessness.

The complementary primary research involved semistructured interviews of people with mental health problems. These interviews expanded the scope of diagnosis and severity of illness covered by the studies in the review to include people with affective and anxiety disorders referred via NHS services, and those with severe and enduring mental health problems and mild to moderate depression/anxiety. We found that our interview data corresponded well with the themes from the review, and any differences tended to be within the themes and related to the degree of impact of the different levels of severity, chronicity and diagnosis. The only change to the themes was that physical health was found to be more important among the interviewees than suggested from the review, so this was added as a seventh theme. The review and interview data found that each theme had a spectrum of positive and negative components.

Mapping between mental health and generic scales

When generic measures are not being used, one solution is to estimate mapping functions where condition-specific measures are regressed onto generic measures to produce health-state values. We estimated mapping functions between widely used condition-specific measures (HADS, CORE-OM, PHQ-9, GAD-7, and GHQ-12) for depression and anxiety and the EQ-5D and SF-6D, using four data sets available to the investigators. Mapping functions were not estimated for other conditions as the psychometric evidence suggested that the generic measures were not appropriate.

The statistical models mapping the HADS onto the EQ-5D had poor predictive performance. The mapping functions for the SF-6D fitted the data sets better, particularly the CORE-OM onto the SF-6D, but they still suffered from some degree of over- and underprediction towards the ends of the ranges. Given the psychometric evidence that the generic measures performed satisfactorily in this group, this result is perhaps surprising. These analyses were limited by the small size of some of the data sets and the condition-specific measures available, with the EQ-5D only being mapped onto the HADS. Furthermore, most of the mental health measures only focus on mental health, and often on very narrow aspects of symptoms of mental health, which may not translate into a general health score. We conclude that these mapping functions should not be used, but that the original generic measures should be used in trials and other clinical applications in order to obtain accurate estimates.

Implications of findings

The content of the EQ-5D and the SF-36 was compared with the seven themes identified in the qualitative research in order to provide an assessment of content validity. In summary, the EQ-5D would seem to cover little of the content of these seven themes owing to its focus on physical health. Only physical health seems to be covered, along with activities and functioning, which is included in a rather crude way through usual activities, and ill-being in terms of depression and anxiety. The SF-36 covers more by having a multi-item dimension on mental health and a vitality dimension that covers more aspects of well-being and ill-being, and some aspects of social functioning. However, like the EQ-5D it fails to include the response of participants and others to activities and social well-being; self-perception and control; autonomy and choice; and hope and hopelessness. It was decided to extend this analysis to a new

measure, the ICECAP-A, as this was derived from interview data, although it was obtained from the general public and not people with mental health problems. The ICECAP-A has greater overlap with the themes, but it is limited to the positive manifestations of these themes, and this may be because people with mental health problems were not targeted in the interview sample used to develop the measure. It should also be emphasised that there is no psychometric evidence on the performance of ICECAP-A in a mental health population, and it may miss important changes, particularly in populations with more severe problems.

Care must be taken in drawing any firm conclusions about the generic preference-based measures of health reviewed in this report due to the following limitations:

- The quantitative psychometric evidence is limited to a comparatively small number of studies that in many cases have small sample sizes. Furthermore, the tests were reliant on measures of symptoms and clinical diagnosis which, although widely accepted in mental health research and validated in the populations concerned, are not measures of the construct of HRQoL.
- The population of people with mental health problems in the qualitative research was not comprehensive despite an extensive review of the literature and an attempt to recruit across the spectrum of conditions. An important problem is the risk of selection bias by professionals in assessing the suitability of service users for interview.
- There has not been time to undertake a full theoretical analysis of the qualitative data. Further work is required to fully understand the themes and consider them in relation to existing models of quality of life and related concepts.

A more general problem arises from the lack of an agreed definition for quality of life or mental health-related quality of life, and hence there is no gold standard for examining the validity of these or any other measures. This is a major handicap for the field. However, improving the mental health-related quality of life of people with mental health problems is a key objective, so it is incumbent on researchers to develop measures and assess their performance as best they can, and to use accepted tests such as known-group validity, convergent validity and content validity to provide support or otherwise for the measures they use. Policy-makers also need guidance on which measures to use for assessing the cost-effectiveness of interventions and monitoring the outcomes of services.

Despite these concerns, there are conclusions to be drawn about the implications of the findings.

- The combined evidence from the psychometric research and the qualitative research suggests that the generic measures of health do not capture many of the concerns of importance to people with mental health problems. While the ICECAP-A is better in covering many of the themes, it does not consider the negative end of the spectrum of these themes, more relevant to people with mental health problems.
- For depression and personality disorders, the generic measures of health would seem to be adequate in picking up differences between known groups, but there is less support for their responsiveness to changes over time. These also exclude key themes.
- For anxiety, bipolar disorder and schizophrenia, the generic measures fail to capture many of the problems that arose in the interviews and this is reflected in the psychometric evidence on validity and responsiveness.

Further research

This has been an extensive and rigorous testing of generic measures of health in mental health population through the application of quantitative and qualitative techniques. However, there are important gaps remaining in the literature and limitations in the research reported here that need to be addressed.

Further research is required to improve the robustness of the findings as follows:

- The quantitative analysis needs to be extended to include further data sets that can allow the further testing of the construct validity and responsiveness of the generic measures. The relevance of the tests would be improved by administering other measures of self-perceived HRQoL rather than relying on clinical indicators to examine construct validity and responsiveness.
- This report has examined tests drawn mainly from classical psychometric theory, but further insight could be gained from the application of modern methods using latent trait models, such as Rasch and item response theory, to examine how well the items used in the generic measures reflect the dimension in general. This would require the collection of generic measures alongside more specific measures of the dimension of interest, such as depression or anxiety, before pooling the items and running these models to examine item performance in terms of characteristics such as model fit, differential item functioning across groups, item severity and ordering of responses.[246]
- Further interviews need to be carried out in those conditions not well represented, such as obsessive compulsive disorder and alcohol and drug misuse. It is also important to recruit some respondents through different channels in order to avoid the risk of a selection bias caused by relying on professionals. This is difficult to achieve from an ethical viewpoint, but there may be scope through patient groups.

Research implications of the findings reported here include the following:

- The analysis of content validity should be extended to mental health-specific outcome measures. This is important for guiding the choice of mental health measures for use in research to measure mental health-related quality of life.
- Consideration should be given to the development of a preference-based measure for calculating QALYs that is more appropriate to mental health. This could be an enhanced version of an existing generic measure. A recent development has been the addition of extra dimensions or 'bolt-ons' in order to make the EQ-5D more relevant.[247] The problem for the EQ-5D and SF-6D is that adding more dimensions makes them difficult to value using one of the preference elicitation techniques (such as TTO). A more fruitful avenue might be to develop preference-based measures more specific to mental health either from existing measures[51,248–252] or by the development of new measures. Developing from an existing measure has the advantage of building on past work and can be applied to existing data sets containing the measure. A recent example of this has been the development of the CORE-6D preference-based measure from the mental health-specific CORE-OM.[52,253] A limitation in this case is that it is concerned with common mental health problems such as depression and anxiety rather than more complex conditions such as schizophrenia and personality disorder. The other option would be to develop a new preference-based measure for use in mental health populations. This could build on the findings from the review of qualitative evidence and the interviews presented in this report and elsewhere, but it would be a major research endeavour.

Conclusion

The results of this mixed-method study are that the generic EQ-5D and SF-36 seem to achieve an adequate level of performance against some psychometric tests and may be acceptable for use in depression, and to some extent in anxiety and personality disorder. However, there are concerns regarding the way depression and anxiety are combined into a single dimension in the EQ-5D and SF-36. Results from the psychometric analyses in schizophrenia and bipolar disorder have been more mixed. These measures provide only a limited coverage of the themes found in the qualitative research carried out with people with mental health problems, and so may present a misleading impression of the impact of these problems on the lives of those affected. This has implications for the validity of economic evaluation in mental health. Recommendations for further work include the development of a new preference-based measure in mental health based on the themes identified in thus research and existing measures.

Acknowledgements

The authors would like to thank all interview participants for agreeing to be interviewed, staff at the Sheffield Community Care Trust (particularly Tom Ricketts) for their support and help in recruitment into the study, and Eleni Chambers who provided advice on the design. Thanks are also extended to Professor Jan van Busschbach, Professor Simon Dixon, Professor Derek King, Dr Louise Longworth, and Associate Professor Jane Morrell for their support in accessing the data sets used in this study.

Contributions of authors

John Brazier was the principal investigator who led the design, conduct and writing up of the research.

Janice Connell undertook the review of the qualitative literature and conducted the interviews of mental health service users.

Diana Papaioannou undertook the reviews of the psychometric literature on personality disorder, schizophrenia and bipolar disorder.

Clara Mukuria undertook psychometric and mapping analyses.

Brendan Mulhern undertook the psychometric analyses.

Tessa Peasgood undertook the reviews of the psychometric literature on depression and anxiety.

Myfawnwy Lloyd Jones and **Suzy Paisley** assisted with the literature reviews.

Alicia O'Cathain supervised the qualitative research.

Michael Barkham provided clinical input into the psychometric analyses and overall design of the study.

Martin Knapp, **Sarah Byford** and **Simon Gilbody** contributed to the design and writing up of the study.

Glenys Parry provided a clinical input into the different aspects of the study and contributed to the design and conduct of the research.

All authors contributed to the writing up of the research.

References

1. Gilbody S, Sheldon T, House A. Outcomes research in mental health. Systematic review. *Br J Psychiatry* 2003;**181**:8–16. http://dx.doi.org/10.1192/bjp.181.1.8

2. Gladis MM, Gosch EA, Dishuk NM, Crits-Cristoph P. Quality of life: expanding the scope of clinical significance. *J Consult Clin Psychol* 1999;**67**:320–31. http://dx.doi.org/10.1037//0022-006X.67.3.320

3. Pharmaceutical Benefits Advisory Committee. *Guidelines for Preparing Submissions to the Pharmaceutical Benefits Advisory Committee: Australia*. Department of Health (Australia); 2008. URL: www.health.gov.au/internet/main/publishing.Nsf/content/pbacguidelines-index (accessed 2 January 2014).

4. International Society for Pharmacoeconomics and Outcomes Research (ISPOR). *General Methods for the Assessment of the Relation of Benefits to Costs: Germany*. ISPOR; 2009. URL: www.ispor.org/PEguidelines/countrydet.asp?c=9&t=1 (accessed 2 January 2014).

5. Health Information and Quality Authority (HIQA). *Guidelines for the Economic Evaluation of Health Technologies In Ireland*. HIQA; 2010. URL: www.hiqa.ie/publication/guidelines-economic-evaluation-health-technologies-ireland (accessed 2 January 2014).

6. International Society for Pharmacoeconomics and Outcomes Research (ISPOR). *China Guidelines for Pharmacoeconomic Evaluations*. ISPOR; 2011. URL: www.ispor.org/PEguidelines/countrydet.asp?c=28&t=4 (accessed 2 January 2014).

7. International Society for Pharmacoeconomics and Outcomes Research (ISPOR). *Health Technology Assessment Guideline: Thailand*. ISPOR; 2008. URL: www.ispor.org/PEguidelines/countrydet.asp?c=36&t=2 (accessed 2 January 2014).

8. International Society for Pharmacoeconomics and Outcomes Research (ISPOR). *Pharmacoeconomic Evaluations in Belgium*. ISPOR; 2008. URL: www.ispor.org/PEguidelines/countrydet.asp?c=3&t=1 (accessed 2 January 2014).

9. International Society for Pharmacoeconomics and Outcomes Research (ISPOR). *Spanish Recommendations on Economic Evaluation of Health Technologies* (Spanish version 2009). ISPOR; 2009. URL: www.ispor.org/PEguidelines/countrydet.asp?c=20&t=4 (accessed 2 January 2014).

10. International Society for Pharmacoeconomics and Outcomes Research (ISPOR). *Guidelines on how to conduct pharmacoeconomic analyses: Norway*. ISPOR; 2012. URL: www.ispor.org/PEguidelines/countrydet.asp?c=15&t=1 (accessed 2 January 2014).

11. Brazier J. Is the EQ-5D fit for purpose in mental health? *Br J Psychiatry* 2010;**197**:348–9. http://dx.doi.org/10.1192/bjp.bp.110.082453

12. Saarni S, Viertio D, Perala J, Koskinen S, Lonqvist J, Suvisaari J. Quality of life of people with schizophrenia, bipolar disorder and other psychotic disorders. *Br J Psychiatry* 2010;**197**:386–94. http://dx.doi.org/10.1192/bjp.bp.109.076489

13. Brazier J, Ratcliffe J, Tsuchiya A, Solomon J. *Measuring and Valuing Health for Economic Evaluation*. Oxford: Oxford University Press; 2007.

14. Williams A. Economics of coronary artery bypass grafting. *Br Med J* 1985;**291**:326–9. http://dx.doi.org/10.1136/bmj.291.6491.326

15. Dolan P. Modelling valuations for EuroQol health states. *Med Care* 1997;**35**:1095–108. http://dx.doi.org/10.1097/00005650-199711000-00002

16. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ* 2002;**21**:271–92. http://dx.doi.org/10.1016/S0167-6296(01)00130-8

17. Feeny D, Furlong W, Torrance GW, Goldsmith CH, Zhu Z, De Pauuw D, *et al.* Multiattribute and single-attribute utility functions for the Health Utilities Index Mark 3 system. *Med Care* 2002;**40**:113–28. http://dx.doi.org/10.1097/00005650-200202000-00006

18. Kaplan R, Bush JW. Health-related quality of life measurement for evaluation research and policy analysis. *Health Psychol* 1982;**1**:61–80. http://dx.doi.org/10.1037/0278-6133.1.1.61

19. Ware JE, Snow KK, Kolinski M, Gandeck B. *SF-36 Health Survey Manual and Interpretation Guide*. Boston, MA: The Health Institute, New England Medical Center; 1993.

20. Brazier J, Deverill M, Harper R, Booth A. A review of the use of health status measures in economic evaluation. *Health Technol Assess* 1999;**3**(9).

21. Brazier J, Roberts J. Estimating a preference-based index from the SF-12. *Med Care* 2004;**42**:851–9.

22. Brown J, Bowling A, Flynn TN. *Models of Quality of Life. A Taxonomy and Systematic Review of the Literature*. European Forum on Population Ageing Research/Quality of Life; 2004. URL: http://eprints.kingston.ac.uk/17177 (accessed 2 January 2014).

23. Johansson S. Conceptualising and measuring quality of life for national policy: from the Swedish Level of Living Survey to an epistemology of the democratic process. *Soc Indic Res* 2002;**58**:13–32.

24. Veenhoven R. Why social policy needs subjective indicators. *Social Indic Res* 2002;**58**:33–45. http://dx.doi.org/10.1007/0-306-47513-8_3

25. Maslow A. *Motivation and Personality*. New York, NY: Harper; 1954.

26. Rettig KD, Leichnetritt RD. A general theory for perceptual indicators of family life quality. *Soc Indic Res* 1999;**47**:307–42.

27. Higgs P, Hyde M, Wiggins RD, Blane DB. Research quality of life in early old age: the importance of the sociological dimension. *Soc Policy Adm* 2003;**37**:239–52. http://dx.doi.org/10.1111/1467-9515.00336

28. Netten A, Burge P, Malley J, Potoglou D, Towers AM, Brazier J, *et al.* Outcomes of social care for adults: developing a preference weighted measure. *Health Technol Assess* 2012;**16**(16).

29. Dolan P, Peasgood T, White M. Do we know what makes us happy? A review of the economic literature on the factors associated with subjective wellbeing. *J Econ Psychol* 2008;**29**:94–122.

30. Bentham J. *Deontology*. Oxford: Clarendon Press; 1983.

31. Kahneman D, Wakker P, Sarin R. Back to Bentham? Explorations of experienced utility. *Q J Econ* 1997;**112**:374–406. http://dx.doi.org/10.1162/003355397555235

32. Diener E, Lucas RE. Subjective emotional well-being. In Lewis M, Haviland JM, editors. *Handbook of Emotions*. 2nd edn. New York, NY: Guilford; 2000. pp. 325–37.

33. Diener E, Emmons RA, Larson RJ, Griffin S. The satisfaction with life scale. *J Pers Assess* 1985;**49**:71–5. http://dx.doi.org/10.1207/s15327752jpa4901_13

34. Fisher I. Is 'utility' the most suitable term for the concept it is used to denote? *Am Econ Rev* 1918;**8**:335–7.

35. Dolan P, Kahneman D. Interpretations of utility and their implications for the valuation of health. *Econ J* 2008;**118**:215–34. http://dx.doi.org/10.1111/j.1468-0297.2007.02110.x

36. Sen AK. *Commodities and Capabilities*. Oxford: Oxford University Press; 1985.

37. Sen AK. Capability and wellbeing. In Nussbaum MC, Sen AK, editors. *The Quality of Life*. New York, NY: Oxford Clarendon Press; 1993. pp. 30–53.

38. Nussbaum MC. Capabilities as fundamental entitlements: Sen and social justice. *Fem Econ* 2003;**9**:33–59. http://dx.doi.org/10.1080/1354570022000077926

39. Grewal I, Lewis J, Flynn TN, Brown J, Coast J. Developing attributes for a generic quality of life measure for older people: preferences or capabilities? *Soc Sci Med* 2006;**62**:1891–901. http://dx.doi.org/10.1016/j.socscimed.2005.08.023

40. Al-Janabi H, Flynn T, Coast J. Development of a self-report measure of capability well-being for adults: the ICECAP-A. *Qual Life Res* 2012;**21**:167–76.

41. Barry MM. *Wellbeing and Life Satisfaction as Components of Quality of Life in Mental Disorders*. Chichester: John Wiley and Sons; 2006.

42. Zissi A, Barry MM, Cochrane R. A mediational model of quality of life for individuals with severe mental health problems. *Psychol Med* 1998;**28**:1221–30. http://dx.doi.org/10.1017/S0033291798007338

43. Bowling A. *Measuring Health: A Review of Quality of Life and Measurement Scales*. Milton Keynes: Oxford University Press; 2002.

44. World Health Organization (WHO). *Constitution of the World Health Organization*. Geneva: WHO; 1948.

45. Leidy NK, Revicki DA, Geneste B. Recommendations for evaluating the validity of quality of life claims for labeling and promotion. *Value Health* 1999;**2**:113–27. http://dx.doi.org/10.1046/j.1524-4733.1999.02210.x

46. Katschnig H. Quality of life in mental disorders: challenges for research and clinical practice. *World Psychiatry* 2006;**5**:139–45.

47. Chisholm D, Healey A, Knapp M. QALYs and mental health care. *Soc Psychiatry Epidemiol* 1997;**32**:68–75. http://dx.doi.org/10.1007/BF00788923

48. van de Willige G, Wiersma D, Nienhuis FJ, Jenner JA. Changes in quality of life in chronic psychiatric patients: a comparison between EuroQol (EQ-5D) and WHOQoL. *Qual Life Res* 2005;**14**:441–51. http://dx.doi.org/10.1007/s11136-004-0689-y

49. Moock JK. Comparing preference-based quality-of-life measures: results from rehabilitation patients with musculoskeletal, cardiovascular, or psychosomatic disorders. *Qual Life Res* 2008;**17**:485–95. http://dx.doi.org/10.1007/s11136-008-9317-6

50. Lamers LM, Bouwmans CA, van Straten A, Donker MC, Hakkaart L. Comparison of EQ-5D and SF-6D utilities in mental health. *Health Econ* 2006;**15**:1229–36. http://dx.doi.org/10.1002/hec.1125

51. Brazier J, Rowen D, Mavranezouli I, Tsuchiya A, Young T, Yang Y. Developing and testing methods for deriving preference-based measures of health from condition-specific measures (and other patient-based measures of outcome). *Health Technol Assess* 2012;**16**(32).

52. Mavranezouli I, Brazier J, Young T, Barkham M. Using Rasch analysis to form plausible health states amenable to valuation: the development of CORE-6D from a measure of common mental health problems. *Qual Life Res* 2011;**20**:321–33. http://dx.doi.org/10.1007/s11136-010-9768-4

53. Fitzpatrick R, Davey C, Buxton M, Jones D. Evaluating patient-based outcome measures for use in clinical trials. *Health Technol Assess* 1998;**6**(29).

54. Brazier J, Deverill M. A checklist for judging preference-based measures of health related quality of life: learning from psychometrics. *Health Econ* 1999;**8**:41–51. http://dx.doi.org/10.1002/(SICI) 1099-1050(199902)8:1%3C41::AID-HEC395%3E3.3.CO;2-R

55. Streiner DL, Norman GR. Health Measurement Scales: a Practical Guide to Their Development and Use. Oxford: Oxford University Press; 1989.

56. Bowling A. *Measuring Health. A Review of Quality of Life Measurement Scales*. Buckingham: Open University Press; 1997.

57. American Psychological Association. *Committee on Psychological Tests. Technical Recommendations for Psychological Tests and Diagnostic Techniques*. Washington, DC: American Psychological Association; 1954.

58. US Department of Health and Human Services Food and Drug Administration (FDA), Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Resesarch (CBER), Center for Devices and Radiological Health (CDRH). *Guidance for Industry: Patient-reported Outcome Measures: Use in Medical Product Development to Support Labelling Claims*. FDA; 2009. URL: www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ ucm193282.pdf (accessed 2 January 2014).

59. Katz JN, Larson MG, Phillips CB, Fossel AH, Liang MH. Comparative measurement sensitivity of short and longer health status instruments. *Med Care* 1992;**30**:917–25. http://dx.doi.org/10.1097/00005650-199210000-00004

60. Papaioannou D, Brazier J, Parry G. How valid and responsive are generic health status measures such as EQ-5D and SF-6D, in schizophrenia? A systematic review. *Value Health* 2011;**14**:907–20.

61. Papaioannou D, Brazier JE, Parry G. How to measure quality of life for cost effectiveness analyses in personality disorders? A systematic review. *J Personal Disord* 2013;**27**:383–401. http://dx.doi.org/10.1521/pedi_2013_27_075

62. Peasgood T, Brazier J, Papaioannou D. *A Systematic Review of the Validity and Responsiveness of the EQ-5D and the SF-6D for Depression and Anxiety*. Health Economics and Decision Science discussion paper 12/15. Sheffield: Health Economics and Decision Science, University of Sheffield; 2012.

63. Fitzsimmons D, Gilbert J, Howse F, Young T, Arraras J, Brident A, *et al.* A systematic review of the use and validation of health-related quality of life instruments in older cancer patients. *Eur J Cancer Care* 2004;**45**:19–32. http://dx.doi.org/10.1016/j.ejca.2008.07.036

64. Walters SJ. *Quality of Life Outcomes in Clinical Trials and Health-care Evaluation: A Practical Guide to Analysis and Interpretation*. Chichester: Wiley-Blackwell; 2009.

65. Cohen J. *Statistical Power Analysis for the Behavioural Sciences*. 2nd edn. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.

66. Lenert LA, Rupnow MF, Elnitsky C. Application of a disease-specific mapping function to estimate utility gains with effective treatment of schizophrenia. *Health Qual Life Outcomes* 2005;**3**:57.

67. Bebbington PE, Angermeyer M, Azorin JM, Marwaha S, Marteau F, Tourni M. Side-effects of antipsychotic medication and health-related quality of life in schizophrenia. *Acta Psychiatr Scand* 2009;**119**:22–8. http://dx.doi.org/10.1111/j.1600-0447.2008.01310.x

68. McCrone P, Patel A, Knapp M, Schene A, Koeter M, Amaddeo F, *et al.* A comparison of SF-6D and EQ-5D utility scores in a study of patients with schizophrenia. *J Ment Health Policy Econ* 2009;**12**:27–31.

69. Revicki D, Genduso LA, Hamilton SH, Ganoczy D, Beasley CM. Olanzapine versus haloperidal in the treatment of schizophrenia and other psychotic disorders: quality of life and clinical outcomes of a randomized controlled trial. *Qual Life Res* 1999;**8**:417–26.

70. Strakowski SM, Johnson JL, Delbello MP, Hamer RM, Green AI, Tohen M, *et al.* Quality of life during treatment with haloperidol or olanzapine in the year following a first psychotic episode. *Schizophr Res* 2005;**78**:161–9. http://dx.doi.org/10.1016/j.schres.2005.04.017

71. Tunis SL, Croghan TW, Heilman DK, Johnston BM, Obenchain RL. Reliability, validity and application of the Medical Outcomes Study 36-Item Short Form Health Survey (SF-36) in schizophrenic patients treated with olanzapine verus haloperidal. *Med Care* 1999;**37**:678–91.

72. Nasrallah HA, Duchesne I, Mehnert A, Janagap C, Eerdekens M, Nasrallah HA, *et al.* Health-related quality of life in patients with schizophrenia during treatment with long-acting, injectable risperidone. *J Clin Psychiatry* 2004;**65**:531–6. [Erratum appears in *J Clin Psychiatry* 2004;**65**:1150]. http://dx.doi.org/10.4088/JCP.v65n0412

73. Russo J, Trujillo CA, Wingerson D, Decker K, Ries R, Wetzler H, *et al.* The MOS 36-Item Short Form Health Survey: reliability, validity and preliminary findings in schizophrenia outpatients. *Med Care* 1998;**36**:752–6. http://dx.doi.org/10.1097/00005650-199805000-00015

74. Kebede D, Alem A, Shibre T, Alemayehu N, Negussie D, Beyero T, *et al.* Health related quality of life (SF-36) survey in Butajira, rural Ethopia: normative data and evaluation of reliability and validity. *Ethiop Med J* 2004;**42**:289–97.

75. Kebede D, Alem A, Shibre T, Negash A, Deyassa N, Beyero T, *et al.* Short-term symptomatic and functional outcomes of schizophrenia in Butajira. *Ethiop Med J* 2005;**78**:171–85. http://dx.doi.org/10.1016/j.schres.2005.05.028

76. Norholm V, Bech P. Quality of life assessment in schizophrenia: applicability of the Lehman Quality of Life Questionnaire (TL-30). *Nord J Psychiatry* 2007;**61**:438–42. http://dx.doi.org/10.1080/08039480701773154

77. Folsom DP, Depp C, Palmer BW, Mausbach BT, Golshan S, Fellows I, *et al.* Physical and mental health-related quality of life among older people with schizophrenia. *Schizophr Res* 2009;**108**:207–13. http://dx.doi.org/10.1016/j.schres.2008.12.008

78. Sciolla A, Patterson TL, Wetherell JL, McAdams LA, Jeste DY. Functioning and well-being of middle aged and older patients with schizophrenia: measurement with the 36-item short form (SF-36) health survey. *Am J Geriatr Psychol* 2003;**11**:629–37. http://dx.doi.org/10.1176/appi.ajgp.11.6.629

79. Pukrop R, Schlaak V, Moller-Leimkuhler AM, Albus M, Czernik A, Klosterkotter J, *et al.* Reliability and validity of quality of life assessed by the Short-Form SF-36 and the Modular System for Quality of Life in patients with schizophrenia and patients with depression. *Psychol Res* 2003;**119**:63–79.

80. Law CW, Chen EY, Cheung EF, Chan RC, Wong JG, Lam CL, *et al.* Impact of untreated psychosis on quality of life in patients with first-episode schizophrenia. *Qual Life Res* 2005;**14**:1803–11. http://dx.doi.org/10.1007/s11136-005-3236-6

81. Meijer CJ, Schene AH, Koeter MW. Quality of life in schizophrenia measured by the MOS SF-36 and the Lancashire Quality of Life Profile: a comparison. *Acta Psychiatr Scand* 2002;**105**:293–300. http://dx.doi.org/10.1034/j.1600-0447.2002.1198.x

82. Wilkinson G, Hesdon B, Wild D, Cookson R, Farina C, Sharma V, *et al.* Self-report quality of life measure for people with schizophrenia: the SQLS. *Br J Psychiatry* 2000;**177**:42–6. http://dx.doi.org/10.1192/bjp.177.1.42

83. Milliken HI, Whitehorn D, Rui Q, Ronson K. SF-36: a measure of recovery in the first six months of treatment for schizophrenia. *Schizophr Bull* 2007;**33**:597–8.

84. Auquier P, Simeoni MC, Sapin C, Reine G, Aghababian V, Cramer J, *et al.* Development and validation of a patient-based health-related quality of life questionnaire in schizophrenia: the S-QoL. *Schizophr Res* 2003;**63**:137–49. http://dx.doi.org/10.1016/S0920-9964(02)00355-9

85. Bobes J, Gonzalez MP, Saiz PA. The SF-36 versus the WHOQOL-100 and -26 in schizophrenic private out-patients. *Qual Life Res* 1997;**6**:126.

86. Prieto L, Sacristan JA, Hormaechea JA, Casado A, Badia X, Gomez JC, *et al.* Psychometric validation of a generic health-related quality of life measure (EQ-5D) in a sample of schizophrenic patients. *Curr Med Res Opin* 2004;**20**:827–35. http://dx.doi.org/10.1185/030079904125003674

87. Badia X, Casado A, Sacristan JA, Gomez JC, Gregor KJ. Antipsychotic treatment and changes in health related quality of life in patients with schizophrenia using EQ-5D. International Society for Pharmacoeconomics and Outcomes Research, Second Annual European Meeting, Edinburgh, November 1999.

88. Reine G, Simeoni MC, Auquier P, Loundou A, Aghababian V, Lancon C. Assessing health-related quality of life in patients suffering with schizophrenia: a comparison of instruments. *Eur Psychiatry* 2005;**20**:510–19. http://dx.doi.org/10.1016/j.eurpsy.2005.05.009

89. Barton GR, Hodgekins J, Mugford M, Jones PB, Croudace T, Fowler D. Measuring the benefits of treatment for psychosis: validity and responsiveness of the EQ-5D. *Br J Psychiatry* 2009;**195**:170–7. http://dx.doi.org/10.1192/bjp.bp.108.057380

90. Konig HH, Roick C, Angermeyer M. Validity of the EQ-5D in assessing and valuing health status in patients with schizophrenia, schizotypal or delusional disorders. *Eur Psychiatry* 2007;**22**:177–87. http://dx.doi.org/10.1016/j.eurpsy.2006.08.004

91. Konig HH, Gunther OH, Angermeyer MC, Roick C. Utility assessment in patients with mental disorders: validity and discriminative ability of the time trade-off method. *Pharmacoeconomics* 2009;**27**:405–19. http://dx.doi.org/10.2165/00019053-200927050-00005

92. Scalone L, Pirfo E, Mencacci C, Ferranini L, Berto P, Sturkenboom MC, *et al.* Relationship between clinical outcomes and patients' reported outcomes in schizophrenia: the contribution of the EQ-5D. *Value Health* 2008;**11**:A592.

93. Dunayevich E, Ascher-Svanum H, Zhao F, Jacobson JG, Phillips GA, Dellva MA, *et al.* Longer time to antipsychotic treatment discontinuation for any cause is assocated with better functional outcomes for patients with schizophrenia, schizophreniform disorder or schizoaffective disorder. *J Clin Psychiatry* 2007;**68**:1163–71.

94. Phillips GA, Van Brunt DL, Roychowdhury SM, Xu W, Naber D. The relationship between quality of life and clinical efficacy from a randomized trial comparing olanzapine and ziprasidone. *J Clin Psychiatry* 2006;**67**:1397–403. http://dx.doi.org/10.4088/JCP.v67n0910

95. Pyne JM, Sullivan G, Kaplan R, Williams DK. Comparing the sensitivity of generic effectiveness measures with symptom improvement in persons with schizophrenia. *Med Care* 2003;**41**:208–17. http://dx.doi.org/10.1097/01.MLR.0000044900.72470.D4

96. Sanderson K, Andrews G, Alem A. Prevalence and severity of mental health-related disability and relationship to diagnosis. *Psychiatr Serv* 2002;**53**:80–6. http://dx.doi.org/10.1176/appi.ps.53.1.80

97. Albert U, Rosso G, Maina G, Bogetto F. Impact of anxiety disorder comorbidity on quality of life in euthymic bipolar disorder patients: differences between bipolar I and II subtypes. *J Affect Disord* 2008;**105**:297–303. http://dx.doi.org/10.1016/j.jad.2007.05.020

98. Arbuckle R, Frye MA, Brecher M, Paulsson B, Rajagopalan K, Palmer S, *et al.* The psychometric validation of the Sheehan Disability Scale (SDS) in patients with bipolar disorder. *Psychiatry Res* 2009;**165**:163–74. http://dx.doi.org/10.1016/j.psychres.2007.11.018

99. Hayhurst H, Palmer S, Abbott R, Johnston T, Scott J. Measuring health-related quality of life in bipolar disorder: relationship of the EuroQol (EQ-5D) to condition-specific measures. *Qual Life Res* 2006;**15**:1271–80. http://dx.doi.org/10.1007/s11136-006-0059-z

100. Hakkaart-van Roijen L, Hoeijenbos MB, Regeer EJ, ten Have M, Nolen WA, Veraart CP, *et al.* The societal costs and quality of life of patients suffering from bipolar disorder in the Netherlands. *Acta Psychiatr Scand* 2004;**110**:383–92. http://dx.doi.org/10.1111/j.1600-0447.2004.00403.x

101. Vojta C, Kinosian B, Glick H, Altshuler L, Bauer MS. Self-reported quality of life across mood states in bipolar disorder. *Compr Psychiatry* 2001;**42**:190–5. http://dx.doi.org/10.1053/comp.2001.23143

102. Arnold LM, Witzeman KA, Swank ML, McElroy SL, Keck PE Jr, Health-related quality of life using the SF-36 in patients with bipolar disorder compared with patients with chronic back pain and the general population. *J Affect Disord* 2000;**57**:235–9. http://dx.doi.org/10.1016/S0165-0327(99)00042-7

103. Bauer MS, Altshuler L, Evans DR, Beresford TP, Williford WO, Hauger R. Prevalence and distinct correlates of anxiety, substance, and combined comorbidity in a multi-site public sector sample with bipolar disorder. *J Affect Disord* 2005;**85**:301–15. http://dx.doi.org/10.1016/j.jad.2004.11.009

104. Cuijpers P, Smits N, Donker T, ten Have M, de Graaf R. Screening for mood and anxiety disorders with the five-item, the three-item and the two-item Mental Health Inventory. *Psychiatry Res* 2009;**168**:250–5. http://dx.doi.org/10.1016/j.psychres.2008.05.012

105. Depp CA, Davis CE, Mittal D, Patterson TL, Jeste DV. Health-related quality of life and functioning of middle-aged and elderly adults with bipolar disorder. *J Clin Psychiatry* 2006;**67**:215–21. http://dx.doi.org/10.4088/JCP.v67n0207

106. Depp CA, Mausbach BT, Eyler LT, Palmer BW, Cain AE, Lebowitz BD, *et al.* Performance-based and subjective measures of functioning in middle-aged and older adults with bipolar disorder. *J Nerv Ment Dis* 2009;**197**:471–5. http://dx.doi.org/10.1097/NMD.0b013e3181ab5c9b

107. Gutierrez-Rojas L, Gurpegui M, yuso-Mateos JL, Gutierrez-Ariza JA, Ruiz-Veguilla M, Jurado D, *et al.* Quality of life in bipolar patients: a comparison with a general population sample. *Bipolar Disord* 2008;**10**:625–34. http://dx.doi.org/10.1111/j.1399-5618.2008.00604.x

108. Kebede D, Alem A, Shibire T, Deyassa A, Negash A, Beyero T, *et al.* Symptomatic and functional outcome of bipolar disorder in Butarija, Ethiopia. *J Affect Disord* 2006;**90**:239–49.

109. Leidy NK, Palmer C, Murray M, Robb J, Revicki D. Health-related quality of life assessment in euthymic and depressed patients with bipolar disorder. Psychometric performance of four self-report measures. *J Affect Disord* 1998;**48**:207–14.

110. Maina G, Albert U, Bellodi L, Colombo C, Faravelli C, Monteleone P, *et al.* Health-related quality of life in euthymic bipolar disorder patients: differences between bipolar I and II subtypes. *J Clin Psychiatry* 2007;**68**:207–12. http://dx.doi.org/10.4088/JCP.v68n0205

111. Namjoshi M, Rajamannar G, Jacobs T, Sanger TM, Risser R, Tohen MF, *et al.* Economic, clinical, and quality-of-life outcomes associated with olanzapine treatment in mania. Results from a randomized controlled trial. *J Affect Disord* 2002;**69**:109–18. http://dx.doi.org/10.1016/S0165-0327(01)00310-X

112. Shi L, Namjoshi MA, Swindle R, Yu X, Risser R, Baker RW, *et al.* Effects of olanzapine alone and olanzapine/fluoxetine combination on health-related quality of life in patients with bipolar depression: secondary analyses of a double-blind, placebo-controlled, randomized clinical trial. *Clin Ther* 2004;**26**:125–34. [Erratum appears in *Clin Ther* 2004;**26**:1934.]

113. Shi L, Namjoshi MA, Zhang F, Gandhi G, Edgell ET, Tohen M, *et al.* Olanzapine versus haloperidol in the treatment of acute mania: clinical outcomes, health-related quality of life and work status. *Int Clin Psychopharmacol* 2002:**17**:227–37. http://dx.doi.org/10.1097/00004850-200209000-00003

114. Sierra P, Livianos L, Rojo L. Quality of life for patients with bipolar disorder: relationship with clinical and demographic variables. *Bipolar Disord* 2005:**7**:159–65. http://dx.doi.org/10.1111/j.1399-5618.2005.00186.x

115. Simon GE, Bauer MS, Ludman EJ, Operskalski BH, Unutzer J. Mood symptoms, functional impairment, and disability in people with bipolar disorder: specific effects of mania and depression. *J Clin Psychiatry* 2007;**68**:1237–45. http://dx.doi.org/10.4088/JCP.v68n0811

116. Yatham LN, Lecrubier Y, Fieve RR, Davis KH, Harris SD, Krishnan AA. Quality of life in patients with bipolar I depression: data from 920 patients. *Bipolar Disord* 2004;**6**:379–85. http://dx.doi.org/10.1111/j.1399-5618.2004.00134.x

117. Zhang H, Wisniewski SR, Bauer MS, Sachs GS, Thase ME. Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD) investigators. Comparisons of perceived quality of life across clinical states in bipolar disorder: data from the first 2000 Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD) participants. *Compr Psychiatry* 2006;**47**:161–8.

118. Jarema M, Konieczynska Z. Quality of life in schizophrenia: impact of psycopathology, patients' gender and antipsychotic treatment. *Int J Psychiatry Clin Pract* 2001;**5**:19–26.

119. Narud K, Mykletun A, Dahl AA. Quality of life in patients with personality disorders seen at an ordinary psychiatric outpatient clinic. *BMC Psychiatry* 2005;**5**:10.

120. Soeteman D, Verheul R, Busschbach J. The burden of disease in personality disorders: diagnosis-specific quality of life. *J Personal Disord* 2008;**22**:259–68. http://dx.doi.org/10.1521/pedi.2008.22.3.259

121. Hueston WJ, Mainous AG III, Schilling R. Patients with personality disorders: functional status, health care utilization, and satisfaction with care. *J Fam Pract* 1996;**42**:54–60.

122. Jackson HJ, Burgess PM. Personality disorders in the community: a report from the Australian National Survey of Mental Health and Wellbeing. *Soc Psychiatry Psychiatr Epidemiol* 2000;**35**:531–8. http://dx.doi.org/10.1007/s001270050276

123. Jackson HJ, Burgess PM. Personality disorders in the community: results from the Australian Survey of Mental Health and Wellbeing Part II. Relationships between personality disorder, Axis I mental disorders and physical conditions with disability and health consultations. *Soc Psychiatry Psychiatr Epidemiol* 2002;**37**:251–60. http://dx.doi.org/10.1007/s001270200017

124. Bartak A, Andrea H, Spreeuwenberg MD, Ziegler UM, Dekker J, Rossum BV, *et al.* Effectiveness of outpatient, day hospital, and inpatient psychotherapeutic treatment for patients with cluster B personality disorders. *Psychother Psychosom* 2011;**80**:28–38. http://dx.doi.org/10.1159/000321999

125. Bartak A, Andrea H, Spreeuwenberg MD, Thunnissen M, Ziegler UM, Dekker J, *et al.* Patients with cluster A personality disorders in psychotherapy: an effectiveness study. *Psychother Psychosom* 2011;**80**:88–99.

126. Bartak A, Spreeuwenberg MD, Andrea H, Holleman L, Rijnierse P, Rossum BV, *et al.* Effectiveness of different modalities of psychotherapeutic treatment for patients with cluster C personality disorders: results of a large prospective multicentre study. *Psychother Psychosom* 2010;**79**:20–30. http://dx.doi.org/10.1159/000254902

127. Soeteman DI, Timman R, Trijsburg RW, Verheul R, Busschbach JJ. Assessment of the burden of disease among inpatients in specialized units that provide psychotherapy. *Psychiatr Serv* 2005;**56**:1153–5. http://dx.doi.org/10.1176/appi.ps.56.9.1153

128. van Asselt AD, Dirksen CD, Arntz A, Giesen-Bloo JH, Severens JL. The EQ-5D: A useful quality of life measure in borderline personality disorder? *Eur Psychiatry* 2009;**24**:79–85. http://dx.doi.org/10.1016/j.eurpsy.2008.11.001

129. Aydemir R, Ergun H, Soygur H, Kesebir S, Tulunay C. [Quality of life in major depressive disorder: a cross-sectional study.] *Turk Psikiyatri Derg* 2009;**20**:205–12.

130. Bosmans JE, Hermens MLM, de Bruijne MC, van Hout HPJ, Terluin B, Bouter LM, *et al.* Cost-effectiveness of usual general practitioner care with or without antidepressant medication for patients with minor or mild-major depression. *J Affect Disord* 2008;**111**:106–12. http://dx.doi.org/10.1016/j.jad.2008.02.002

131. Caruso R, Rossi A, Barraco A, Quail D, Grassi L. The Factors Influencing Depression Endpoints Research (FINDER) study: final results of Italian patients with depression. *Ann Gen Psychiatry* 2010;**9**:33.

132. Ergun H, Aydemir O, Kesebir S, Soygur H, Tulunay FC. SF-36 and EQ-5D quality of life instruments in major depressive disorder patients: comparisons of two different treatment options. *Value Health* 2007;**10**:A303. http://dx.doi.org/10.1016/S1098-3015(10)65099-1

133. Fernandez JL, Montgomery S, Francois C. Evaluation of the cost effectiveness of escitalopram versus venlafaxine XR in major depressive disorder. *Pharmacoeconomics* 2005;**23**:155–67. http://dx.doi.org/10.2165/00019053-200523020-00007

134. Gunther OH, Roick C, Angermeyer MC, Konig HH. The responsiveness of EQ-5D utility scores in patients with depression: a comparison with instruments measuring quality of life, psychopathology and social functioning. *J Affect Disord* 2008;**105**:87–91.

135. Konig HH, Born A, Heider D, Matschinger H, Heinrich S, Riedel-Heller SG, *et al.* Cost-effectiveness of a primary care model for anxiety disorders. *Br J Psychiatry* 2009;**195**:308–17. http://dx.doi.org/10.1192/bjp.bp.108.058032

136. Konig HH, Born A, Gunther O, Matschinger H, Heinrich S, Riedel-Heller SG, *et al.* Validity and responsiveness of the EQ-5D in assessing and valuing health status in patients with anxiety disorders. *Health Qual Life Outcomes* 2010;**8**:47. http://dx.doi.org/10.1186/1477-7525-8-47

137. Mann R, Gilbody S, Richards D. Putting the 'Q' in depression QALYs: a comparison of utility measurement using EQ-5D and SF-6D health related quality of life measures. *Soc Psychiatry Psychiatr Epidemiol* 2009;**44**:569–78. http://dx.doi.org/10.1007/s00127-008-0463-5

138. Mychaskiw M, Hoffman D, Dodge W. EQ-5D index scores by remission status in patients with generalised anxiety disorder. *Int J Neuropsychopharmacol* 2008;**11**:279.

139. Petrou S, Morrell J, Spiby H. Assessing the empirical validity of alternative multi-attribute utility measures in the maternity context. *Health Qual Life Outcomes* 2009;**7**:40. http://dx.doi.org/10.1186/1477-7525-7-40

140. Peveler R, Kendrick T, Buxton M, Longworth L, Baldwin D, Moore M, *et al.* A randomised controlled trial to compare the cost-effectiveness of tricyclic antidepressants, selective serotonin reuptake inhibitors and lofepramine. *Health Technol Assess* 2005;**9**(16).

141. Pyne JM, Fortney JC, Tripathi SP, Maciejewski ML, Edlund MJ, Williams DK. Cost-effectiveness analysis of a rural telemedicine collaborative care intervention for depression. *Arch Gen Psychiatry* 2010;**67**:812–21. http://dx.doi.org/10.1001/archgenpsychiatry.2010.82

142. Reed C, Monz BU, Perahia DG, Gandhi P, Bauer M, Dantchev N, *et al.* Quality of life outcomes among patients with depression after 6 months of starting treatment: results from FINDER. *J Affect Disord* 2009;**113**:296–302. http://dx.doi.org/10.1016/j.jad.2008.05.021

143. Revicki DA, Brandenburg N, Matza L, Hornbrook MC, Feeny D. Health-related quality of life and utilities in primary-care patients with generalised anxiety disorder. *Qual Life Res* 2008;**17**:1285–94.

144. Saarni SI, Suvisaari J, Sintonen H, Pirkola S, Koskinen S, Aromaa A, *et al.* Impact of psychiatric disorders on health-related quality of life: general population survey. *Br J Psychiatry* 2007;**190**:326–32. http://dx.doi.org/10.1192/bjp.bp.106.025106

145. Sapin C, Fantino B, Nowicki ML, Kind P. Usefulness of EQ-5D in assessing health status in primary care patients with major depressive disorder. *Health Qual Life Outcomes* 2004;**2**:20.

146. Sobocki P, Ekman M, Agren H, Krakau I, Runeson B, Martensson B, *et al.* Health-related quality of life measured with EQ-5D in patients treated for depression in primary care. *Value Health* 2007;**10**:153–60. http://dx.doi.org/10.1111/j.1524-4733.2006.00162.x

147. Stein MB, Roy-Byrne PP, Craske MG, Bystritsky A, Sullivan G, Pyne JM, *et al.* Functional impact and health utility of anxiety disorders in primary care outpatients. *Med Care* 2005;**43**:1164–70. http://dx.doi.org/10.1097/01.mlr.0000185750.18119.fd

148. Supina AL, Johnson JA, Patten SB, Williams JVA, Maxwell CJ. The usefulness of the EQ-5D in differentiating among persons with major depressive episode and anxiety. *Qual Life Res* 2007;**16**:749–54. http://dx.doi.org/10.1007/s11136-006-9159-z

149. Zivin K, McCarthy JF, McCammon RJ, Valenstein M, Post EP, Welsh DE, *et al.* Health-related quality of life and utilities among patients with depression in the Department of Veterans Affairs. *Psychiatr Serv* 2008;**59**:1331–4. http://dx.doi.org/10.1176/appi.ps.59.11.1331

150. Kendrick T, Peveler R, Longworth L, Baldwin D, Moore M, Chatwin J, *et al.* Cost-effectiveness and cost-utility of tricyclic antidepressants, selective serotonin reuptake inhibitors and lofepramine. *Br J Psychiatry* 2006;**188**:337–45.

151. Morrell CJ, Slade P, Warner R, Paley G, Dixon S, Walters SJ, *et al.* Clinical effectivneness of health visitor training in psychologically informed approaches for depression in postnatal women: pragmatic cluster randomised trial in primary care. *BMJ* 2009;**338**:a3045.

152. Parry G, Barkham M, Brazier J, Dent-Brown K, Hardy G, Kendrick T, *et al.* An evaluation of a new service model: Improving Access to Psychological Therapies demonstration sites 2006–2009. Final report. NIHR Service Delivery and Organisation Programme; 2011.

153. Tyrer P, Tom B, Byford S, Schmidt U, Jones V, Davidson K, *et al.* Differential effects of manual assisted cognitive behavior therapy on the treatment of recurrent deliberate self-harm and personality disturbance: the POPMACT study. *J Personal Disord* 2004;**18**:102–16. http://dx.doi.org/10.1521/pedi.18.1.102.32770

154. Gray R, Leese M, Bindman J, Becker T, Burti L, David A. Adherence therapy for people with schizophrenia: European multicentre randomised controlled trial. *Br J Psychiatry* 2006;**189**:508–14. http://dx.doi.org/10.1192/bjp.bp.105.019489

155. Crawford M, Killaspy H, Kalaitzaki E, Barrett B, Byford S. The MATISSE study: a randomised trial of group art therapy for people with schizophrenia. *BMC Psychiatry* 2010;**10**:65. http://dx.doi.org/10.1186/1471-244X-10-65

156. Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. *Acta Psychiatr Scand* 1983;**67**:361–70. http://dx.doi.org/10.1111/j.1600-0447.1983.tb09716.x

157. Bjelland I, Dahl AA, Haug TT, Neckelmann D. The validity of the Hospital Anxiety and Depression Scale: an updated literature review. *J Psychosom Res* 2002;**52**:69–77.

158. Lisspers J, Nygren A, Soderman E. The Hospital Anxiety and Depression Scale (HAD): some psychometric data for a Swedish sample. *Acta Psychiatr Scand* 1997;**96**:281–6. http://dx.doi.org/10.1111/j.1600-0447.1997.tb10164.x

159. Suarez-Mendoza AA, Cardiel MH, Caballero-Uribe CV, Ortega-Soto HA, Marquez-Marin M. Measurement of depression in Mexican patients with rheumatoid arthritis: validity of the Beck Depression Inventory. *Arthritis Care Res* 1997;**10**:194–9. http://dx.doi.org/10.1002/art.1790100307

160. Fossa SD, Dahl AA. Short Form 46 and Hospital Anxiety and Depression Scale. A comparison based on patients with testicular cancer. *J Psychosom Res* 2002;**52**:79–87.

161. Barkham M, Evans C, Margison F, McGrath G, Mellor-Clark J, Connell J, *et al.* The rationale for developing and implementing core batteries for routine use in service settings and psychotherapy outcome research. *J Ment Health* 1998;**7**:35–47.

162. Barkham M, Margison F, Leach C, Lucock M, Mellor-Clark J, Evans C, *et al.* Service profiling and outcome benchmarking using the CORE-OM: towards practice based evidence in the psychological therapies. *J Consult Clin Psychol* 2001;**69**:184–96. http://dx.doi.org/10.1037//0022-006X.69.2.184

163. Barkham M, Gilbert N, Connell J, Marshall C, Twigg E. Suitability and utility of the CORE-OM and CORE-A for assessing severity of presenting problems in psychological therapy services based in primary and secondary care settings. *Br J Psychiatry* 2005;**186**:239–46. http://dx.doi.org/10.1192/bjp.186.3.239

164. Evans C, Connell J, Barkham M, Margison F, McGrath G, Mellor-Clark J, *et al.* Towards a standardised brief outcome measure: psychometric properties and utility of the CORE-OM. *Br J Psychiatry* 2002;**180**:51–60. http://dx.doi.org/10.1192/bjp.180.1.51

165. Evans C, Mellor-Clark J, Margison K, Barkham M, Audin K, Connell J, *et al.* CORE: Clinical Outcomes in Routine Evaluation. *J Ment Health* 2000;**9**:247–55. http://dx.doi.org/10.1080/713680250

166. Barkham M, Mullin T, Leach C, Stiles WB, Lucock M. Stability of the CORE-OM and the BDI-I prior to therapy: evidence from routine practice. *Psychol Psychother* 2007;**80**:269–78. http://dx.doi.org/10.1348/147608306X148048

167. Gilbody S, Richards D, Barkham M. Diagnosing depression in primary care using self-completed instruments: UK valuations of PHQ-9 and CORE-OM. *Br J Gen Pract* 2007;**57**:650–2.

168. Ventura J, Lukoff D, Neuchterlein K, Liberman R, Green M, Shaner A, *et al.* Brief Psychiatric Rating Scale (BPRS) Expanded Version (4.0) scales, anchor points and administration manual. *Int J Psychiatry Clin Pract* 1993;**3**:227–43.

169. Mortimer M. Symptom rating scales and outcome in schizophrenia. *Br J Psychiatry* 2007;**191**:S7–14. http://dx.doi.org/10.1192/bjp.191.50.s7

170. Kay SR, Fiszbein A, Opler LA. The Positive and Negative Syndrome Scale (PANSS) for schizophrenia. *Schizophr Bull* 1987;**13**:261–76. http://dx.doi.org/10.1093/schbul/13.2.261

171. Kay SR, Opler LA, Lindenmayer JP. Reliability and validity of the positive and negative syndrome scale for schizophrenics. *Psychiatry Res* 1988;**23**:99–110. http://dx.doi.org/10.1016/0165-1781(88)90038-8

172. Byford S, Knapp M, Greenshields J, Ukoumunne OC, Jones V, Thompson S, *et al.* Cost-effectiveness of brief cognitive behaviour therapy versus treatment as usual in recurrent deliberate self-harm: a rational decision making approach. *Psychol Med* 2003;**33**:977–86.

173. Cleveland W. Robust locally weighted regression and smoothing scatterplots. *J Am Stat Assoc* 1979;**74**:829–36. http://dx.doi.org/10.2307/2286407

174. Leucht S, Kane JM, Kissling W, Hamann J, Etschel E, Engel R. Clinical implications of Brief Psychiatric Rating Scale scores. *Br J Psychiatry* 2005;**187**:366–71. http://dx.doi.org/10.1192/bjp.187.4.366

175. Leucht S, Kane JM, Kissling W, Hamann J, Etschel E, Engel R. What does the PANSS mean? *Schizophr Res* 2005;**79**:231–8. http://dx.doi.org/10.1016/j.schres.2005.04.008

176. Terwee CB, Dekker FW, Wiersinga WM, Prummel MF, Bossuyt PM. On assessing responsiveness of health-related quality of life instruments: guidelines for instrument evaluation. *Qual Life Res* 2003;**12**:349–62.

177. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, *et al.* Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res* 2011;**10**:1727–36. http://dx.doi.org/10.1007/s11136-011-9903-x

178. Brazier J, Yang Y, Tsuchiya A, Rowen D. A review of studies mapping (or cross walking) non-preference based measures. *Eur J Health Econ* 2010;**11**:215–25. http://dx.doi.org/10.1007/s10198-009-0168-z

179. King M, Nazroo J, Weich S, McKenzie K, Bhui K, Karlsen S. Psychotic symptoms in the general population of England: a comparison of ethnic groups (the EMPIRIC study). *Soc Psychiatry Psychiatr Epidemiol* 2005;**40**:375–81. http://dx.doi.org/10.1007/s00127-005-0900-7

180. Longworth L, Rowen D. *The Use of Mapping Methods to Estimate Health State Utility Values. NICE Decision Support Unit (DSU) Technical Support Document 10*. Sheffield: DSU, School of Health and Related Research, University of Sheffield; 2011.

181. Gray AM, Rivero-Arias O, Clark PM. Estimating the association between SF-12 responses and EQ-5D utility values by response mapping. *Med Decis Making* 2006;**26**:18–29. http://dx.doi.org/10.1177/0272989X05284108

182. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001;**16**:606–13. http://dx.doi.org/10.1046/j.1525-1497.2001.016009606.x

183. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self report version of PRIME-MD: the PHQ primary care study. *Arch Intern Med* 1999;**166**:1092–7.

184. Martin A, Rief W, Klaiberg A, Braehler E. Validity of the Brief Patient Health Questionnaire Mood Scale (PHQ-9) in the general population. *Gen Hosp Psychiatry* 2006;**28**:71–7. http://dx.doi.org/10.1016/j.genhosppsych.2005.07.003

185. Spitzer RL, Kroenke K, Williams JB, Lowe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006;**166**:1092–7. http://dx.doi.org/10.1001/archinte.166.10.1092

186. Swinson RP. The GAD-7 scale was accurate for diagnosing generalised anxiety disorder. *Evid Based Med* 2006;**11**:184. http://dx.doi.org/10.1136/ebm.11.6.184

187. Goldberg DP, Williams P. *A User's Guide to the General Health Questionnaire*. Windsor: Nelson; 1988.

188. Hardy G, Shapiro DA, Haynes CE, Rick JE. Validation of the General Health Questionnaire-12: Using a sample of employees from England's health care services. *Psychol Assess* 1999;**11**:159–65. http://dx.doi.org/10.1037//1040-3590.11.2.159

189. Tyrer P, Thompson S, Schmidt U, Jones V, Knapp M, Davidson K, *et al.* Randomised controlled trial of brief cognitive behaviour therapy versus treatment as usual in recurrent deliberate self-harm: the POPMACT study. *Psychol Med* 2003;**33**:977–86.

190. Lewis G, Pelosi AJ, Araya R, Dunn G. Measuring psychiatric disorder in the community: a standardised assessment for use by lay interviewers. *Psychol Med* 1992;**22**:465–86.

191. Mortimer D, Segal L, Hawthorne G, Harris A. Item-based versus subscale-based mapping form the SF-36 to a preference-based quality of life measure. *Value Health* 2007;**10**:398–470.

192. Versteegh MM, Rowen D, Brazier J, Stolk EA. Mapping onto EQ-5D for patients in poor health. *Health Qual Life Outcomes* 2010;**8**:141.

193. Lin F, Longworth L, Pickard S. Evaluation of content on EQ-5D as compared to disease-specific utility measures. *Qual Life Res* 2012;**22**:853–74. http://dx.doi.org/10.1007/s11136-012-0207-6

194. Serrano-Aguilar P, Ramallo-Farina Y, Trujillo-Martin Mdel M, Munoz-Navarro SR, Perestelo-Perez L, de las Cuevas-Castresana C. The relationship among mental health status (GHQ-12), health related quality of life (EQ-5D) and health-state utilities in a general population. *Epidemiol Psichiatr Soc* 2009;**18**:229–39.

195. Jacobs R. *Investigating Patient Outcome Measures in Mental Health.* Research report for the Office of Health Economics Commission on NHS Productivity. York: Centre for Health Economics; 2007.

196. Barnett-Paige E, Thomas J. Methods for the synthesis of qualitative research: a critical review. *BMC Med Res Methodol* 2009;**9**:59. http://dx.doi.org/10.1186/1471-2288-9-59

197. Paterson DL, Thomas SE. *Meta-Study of Qualitative Health Research: A Practical Guide to Meta-Analysis and Meta-Synthesis*. Thousand Oaks, CA: Sage; 2001.

198. Ritchie J, Spencer L. *Qualitative Data Analysis for Applied Policy Research*. London: Routledge; 1994.

199. Connell J, Brazier JE, O'Cathain A, Lloyd-Jones M, Paisley S. Quality of life of people with mental health problems: a synthesis of qualitative research. *Health Qual Life Outcomes* 2012;**10**:138. http://dx.doi.org/10.1186/1477-7525-10-138

200. Sandelowski M, Docherty S, Emden C. Qualitative meta-synthesis: issues and techniques. *Res Nurs Health* 1997;**20**:365–71.

201. Hagerty BMK, Lynch-Sauer J, Patusky KL, Bouwsema M, Collier P. Sense of belonging: a vital mental health concept. *Arch Psychiatr Nurs* 1992;**6**:172–217. http://dx.doi.org/10.1016/0883-9417(92)90028-H

202. Wilcock AA. Reflections on doing, being and becoming. *Can J Occup Ther* 1998;**65**(5):279–85.

203. de Leval N. The Three Time Dimensions Synoptic Scale (3TSS) for Depressive Population. *Quality of Life Newsletter* 2001;**26**:15–16.

204. Diener E. Guidelines for National Indicators of Subjective Well-Being and Ill-Being. *Appl Res Qual Life* 2006;**1**:151–7. http://dx.doi.org/10.1007/s11482-006-9007-x

205. Clarke DM, Kissane DW. Demoralization: its phenomenology and importance. *Aust N Z J Psychiatry* 2002;**36**:733–42. http://dx.doi.org/10.1046/j.1440-1614.2002.01086.x

206. Corring DJ. *'Being Normal': Quality of Life Domains for Persons with a Mental Illness*. Dissertation Abstracts International Section B: Sciences and Engineering 2006;**67**.

207. Michalak EE, Yatham LN, Maxwell V, Hale S, Lam RW. The impact of bipolar disorder upon work functioning: a qualitative analysis. *Bipolar Disord* 2007;**9**:126–43. http://dx.doi.org/10.1111/j.1399-5618.2007.00436.x

208. Cook S, Chambers E. What helps and hinders people with psychotic conditions doing what they want in their daily lives. *Br J Occup Ther* 2009;**72**:238–48.

209. Corring DJ, Cook JV. Use of qualitative methods to explore the quality-of-life construct from a consumer perspective. *Psychiatr Serv* 2007;**58**:240–4. http://dx.doi.org/10.1176/appi.ps.58.2.240

210. Fisher MA, Mitchell GJ. Patients' views of quality of life: transforming the knowledge base of nursing. *Clin Nurse Spec* 1998;**12**:99–105.

211. Gee L, Pearce E, Jackson M. Quality of life in schizophrenia: a grounded theory approach. *Health Qual Life Outcomes* 2003;**1**:31.

212. Gould A, DeSouza S, Rebeiro-Gruhl KL. And then I lost that life: a shared narrative of four young men with schizophrenia. *Br J Occup Ther* 2005;**68**:467–73.

213. Hamer HP, McCallin AM, Garrett N. Searching for self: the layers and labels of panic disorder: a New Zealand study. *Nurs Health Sci* 2009;**11**:51–7. http://dx.doi.org/10.1111/j.1442-2018.2009.00430.x

214. Hedberg L, Skarsater I. The importance of health for persons with psychiatric disabilities. *J Psychiatr Ment Health Nurs* 2009;**16**:455–61. http://dx.doi.org/10.1111/j.1365-2850.2009.01400.x

215. Laliberte-Rudman D, Yu B, Scott E, Pajouhandeh P. Exploration of the perspectives of persons with schizophrenia regarding quality of life. *Am J Occup Ther* 2000;**54**:137–47. http://dx.doi.org/10.5014/ajot.54.2.137

216. Lim L, Nathan P, O'Brien-Malone A, Williams S. A qualitative approach to identifying psychosocial issues faced by bipolar patients. *J Nerv Ment Disord* 2004;**192**:810–17. http://dx.doi.org/10.1097/01.nmd.0000146734.39501.57

217. Mayers CA. Quality of life: priorities for people with enduring mental health problems. *Br J Occup Ther* 2000;**63**:591–6.

218. Michalak EE, Yatham LN, Kolesar S, Lam RW. Bipolar disorder and quality of life: a patient-centered perspective. *Qual Life Res* 2006;**15**:25–37. http://dx.doi.org/10.1007/s11136-005-0376-7

219. Rusner M, Carlsson G, Brunt D, Nystrom M. A dependence that empowers – the meaning of the conditions that enable a good life with bipolar disorder. *Int J Qual Stud Health Well-being* 2010;**5**(1). http://dx.doi.org/10.3402/qhw.v5i1.4653

220. Vallenga BA, Christenson J. Persistent and severely mentally ill clients' perceptions of their mental health. *Issues Ment Health Nurs* 1994;**15**:359–71.

221. Goffman E. *Stigma: Notes on the Management of Spoiled Identity*. London: Penguin Books; 1963.

222. Erikson EH. *Childhood and Society*. New York, NY: WW Norton; 1963.

223. Baumeister R, Leary M. The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychol Bull* 1995;**117**:497–529. http://dx.doi.org/10.1037//0033-2909.117.3.497

224. Anant SS. The need to belong. *Can Ment Health* 1966;**14**:21.

225. Hansson L. Determinants of quality of life in people with severe mental illness. *Acta Psychiatr Scand* 2006;**113**(Suppl. 429):46–50. http://dx.doi.org/10.1111/j.1600-0447.2005.00717.x

226. El-Badri S, Mellsop G. Stigma and quality of life as experienced by people with mental illness. *Australas Psychiatry* 2007;**15**:195–200. http://dx.doi.org/10.1080/10398560701320089

227. Newsom JT, Rook KS, Nishishiba M, Sorkin DH, Mahan TL. Understanding the relative importance of positive and negative social exchanges: examining specific domains and appraisals. *J Gerontol B Psychol Sci Soc Sci* 2005;**60**:304–12. http://dx.doi.org/10.1093/geronb/60.6.P304

228. Reis HT. The role of intimacy in interpersonal relations. *J Soc Clin Psychol* 1990;**9**:15–30. http://dx.doi.org/10.1521/jscp.1990.9.1.15

229. Nagle S, Cook JV, Polatajko HJ. I'm doing as much as I can: occupational choices of persons with a severe and persistent mental illness. *J Occup Sci* 2002;**9**:72–81. http://dx.doi.org/10.1080/14427591.2002.9686495

230. Romm KL, Rossberg JI, Hansen CF, Haug E, Andreasen OA, Melle I. Self-esteem is associated with premorbid adjustment and positive psychotic symptoms in early psychosis. *BMC Psychiatry* 2011;**11**:136. http://dx.doi.org/10.1186/1471-244X-11-136

231. Frank JD. Psychotherapy: the restoration of morals. *Am J Psychiatry* 1974;**131**:271–4.

232. Strada AE. Grief, demoralization and depression: diagnostic challenges and treatment modalities. *Prim Psychiatry* 2009;**16**:49–55.

233. Leamy M, Bird V, Le Boutillier C, Williams J, Slade M. A conceptual framework for personal recovery in mental health: systematic review and narrative synthesis. *Br J Psychiatry* 2011;**199**:445–52. http://dx.doi.org/10.1192/bjp.bp.110.083733

234. Anthony WA. Recovery from mental illness: the guiding vision of the mental health service system in the 1990s. *Psychosoc Rehabil J* 1993;**16**:11–23. http://dx.doi.org/10.1037/h0095655

235. Care Services Improvement Partnership, Royal College of Psychiatrists, Social Care Institute for Excellence. *A Common Purpose: Recovery in Future Mental Health Services*. Leeds: Social Care Institute for Excellence; 2007.

236. Tosh J, Brazier J, Evans P, Longworth L. A review of generic preference-based measures of health-related quality of life in visual disorders. *Value Health* 2012;**15**:118–27. http://dx.doi.org/10.1016/j.jval.2011.08.002

237. Qureshi H, Patmore C, Nicholas E. *Overview Outcomes of Social Care for Older People and Carers*. York: Social Policy Research Unit, University of York; 1998.

238. Huppert FA, Whttington JE. Evidence for the independence of positive and negative well-being: implications for quality of life assessment. *Br J Health Psychol* 2003;**8**:107–22. http://dx.doi.org/10.1348/135910703762879246

239. Bowling A. A taxonomy and overview of quality of life. In Brown J, Bowling A, Flynn T, editors. *Models of Quality of Life: A Taxonomy and Systematic Review of the Literature*. Sheffield: European Forum on Population Ageing Research Project, University of Sheffield; 2004.

240. Headey B, Kelley WJ, Wearing AJ. Dimensions of mental health: life satisfaction, positive affect, anxiety, and depression. *Soc Indic Res* 1993;**29**:63–82. http://dx.doi.org/10.1007/BF01136197

241. Deci E, Ryan RM. The 'what' and 'why' of goal pursuits: human needs and the self-determination of behavior. *Psychol Inq* 2000;**11**:319–38.

242. Maslow A. *Toward a Psychology of Being*. 2nd edn. New York, NY: Van Nostrand; 1968.

243. Keyes C, Shmotkin D, Ryff C. Optimizing wellbeing: the empirical encounter of two traditions. *J Pers Soc Psychol* 2002;**82**:1007–22.

244. Ryff CD. Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *J Pers Soc Psychol* 1989;**57**:1069–81. http://dx.doi.org/10.1037//0022-3514.57.6.1069

245. Waterman A. *The Psychology of Individualism*. New York, NY: Praeger; 1984.

246. Tennant A, McKenna SP, Hagell P. Application of Rasch analysis in the development and application of quality of life instruments. *Value Health* 2004;**7**(Suppl. 1):22–26. http://dx.doi.org/10.1111/j.1524-4733.2004.7s106.x

247. Longworth L, Yang Y, Young T, Mulhern B, Hernández Alava M, Mukuria C, *et al.* Use of generic and condition-specific measures of health-related quality of life in NICE decision-making: a systematic review, statistical modelling and survey. *Health Technol Assess* 2014;**18**(9).

248. Young T, Yang Y, Brazier J, Tsuchiya A. Use of Rasch analysis in reducing a large condition specific instrument for preference valuation: the case of moving from AQLQ to AQL-5D. *Med Decis Making* 2011;**31**:195–210. http://dx.doi.org/10.1177/0272989X10364846

249. Yang Y, Brazier J, Tsuchiya A, Young T. Estimating a preference-based index for a 5-dimensional health state classification for asthma (AQL-5D) derived from the Asthma Quality of Life Questionnaire (AQLG). *Med Decis Making* 2011;**31**:281–91.

250. Mulhern B, Smith SC, Rowen D, Brazier JE. Improving the measurement of QALYs in dementia: developing patient- and carer-reported health state classification systems using Rasch analysis. *Value Health* 2012;**15**:346–56.

251. Rowen D, Brazier J, Young T, Gaugris S, Craig BM, King MT, *et al.* Deriving a preference-based measure for cancer using the EORTC QLQ-C30. *Value Health* 2011;**14**:721–31. http://dx.doi.org/10.1016/j.jval.2011.01.004

252. Rowen D, Mulhern B, Banerjee S, Hout BV, Young T, Knapp M, *et al.* Estimating preference-based single index measures for dementia using DEMQOL and DEMQOL-Proxy. *Value Health* 2012;**15**:346–56. http://dx.doi.org/10.1016/j.jval.2011.10.016

253. Mavranezouli I, Brazier J, Rowen D, Barkham M. Estimating a preference-based index from the Clinical Outcomes in Routine Evaluation-Outcome Measure (CORE-OM): valuation of CORE-6D. *Med Decis Making* 2013;**33**:381–95. http://dx.doi.org/10.1177/0272989X12464431

Appendix 1 Study proposal

1. Title: Validating generic preference-based measures of health in mental health populations and estimating mapping functions for widely used specific measures

2. Importance

The last decade has seen the increased use of economic evaluation and particularly the use of cost effectiveness, where interventions are assessed in terms of their cost per Quality Adjusted Life Year (QALY). The QALY provides a way of measuring the benefits of health care interventions, including improvements in health related quality of life (HRQL). The most commonly used method for putting the 'q' into the QALY is to use a generic preference-based instrument such as the EQ-5D (Brook, 1996), as preferred by NICE. It is claimed that the EQ-5D and other 'generic' preference-based measures such as SF-6D (Brazier *et al.*, 2002) are applicable to all interventions and patient groups. This claim has support in many physical conditions where these instruments have managed to pass psychometric tests of reliability and validity (e.g. Marra *et al.*, 2006). For other conditions the claim has not be substantiated, such as in relation to visual impairment in Macular Degeneration (Espallargues et al, 2006) and hearing loss (Barton *et al.*, 2004).

For mental health the evidence is rather limited, but suggests a potentially mixed picture. There is evidence that generic instruments are able to reflect the impact of common conditions, such as mild to moderate depression and anxiety (Lamer *et al.*, 2006), while in a study of chronic schizophrenia using measures of psychopathology and functioning to establish change, the EQ-5D did not have a significant correlation with negative symptoms, disorganisation, depression, excitement and general symptoms (Willige *et al.*, 2005). The impact of a range of mental disorders on the generic preference-based SF-6D has been modelled by researchers in Sheffield using data from the 8,580 respondents to the ONS Psychiatric Morbidity Survey (Towers *et al.*, 2006). After adjusting for covariates, major anxiety disorders and depressive episodes were found to have a significant impact on SF-6D. However, obsessive compulsive disorder, personality disorder and probable psychosis were not significant.

Generic HRQL measures are rarely used in mental health research and this has been seen as a major limitation to conducting economic evaluations in mental health (Gilbody *et al.*, 2003). This is either because: 1) these measures are not regarded as valid in the patient group or 2) investigators are keen to keep patient burden down and prefer to use a more specific measure. There are many outcome measures specific to mental health and these make an important contribution to clinical research. The Department of Health's Expert Group on mental health outcomes identified the more widely used measures, including HoNOS, CORE-OM, GHQ, BDI, Lancashire Quality of Life Scale, CAN, FACE, MHI-5 (from the SF-36) and MANSA (DH, 2002). More recently the Improving Access to Psychological Therapies initiative has made a Minimum Data Set mandatory that includes the PHQ-9 and GAD-7 (for depression and anxiety respectively). However, these condition specific measures are not suitable for use in economic evaluation since they are not preference-based, that is they have not been scored with the values of the general public obtained using a recognised elicitation technique (as required by NICE and similar agencies around the world).

This proposal aims to address two issues: 1) to examine whether or not the low take-up of generic measures in mental health research is justified; and 2) to examine methods for overcoming the low use of generics by mapping relationships between widely used condition-specific measure and generic measures in order to permit evidence synthesis and cost-effectiveness analysis. This is important methodological ground work for conducting economic evaluation in mental health in a way that is consistent with other areas of health care and with the methods preferred by NICE. This will also provide important evidence on

the appropriateness of existing generic preference-based measures in mental health and whether it is necessary to adapt them in some way or develop a new preference-based measure in mental health.

3. Scientific potential

3.1 People and track record

Applicants: The team has the expertise required to undertake the study. John Brazier is a recognised international expert on the development of preference-based measures for use in economic evaluation. His work includes the development of the SF-6D and a number of condition specific measures. He currently has an MRC grant to develop a new method for mapping between measures of HRQL. Alicia O'Cathain, is a mixed methodologist with special expertise in qualitative methods and has worked in the area of quality of life research for many years. Michael Barkham is also a clinical psychologist with an international reputation in outcome measurement in mental health including the development of the CORE-OM (Barkham *et al.*, 2001). Myfanwy Lloyd Jones has extensive experience in conducting systematic reviews, including in the area of qualitative evidence. Donna Rowen has expertise in estimating mapping functions. Simon Gilbody is an expert in all aspects of evidence synthesis and mental health outcome measurement, and is Editor of the Cochrane Depression and Anxiety Group. He is also an Honorary Consultant psychiatrist and psychotherapist within the local NHS trust. David Richards has a background in nursing and behaviour therapy, and has been a leading figure in the expansion of provision of psychosocial care for common mental health problems in response to the Layard agenda. He has developed and implemented low intensity-high volume care in the National Improving Access to Psychological Therapies (IAPT) demonstration site in Doncaster. Sarah Byford and Martin Knapp are senior mental health economists with considerable experience with using the EQ-5D and SF-36 across a range of MH services.

Collaborators: Glens Parry has considerable research and policy experience and is leading on the SDO-funded evaluation of the DH demonstration sites for increasing access to psychological therapies. Eleni Chambers is a user researcher with expertise in advising on issues related to public and patient involvement at all stages of the research process, from pre-protocol planning to publications/dissemination. She is a member of various national and local networks and user groups. Tony O'Hagan is an Emeritus Professor of Bayesian Statistics who has worked with JB for many years on mapping these types of data and will provide advice on key statistical aspects of the mapping.

3.2 Environment

The project will be led from the School of Health and Related Research (ScHARR) at the University of Sheffield, but is very much collaboration with the University of York and Kings College London, Institute of Psychiatry. ScHARR is a large multi-disciplinary School within the Faculty of Medicine at the University of Sheffield with over 200 staff of social scientists, epidemiologists, statisticians, public health specialists and health economists. It provides a stimulating and supportive environment in which to conduct applied and methodological health services research. The health economics groups has developed a special expertise in the development and testing of outcome measures, including preference-based measures. The Mental Health Section provides the local clinical and academic support needed to conduct the study. ScHARR is able to provide the full range of facilities, including excellent library and information resources, office space and secretarial support.

3.3 Research Plans

3.3.1 Research Objectives

A previous proposal to develop a new preference-based measure in mental health was rejected by the MRC Board on the grounds that they did not believe there was sufficient evidence of the need for such a measure and they recommended pilot and feasibility work to establish this. Our proposal is partly aimed at addressing these suggestions, but more specifically it aims to address two issues: 1) to examine whether or not the low take-up of generic measures is justified and 2) to examine methods for overcoming the low

use of generics by mapping relationships between widely used condition specific measure and generic measures. Specific objectives are as follows:

- To establish the domains of quality of life that are important to people with mental health disorders from the literature and interviews with users
- To examine the extent to which EQ-5D and SF-36 (12) cover these domains
- To examine the psychometric properties of the EQ-5D and SF-36 across a range of mental health problems
- To estimate mapping functions between the measures commonly used in mental health trials and the EQ-5D and SF-6D (a derivative of the SF-36) and make these publicly available on the internet.

3.3.2 Plan of investigation

To address the two main aims of the study there will be two parts to the plan of investigation. The first part is concerned with testing the psychometric properties of the EQ-5D and SF-6D in mental health populations. The second part will be the estimation of mapping functions between commonly used measures in mental health and the generic measures (in those disorders where they seem to have acceptable properties). Part 1 uses a combination of qualitative and quantitative methods including: 1) a review of the qualitative evidence on the domains of quality of life raised by mental health service users; 2) qualitative interviews with a small number of users in order to examine the coverage and relevance of the generic measures and those raised in the literature; 3) review of the literature on the psychometric properties of these generic measures; and 4) a quantitative examination of the psychometric properties of the measures using existing data sets.

Study design

Part 1: Testing the properties of two generic measures in mental health

This will adopt the standard criteria for assessing measures recommended by Fitzpatrick *et al.* (1999) and adapted for preference-based measures by Brazier *et al.* (1999) that include feasibility, reliability, validity (content validity, face validity and construct validity) and responsiveness.

Feasibility is a problem general to all instruments, but some are more prone to non-response and missing data than others due to length, complexity and relevance to patients (McColl *et al.*, 2000). This has important implications for the power and validity of studies. Reliability over time is a desirable characteristic in any measure but it is strictly encapsulated in responsiveness. An unreliable instrument will not be very responsive. Internal reliability is an even less direct way at measuring the key property of responsiveness, though it can provide useful evidence of item functioning across different populations in different settings. Validity is the extent to which an instrument measures what it is intended to measure. In the absence of a gold standard measure for quality of life, various indirect ways of establishing validity will be used in this study including content and face validity, construct validity and concurrent or convergent validity (Streiner and Norman, 1989; Bowling, 2002). Content and face validation are best tested using qualitative methods based on interviews with the relevant population. Construct and current validation require data sets with indicators or measures of severity that can be used to assess group differences and degree of association.

Responsiveness is the ability of an instrument to measure clinically significant changes in health. Responsiveness is usually assessed statistically using measures such as the 'effect size', where the mean change in score is divided by either the standard deviation at baseline or the standard deviation of the change. A common assumption in the assessment of responsiveness is that for a given health change, the measure with the largest effect size is the better measure (Katz *et al.*, 1994). Where the objective is to minimise sample size this makes sense and maximises the ability to detect differences in the context of performance assessment. For economic evaluation this is less important, and instead it is simply a case of establishing that the descriptive content is able to detect change of significance to users. Related to this is precision which is

concerned with the ability of a measure to detect changes over the range experienced by patients being assessed. This requires the items and levels of the domains to be well spread over the range of the measure.

These criteria are going to be assessed using evidence from a range of different sources: the literature, interviews with users and a number of existing data sets. They will be applied to three broad groups of people with mental health problems: 1) common mental health problems (e.g. mild to moderate depression, anxiety, obsessive compulsive disorder and panic disorders); 2) severe and complex non-psychotic disorders (e.g. personality disorder); 3) schizophrenia and other psychotic disorders. These groups are not mutually exclusive, but are designed to help achieve a sample that covers the main groups of users. It is anticipated that much of the existing evidence will be by specific mental health disorder.

1) Review of qualitative evidence on dimensions of quality of life important to people with mental health disorders

Comprehensive systematic searches of the major relevant electronic bibliographic databases will be undertaken: these will include MEDLINE, PreMEDLINE, CINAHL and EMBASE. The reference lists of relevant identified articles and reviews will be searched for additional studies and authors and, where appropriate, citation searching will be conducted using CINAHL and the Web of Science. The search strategies will employ a combination of free-text and, where available, keyword searching. The terms to be used will include terms related to quality of life and mental health disorders, and terms or filters designed to identify qualitative research studies. The search strategies will be modified for different databases to take account of the thesaurus and limitations of each. The searches will not be restricted by date or language.

References identified by the literature searches will be screened for relevance in three stages, first by title, then by abstract, and finally by full text, excluding at each stage those which clearly do not satisfy the inclusion and exclusion criteria (Meade & Richardson 1997). Those studies which meet the inclusion criteria will then be appraised, and the data extracted, using a customised appraisal and data extraction form which draws on the primary research appraisal tool proposed by Paterson et al (Paterson *et al.* 2001) and on Sandelowski and Barroso's guide for reading qualitative studies (Sandelowski & Barroso 2002). The purpose of this appraisal is not to critique the quality of individual reports but to achieve an understanding of each study on its own terms (Sandelowski, Docherty, & Emden 1997), thus enabling consideration of the ways in which the methodology used has shaped understanding (Paterson *et al.* 2001). Studies will be excluded for reasons of quality only if the researcher's 'political' agenda is evident throughout, or if the depth and breadth of the reported data is insufficient to suggest that the findings are trustworthy (Paterson *et al.* 2001). Data will be analysed and synthesised using meta-study techniques developed by Paterson *et al.* (2001) facilitated by use of the Ritchie and Spencer's Framework approach (Ritchie and Spencer,1994).

2) Qualitative interviews with mental health service users

The aims of the interviews are: 1) to explore the dimensions found from the qualitative literature and 2) to explicitly examine whether users feel the generic measures cover the aspects of their quality of life affected by their condition. There will be around 24 semi-structured interviews conducted across the three broad categories of mental health disorders identified above. Sampling within each group will aim to achieve maximum variation until saturation of themes in the qualitative analysis is achieved. To this end, people will be recruited in a purposive manner from different settings, disorders and backgrounds. Sampling will be purposive to ensure all the main types of disorders in each group are covered. The numbers are small, but the aim is mainly to examine the extent to which the content of existing measures cover their concerns, rather than provide a definitive study of quality life in these groups of users.

Participants will be recruited in Sheffield from a number of different settings, specifically primary care for the common mental health problems, and the other two groups will be recruited from a combination of community mental health teams and inpatient services. Professionals responsible for the care of the users will be asked to approach them regarding their willingness to participate in the study and obtain their consent to be seen by an interviewer. These details will be sent to the interviewer at the University of Sheffield who will contact them to arrange the interview. Users in primary care will be identified by

considering those referred to brief psychological interventions within GP practices or to the primary care counselling services and by considering their medication. Secondary care users will be identified using the Sheffield patient database (known as 'Insight') which records service utilisation, demographic details and, in a number of services diagnoses and routinely collected assessment information.

Participants will be asked whether they are comfortable with the interview being recorded and transcribed verbatim, if not notes will be taken. A topic guide will be used to ensure that a common set of questions are asked about the person, including important background information, current activities, positive and negative influences on quality of life. The topic guide will be agreed by a project team (including user representation) based on the results of the qualitative review and the content of the EQ-5D and the SF-36. Interviews will be conducted sensitively, starting with general questions about their day to day lives. By the use of responsive questions the interviewer will probe those aspects of their quality of life identified and why these aspects are important. The interviews will be analysed using the 'framework' approach that helps to classify and progressively organise the data according to emergent themes (Ritchie and Lewis, 2003). The aim will be to identify a common set of dimensions across the groups and disorders. It will examine aspects of QL identified by the review and the two generic measures.

The use of in depth interviews to obtain people's views on what matters to their quality of life has become a key component in the development of quality of life measures in this field (Bowling, 2002; FDA, 2006). There is a growing body of literature on successful application of quantitative methods in various mental health populations, including schizophrenia (Gee *et al.*, 2003; Rose *et al.*, 2008; Michalak *et al.*, 2005).

3) Reviews of empirical studies using EQ-5D and SF-36 in mental health

Systematic searches of the main literature databases will be undertaken including MEDLINE, EMBASE, NHS Economic Evaluation Database (EED) and HTA, OHE Health Economic Evaluations Database (HEED), Quality of Life Database (QOLID), Cochrane Central Register of Controlled Trials, Science Citation Index (SCI), Social Sciences Citation Index (SSCI), Cochrane Database of Systematic Reviews (CDSR). The search terms will be based on those described for the qualitative review and various terms for SF-36, SF-12, SF-6D and EQ-5D. All papers searched will be sifted by title and then by abstract. All empirical studies using one or both of these measures in mental health populations will be included. Papers meeting the inclusion criteria will be reviewed in detail in order to extract data on the following for each health state value reported: disorder, severity of disorder, demographics of patients, measures, descriptive statistics for the values reported (mean, SD, median, IQ etc.) and evidence on validity and responsiveness.

4) Psychometric testing of the two instruments on existing data sets

To undertake new surveys of patients across the 3 groups would be difficult, complex and expensive. It would be hard to justify a new primary study, when there are a large and varied range of existing data sets available to the applicants from past trials and other studies that have not been exploited to address these methodological questions. The data sets currently available to the investigators have been listed in the Table below in terms of the study, sample (including size), and different instruments used. The review is expected to identify more data sets and access to these will be sought from the investigators.

The practicality of the instrument will be assessed in terms of the proportion of returned questionnaires and the proportion of questionnaires completed. Reliability will be examined in those data sets where there have been re-tests and evidence of no change in health status. Agreement will be examined in terms of their intra class correlation coefficient and Bland-Altman plots. Construct validation will be tested between groups with known differences defined using indicators of severity and mental health specific instruments. Testing will be undertaken on the preference-based index scores of the EQ-5D and SF-6D and also at the dimension levels. Statistical tests will include standard t-tests for the indices and K-W tests for the dimension levels. Responsiveness of the instrument will be examined in data sets that have some indicator of change. Changes in scores will be tested for statistical significance at the 5% level and the clinical importance of any difference will be assessed using the standard response mean (mean change over time divided by the standard deviation of the change).

Part II: Estimation of mapping functions between mental health specific measures and the generics

For those situations where the EQ-5D and/or SF-6D prove to be acceptable in terms of their psychometric properties, then the next stage in the project will be to estimate mapping functions between those mental health specific measures widely used in clinical studies onto the two generic measures. The PI recently completed a review of mapping functions (Brazier *et al.*, 2007) and will be drawing on the best practice identified in that review.

Regression analysis will be used to estimate the relationship between the preference-based scores of these two measures and the various mental health measures across different mental health disorders and conditions. One set of models will use the dependent variables of the EQ-5D and SF-6D indices. The explanatory variables will be the various mental health instruments entered into the models a variety of ways including dimension scores with main effects, squared terms and interaction terms. Models will also explore using item and item response variables. The inclusion of other covariates, including age and gender, will also be examined.

More complex specifications will include random effects generalised least squares (to allow for clustering in the data – usually around the individual), random effects (RE) tobit and censored least absolute deviations (CLAD) models. The RE tobit model is an appropriate alternative for estimating models using censored data (Sullivan and Ghushchyan, 2006) and as EQ-5D utility scores exhibit a ceiling effect, where a large proportion of subjects rate themselves in full health with a utility score of 1. CLAD will also used since its estimator does not depend upon assumptions of homoscedasticity and normality and hence produces consistent estimates in the presence of heteroscedasticity and non-normality (see Powell, 1984, and Sullivan and Ghushchyan, 2006). Another approach will be explored using EQ-5D and SF-6D dimensions as the dependent variables. A multinomial logit model will be used to estimate the probability that a respondent will choose a particular level for each dimension of each instrument (Gray *et al.* (2006)). Subsequently predicted EQ-5D level responses for each dimension are generated using Monte Carlo simulation methods and the corresponding preference-based index utility score for that health state is calculated using the relevant value sets. Bayesian approaches to modelling the data will also be developed drawing on work with modelling SF-6D using Gaussian approaches (Kharoubi *et al.*, 2007) that proved more flexible and better at predicting health state values.

Model performance will be mainly assessed in terms of their ability to predict EQ-5D/SF-6D index values in independent data sets, including the difference between predicted and observed values at either the aggregate health state level by calculating Mean Error (ME); or the individual level by calculating the Mean Absolute Error (MAE) or the Root Mean Squared Error (RMSE). Mean error will be examined by severity of health problem. The main purpose is to predict differences in values across sub-groups of patients, such as between arms of a trial or over time. So models will be examined in terms of the size and direction of error in their prediction of differences between sub-groups in independent data.

TABLE 1 Timetable of project (months)

Phase 1: Validation of measures	
Preparatory work – liaising with SCT and GP, research governance and ethics	1–3
I. Literature reviews	1–6
II. Qualitative interviews	6–15
III. Psychometric testing	1–7
IV. Mapping	8–15
Final report	15

4. Ethical and other implications

For the qualitative interviews it will be necessary to obtain informed consent. Researchers will only have access to potential participants' names after the clinical person responsible for their care has agreed and given consent. The research team will not have access to NHS databases of patients and their diagnoses (under Data Protection Act legislation). The interviewer will have experience with the client group, able to minimise any possible anxiety or distress during the interview process. During the interviews there will be an emphasis on gaining trust and building rapport. Interviews held in place of participant's choosing with someone present if they wish, positive feedback given at end of interview. User representatives ('Experts with experience') will help with preparing written materials and verbal explanations. At each interview researchers will check that the participant continues to give consent and understands that he or she can withdraw at any time. Data will be kept in line with the Data Protection Act. Names and addresses will be anonymised in the data base. Qualitative and quantative data will be reported in an anonymised form. Some mental health service users are at risk of relapse, self-harm or suicide and occasionally risk of harm to others. If such risks are revealed to researchers, the user's clinician will be informed. This potential need to break confidentiality will be clearly explained in the patient information sheet and consent form.

5. Public engagement

The proposed research has implications for the public in terms of its impact on policy decision-making processes in the Department of Health, NICE and the NHS. Service users have been involved in the development of this bid and will continue to be involved throughout the study. Users will also be on the Advisory Group. A summary of the findings will be advertised through user organisations. The results will be presented at a number of local workshops involving users up and down the country.

6. Exploitation and dissemination

We will offer to make presentations to NICE, Department of health and other interested public agencies on the results of this work. The qualitative work will be of interest in its own right to mental health services researchers and mental health professionals, so the results will be submitted to key national and international conferences. The results will also be written up for publication in academic journals. A key output of this part of the project will be a set of mapping functions programmed in standard software, including spss and excel, for researchers around the world to use and downloadable from the web. Dissemination will also be via user-focussed publications and through user networks.

7. References

Barton GR, Bankart J, Davis AC, *et al.* (2004) Comparing utility scores before and after hearing-aid provision. *Appl Health Econ Health Policy* vol. 3:(2), pp. 103–105.

Bower P, Byford S, Sibbald B, *et al.* Randomised controlled trial of non-directive counselling, cognitive-behaviour therapy, and usual general practitioner care for patients with depression. II Cost-effectiveness. *British Medical Journal*, 2000;321:1389–1392.

Bowling A (2002). *Measuring health: a review of quality of life and measurement scales*. Milton Keynes: OUP.

Brazier JE, Deverill (1999) A checklist for judging preference-based measures of health related quality of life: learning from psychometrics. *Health Economics* 8:41–51

Brazier J, Roberts J, Deverill M (2002). The estimation of a preference-based single index measure for health from the SF-36. *Journal of Health Economics* 21(2):271–292.

Byford S, Knapp M, Greenshields J, Ukoumunne OC, Jones V, Thompson S, Tyrer P, Schmidt U, Davidson K, Catalan J (on behalf of the POPMACT Group). Cost-effectiveness of brief cognitive behaviour therapy versus treatment as usual in recurrent deliberate self-harm: a rational decision making approach. *Psychological Medicine*, 2003, 33, 977–986.

Brooks R and EuroQol Group (1996). EuroQol: the current state of play. *Health Policy* 37:53–72

Savitha Eranti, Andrew Mogg, Graham Pluck, Sabine Landau, Rick Purvis, Richard Brown, Robert Howard, Martin Knapp, Michael Philpot, Sophia Rabe-Hesketh, Renee Romeo, John Rothwell, Denzil Edwards and Declan McLoughlin (2007) A randomised controlled trial with six-month follow-up of repetitive transcranial magnetic stimulation and electroconvulsive therapy for severe depression, *American Journal of Psychiatry*, 164, 73–81.

Espallargues M, Czoski-Murray C, Bansback N, *et al.* (2005) The impact of age related macular degeneration on health state utility values. *Investigative Ophthalmology and Visual Science* 46: 4016–4023.

Fitzpatrick R, Davey C, Buxton M, Jones D. Evaluating patient-based outcome measures for use in clinical trials. *Health Technology Assessment* 1998; 6(29).

Gee L, Pearce E, Jackson M. Quality of life in schizophrenia: a grounded theory approach. *Health and Quality of Life Outcomes* 2003,1:31

Goodyer I, Dubicka B, Wilkinson P, *et al.*

Gray, A. M., Rivero-Arias, O., Clarke, P. M., 2006. Estimating the Association between SF-12 Responses and EQ-5D Utility Values by Response Mapping. *Medical Decision Making* 26, 18–29.

Katz JN, Larson MG, Phillips CB, Fossel AH, Liang MH. Comparative measurement sensitivity of short and longer health status instruments. *Med Care* 1992;30:917–25.

Kharroubi SA, O'Hagan A, Brazier JE. Estimating utilities from individual health preference data: a nonparametric Bayesian method. *Applied Statistics* 2005, 54: 879–895.

King Fund (2006) Local Variations in NHS spending priorities – Briefing

Knapp M, Romeo R, Mogg A *et al.* (2008) Cost-effectiveness of transcranial magnetic stimulation vs. electroconvulsive therapy for severe depression: a multi-centre randomised controlled trial, *Journal of Affective Disorders*, forthcoming

Ritchie J, Lewis J. *Qualitative Research Practice*. London: Sage

Sainsbury Centre for mental health (2003) The economic and social costs of mental illness.

Kuyken W, Byford S, Watkins E, *et al.* Mindfulness-based cognitive therapy to prevent relapse in recurrent depression. *Journal of Consulting and Clinical Psychology*, 2008 (submitted)

Leech A, Rothwell J, White L, Harrington R. Selective serotonin reuptake inhibitors (SSRIs) and routine specialist care with and without cognitive behaviour therapy in adolescents with major depression: randomized controlled trial. *British Medical Journal*, 2007, 335, 142–146.

Lamers LM, Bouwmans CAM, van Straten A *et al.* (2006) Comparison of EQ-5D and SF-6D utilities in mental health. *Health Economics* (in press)

McColl E, Jacoby A, Thomas L, Soutter J, Bamford C, Steen N, Thomas R, Harvey E, Garratt A and Bond J. (2001). Design and use of questionnaires: a review of best practice applicable to surveys of health service staff and patients. *Health Technology Assessment*, 5(31), 1–256

Meade, M. O. & Richardson, W. S. Selecting and appraising studies for a systematic review. *Annals of Internal Medicine*. 1997; 127, 7:531–537.

Morrell J, Warner, Slade P. *et al.* Psychological Interventions for Postnatal Depression: Cluster Randomised Trial and Economic Evaluation. *Health Technology Assessment* (in press)

Michalak EE, Yatham LN, Kolesar S, Lam RW. Bipolar and quality of life: a patient centred perspective. *Quality of Life Research* 15(1):25–37.

Paterson, B. L., Thorne, S. E., Canam, C., & Jillings C., *Meta-study of qualitative health research*. 1 edn, Sage Publications, Inc., Thousand Oaks. 2001

Powell, J. L., 1994. Least Absolute Deviations Estimation for the Censored Regression Model. *Journal of Econometrics* 25, 303–325.

Ritchie J, Spencer L. Qualitative data analysis for applied policy research. In Bryman A, Burgess RG, eds. *Analysing qualitative data*, pp 173–94. London: Routledge, 1994.

Sandelowski, M., Docherty, S., & Emden, C. Qualitative metasynthesis: issues and techniques. *Research in Nursing & Health*. 1997; 20, 4:365–371.

Sullivan, P. W., Ghushchyan, V., 2006. Mapping the EQ-5D Index from the SF-12: US General Population Preferences in a Nationally Representative Sample. *Medical Decision Making* 26, 401–409.

Towers I, Brazier J, Walters S. *et al.* (2006) Modelling the impact of mental health disorders on health state values. 2006 International Society of Quality of Life Research meeting abstracts. [www.isoqol.org/2006mtgabstracts].

Tyrer P, Thompson S, Schmidt U, Jones V, Knapp M, *et al.* Randomised controlled trial of brief cognitive behaviour therapy versus treatment as usual in recurrent deliberate self-harm: the POPMACT study. *Psychological Medicine*, 2003, 33, 977–986.

US Dept. Health and Human Sciences. *Guidance for industry. Patient-reported outcomes measures: use in medical product development to support labelling claims*. Rockville: FDA.

Willige G, Wiersma D, Nienhuis FJ, *et al.* (2005) Changes in quality of life in chronic psychiatric patients: a comparison between EQ-5D and WHOQoL. *Qual Life Res.* 14:441–451.

Ward E, King M, Lloyd M *et al.* Randomised controlled trial of non-directive counselling, cognitive-behaviour therapy, and usual general practitioner care for patients with depression. I: Clinical effectiveness. *BMJ* 2000;321;1383–1388

TABLE 2 Data sets for examining the psychometric properties of the EQ-5D and SF-6D and estimating mapping functions

1) Common mental health problems	Patient population (n)	Instruments
Improving Access for Psychological Therapy in Doncaster; Richards *et al.* (2008)	Depression, Anxiety etc. n>3000	SF-12, EQ-5D PHQ-9, GAD-7, CORE-OM
McCabe *et al.* (1996)	General population n = 1896	SF-36 (SF-6D) GHQ-12
PONDER trial. Morrell *et al.* (2008)	Mothers at 6 and 12 months post-partum n = 2600 and 1700	SF-12, EPDS, CORE-OM
Mindfulness-based cognitive therapy to prevent relapse in recurrent depression (Kuyken *et al.*, submitted)	Depression (n = 123)	EQ-5D, HRSD, BDI, WHO-brief QoL
SSRIs and routine care with and without CBT in adolescents with major depression (Goodyer *et al.*, 2007)	Depression (n = 208)	HoNOS,CA, EQ-5D, MFQ, CDRS-R, CGAS, CGI-I
RCT of counselling, CBT and usual GP care for patients with depression. Ward *et al.* (2000)	Depression (n = 197)	EQ-5D, BDI, Brief symptom inventory, modified social adjustment scale
2) Severe and complex non-psychotic		
RCT of brief CBT versus treatment as usual in recurrent deliberate self-harm: the POPMACT study. Tyrer *et al.* (2003)	Recurrent deliberate self-harm (n-480)	EQ-5D, HADS, GAF, SFQ, PAS
Patient attending a centre in the Netherlands Soeteman *et al.* (2008)	Patients with personality disorder n = 1708	EQ-5D
TMSECT: treatment of people with severe depression with ECT or transcranial magnetic simulation. Eranti *et al.* (2007); Knapp *et al.* (2008)	Patients referred for RCT with severe depressive episodes (bipolar or unipolar) n = 46	HRSD, BDI, BPRS, CAMCOG, MMSE, SF-36 (SF-6D)
3) Schizophrenia and other psychotic problems		
SCAP (Lilly). Mangalore *et al.* (2006)	Schizophrenia n = 600	EQ-5D, SF-12 PANSS, MADRS, GAF, QLS, AIMS
QUATRO.Gray *et al.* (2006)	Schizophrenia n = 409	EQ-5D, SF-36, MANSA, BPRS, SAI, MAQ,

Appendix 2 *Chapter 2* appendices

Quantitative review search strategy

1. (sf36 or sf 36 or short form 36 or shortform 36 or sf thirtysix or sf thirty six or shortform thirtysix or shortform thirty six or short form thirtysix or short form thirty six).tw. (9816)
2. (sf6 or sf 6 or short form 6 or shortform 6 or sf six or sfsix or shortform six or short form six).tw. (976)
3. (sf12 or sf 12 or short form 12 or shortform 12 or sf twelve or sftwelve or shortform twelve or short form twelve).tw. (1373)
4. (euroqol or euro qol or eq5d or eq 5d).tw. (1748)
5. exp Mental Disorders/ (748,722)
6. exp mood disorders/ or exp affective disorders, psychotic/ or exp bipolar disorder/ or exp depressive disorder/ (88,818)
7. exp Schizophrenia/ (70,221)
8. exp self-injurious behavior/ or exp suicide/ (43,194)
9. exp personality disorders/ (26990)
10. exp Eating Disorders/ (17,636)
11. Stress Disorders, Post-Traumatic/ (14,172)
12. Gambling/ (2042)
13. exp Anxiety Disorders/ (50,236)
14. exp "sexual and gender disorders"/ or exp sexual dysfunctions, psychological/ or exp paraphilias/ (23,682)
15. exp Sleep Disorders/ (46,193)
16. ((mental or psychiatric) adj2 (disorder$ or condition$ or disease$)).tw. (39,628)
17. or/6–16 (350,409)
18. or/1–4 (13,357)
19. 17 and 18 (1155)
20. limit 19 to english language (1082)

Schizophrenia review studies

TABLE 32 Schizophrenia review studies

Study details	Population characteristics	Properties measured
EQ-5D validity and responsiveness		
Auquier (2003)[84]	DSM-IV schizophrenia	Convergent validity
France	Inpatients and outpatients (numbers not reported)	
	n = 207 (141 males and 66 females)	
	Mean age = 37.3 (SD = 10.9) years (range 18–70 years)	
Badia (2000)[87]	Schizophrenia (classification not reported)	Responsiveness
Country not reported	*N* = approx. 2949 (*n* = 2128 on olanzapine; *n* = 821 on risperidone or haloperidol; small numbers on other antipsychotics)	
	No age, sex or inpatient/outpatient status reported	
Barton (2009)[89]	Non-affective psychosis diagnosis (criteria not specified). Includes schizophrenia, schizoaffective disorder, bipolar disorder and psychotic depression	Known-group validity
UK		Convergent validity
	Participants had to screen positive for psychotic symptoms and in relative remission ($\leq$ 4 on PANSS)	Responsiveness
	n = 77 (55 males, 22 females)	
	Mean age = 28.9 years (range 18–52 years) 50/77 had a diagnosis of non-affective psychosis	
	Inpatient/outpatient status not reported	
Konig (2007)[90]	ICD-10: schizophrenia, schizotypal or delusional disorders	Convergent validity
Germany	49.4% outpatient, 41.6% inpatient and 9.0% day clinic	
	n = 166 (97 males, 69 females)	
	Mean age = 40.5 (SD = 11.1) years (range 21–80 years)	
Konig (2009)[91]	ICD-10: schizophrenia, schizotypal or delusional disorders	Convergent validity
Germany	51.7% outpatient, 38.5% inpatient and 9.8% day clinic	
	n = 166 (97 males, 69 females)	
	Mean age = 40.4 (SD = 11.6)	
McCrone (2009)[68]	SCAN interview diagnosed schizophrenia (classification scheme not specified)	Convergent validity
Netherlands, Germany, UK and Italy		Responsiveness
	'Chronic high disability sample' based on number of years on medication, number of psychiatric inpatient days last year and GAF score	
	n = 409 (245 males and 164 females)	
	Mean age = 41.5 (SD = 11.5) years (no range reported)	

TABLE 32 Schizophrenia review studies (*continued*)

Study details	Population characteristics	Properties measured
Prieto (2004)[86] Spain	ICD-10 schizophrenia $n = 2657$ (1691 males, 966 females) Not stated if inpatient or outpatient $n = 2128$ on olanzapine; $n = 417$ on risperidone; $n = 112$ on haloperidol Mean age = 35.32 (SD = 11.57) years (range not reported)	Convergent validity
Scalone (2008)[92] Italy	$N = 637$ ($n = 551$ with schizophrenia; $n = 86$ with schizophreniform disorder) 414 males, 223 females Aged 18 to 40 years (no mean age reported) Inpatient/outpatient status not reported	Convergent validity
van de Willige (2005)[48] Netherlands	DSM-IV schizophrenia (described as chronic sample) Auditory hallucinations for > 2 years after adequate treatment Use of at least two antipsychotic drugs Inpatients and outpatients – numbers not reported $n = 76$ (42 males, 34 females) Mean age = 36 (SD = 11.2) years	Responsiveness
SF-36		
Auquier (2003)[84] France	DSM-IV schizophrenia inpatients and outpatients (numbers not reported) $n = 207$ (141 males, 68 females) Mean age = 37.3 (SD = 10.9) years (range 18–70 years)	Convergent validity
Bebbington (2009)[67] UK, France and Germany	DSM-IV schizophrenia $n = 1208$ (743 males, 465 females) Aged 18 to 64 years (no mean age reported) Outpatients all on antipsychotic medication (type not specified)	Known-group validity
Bobes (1997)[85] Spain	ICD-10 schizophrenia $n = 60$ All outpatients measured by PANSS as out of acute phase; on maintenance treatment with different neuroleptics	Convergent validity

TABLE 32 Schizophrenia review studies (*continued*)

Study details	Population characteristics	Properties measured
Dunayevich (2007)[93]	DSM schizophrenia	Responsiveness
	n = 1278	
Folsom (2009)[77]	DSM-IV schizophrenia or schizoaffective disorder	Known-group validity
USA	*n* = 486 outpatients (317 males, 169 females)	
	n = 164 (42.3%) on typical antipsychotics; *n* = 198 (51.1%) on atypical antipsychotics Mean age = 54.7 (SD = 8.6) years (no age range given)	
Kebede (2004)[74]	ICD-10 schizophrenia	Known-group validity
Ethiopia	*n* = 307 outpatients (253 males, 54 females)	
	Mean age and age range not reported	
Kebede (2005)[75]	*n* = 271/321 available for follow-up analysis (4 years)	Known-group validity
Ethiopia	Outpatients	Responsiveness
	220 males, 51 females	
	208 defined as being long-standing cases (duration of illness between 3 and 30 years) and 63 defined as having recent-onset illness (onset within the last 2 years)	
	Mean age = 30.8 (SE = 0.5) years (age range not given)	
Law (2005)[80]	ICD-10 schizophrenia. Untreated psychosis – participants experiencing first episode	Known-group validity
Hong Kong		Convergent validity
	Inpatients and outpatients (mostly outpatients, numbers not reported)	
	n = 117 (63 males, 54 females)	
	Mean age = 20.3 (SD = 3.6) years (range 14–28 years)	
Lenert (2005)[66]	DSM-IV schizophrenia or schizoaffective disorder	Convergent validity
Europe and Canada	*N* = 725 completed the trial (*n* = 615 schizophrenia; *n* = 110 schizoaffective disorder)	
	312 males and 164 females completed the trial	
Meijer (2002)[81]	ICD-10 schizophrenia/SCAN-diagnosed schizophrenia	Convergent validity
Netherlands	*n* = 143	
	Aged 18–65 years. No mean age or sex data reported	
	Inpatients and outpatients (numbers not reported)	

TABLE 32 Schizophrenia review studies (*continued*)

Study details	Population characteristics	Properties measured
Milliken (2007)[83] Canada	Schizophrenia spectrum disorder (classification not specified) Treatment: 6 months' exposure to antipsychotic medication Inpatients/outpatients status not reported n = 15 (first patients with complete SF-36 data) No age or sex data reported	Convergent validity Responsiveness
Nasrallah (2004)[72] USA	DSM-IV schizophrenia Treatment: long-acting injection and oral risperidone vs. oral risperidone n = 369 (275 males, 94 females) Inpatients or outpatients (numbers not reported) Mean age across four treatment groups ranged between 36.0 (SD = 1.0) and 39.0 (SD = 1.1) years	Known-group validity
Norholm (2007)[76] Denmark	ICD-10 schizophrenia n = 56 (23 males, 33 females) Inpatients – during hospitalisation. Usual treatment including antipsychotic medication, psychosocial support and a rehabilitation programme Mean age = 39 (SD = 10.9) years (range 21–65 years) Compared with an age-matched sample of Danish population (n = 2650)	Known-group validity
Phillips (2006)[94] USA	DSM-IV schizophrenia treatment; olanzapine vs. ziprasidone N = 548 (n = 277 assigned to olanzapine and n = 271 assigned to ziprasidone) Mean age in olanzapine groups = 40.1 years and mean age in ziprasidone group = 38.2 years No data on sex reported Inpatients and outpatients (numbers not reported)	Responsiveness
Pukrop (2003)[79] Germany	Inpatients n = 91 (47 males, 44 females) Mean age = 33.73 (SD = 9.67) years	Known-group validity Convergent validity Responsiveness

TABLE 32 Schizophrenia review studies (*continued*)

Study details	Population characteristics	Properties measured
Pyne (2003)[95] USA	DSM-IV schizophrenia or schizoaffective disorder N = 134 (n = 85 schizophrenia; n = 49 schizoaffective disorder) Inpatients and outpatients (numbers not reported) 126 males, 8 females Mean age = 46.8 (SD = 8.1) years (no range reported)	Responsiveness
Reine (2005)[88] France	DSM-IV schizophrenia 36% outpatients and 64% inpatients Phase: first episode (7%); readmission period for an acute phase (31%); stabilised (62%) N = 205 (139 males, 66 females) Mean age = 37.4 (SD = 10.9) years (range 18–70 years)	Convergent validity Responsiveness
Revicki (1999)[69] Austria, Belgium, France, Germany, Italy, Poland, Spain, USA, Canada and UK	DSM-III-R schizophrenia, schizophreniform disorder and schizoaffective disorder Inpatients or outpatients (numbers not reported). BPRS score ≤ 18 Treatment: olanzapine vs. haloperidol (RCT) n = 1159 (787 olanzapine patients and 372 haloperidol patients) Age and sex not reported for this group	Convergent validity Responsiveness
Russo (1998)[73] USA	DSM-III-R schizophrenia or schizoaffective disorder Outpatients n = 36 (25 males, 11 females) Mean age = 40.0 (SD = 8.07) years (range 19–53 years)	Known-group validity Convergent validity
Scalone (2008)[92] Italy	N = 637 (n = 551 with schizophrenia; n = 86 with schizophreniform disorder) 414 males, 223 females Aged 18 to 40 years (no mean age reported) Inpatient/outpatient status not reported	Convergent validity
Sciolla (2003)[78] USA	DSM-IV schizophrenia or schizoaffective disorder. 'Medically, psychologically and pharmacologically stable' Mostly outpatients (numbers not reported) n = 137 (91 males, 46 females); mean age = 57.9 (SD = 8.9) years n = 77 'normal' participants (30 males, 47 females); mean age = 66 (SD = 10.6) years	Known-group validity

TABLE 32 Schizophrenia review studies (*continued*)

Study details	Population characteristics	Properties measured
Strakowski (2005)[70] USA and Western Europe	DSM-IV schizophrenia, schizoaffective disorder and schizophreniform disorder First episode of psychosis-active psychotic symptoms (at least two PANSS psychosis items ≥ 4 or one psychosis item ≥ 5, and a CGI score ≥ 4) Inpatient/outpatient status not reported RCT olanzapine vs. haloperidol n = 195 (156 males, 39 females) Mean age = 24 (SD = 5) years (no range reported) Normative data were taken from published tables (Ware JE, Jr, Kosinski M, Keller SD. *SF-36 Physical and Mental Health Summary Scales: A User's Manual*. Boston, MA: Health Assessment Lab; 1994) and were weighted to the age and sex of the study sample	Known-group validity
Tunis (1999)[71] USA, UK and Canada	DSM-III-R schizophrenia, schizophreniform disorder or schizoaffective disorder. Experiencing 'clinically significant psychosis' Inpatients and outpatients (numbers not reported) n = 1155 (802 males, 353 females) Mean age = 39.29 (SD = 11.32) years (range not reported) n = 935 schizophrenia and n = 220 schizoaffective disorder	Known-group validity Convergent validity
Wilkinson (2000)[82] UK	Schizophrenia (unknown classification) Clinic and home-based settings n = 78 (male/female not reported) Mean age = 40 (SD = 11.9) years (range 18–64 years) Inpatient/outpatient status not reported	Convergent validity

CGI, Clinical Global Impression; DSM-III-R, *Diagnostic and Statistical Manual of Mental Disorders* – Third Edition-Revised; SCAN, Schedules for Clinical Assessment in Neuropsychiatry; SE, standard error.

Bipolar disorder review studies

TABLE 33 Bipolar disorder review studies

Study details	Population characteristics	Properties measured
EQ-5D		
Arbuckle (2009)[98] USA	Bipolar disorder – type I and II $N = 225$ Type I: $n = 150$ Type II: $n = 31$ Missing data: $n = 31$ Mean age = 41.42 (SD = 11.49) years 75 males and 149 females	Convergent validity
Hayhurst (2006)[99] UK	Bipolar disorder – type not specified $n = 221/253$ completed the study Mean age = 41.5 (SD = 11.0) years 79 males and 142 females	Known-group validity Convergent validity
Hakkaart-van Roijen (2004)[100] Netherlands	Bipolar disorder – type I, II, NOS, cyclothymia and substance-induced $N = 40$ Bipolar disorder type I: $n = 14$ Bipolar disorder type II: $n = 14$ Bipolar disorder NOS: $n = 7$ Cyclothymia: $n = 3$ Bipolar disorder, substance-induced: $n = 2$ Mean age = 43.84 (SD = 10.88) years 15 males and 25 females	Known-group validity
Vojta (2001)[101] USA	Bipolar disorder – type not specified $n = 86$ (70 males and 14 females) All completed Mean age = 47.3 (SD = 6.9) years	Known-group validity Convergent validity
SF-36		
Albert (2008)[97] Italy	Bipolar disorder – type I and II $N = 105$ $n = 44$ with bipolar I; $n = 61$ with bipolar II Mean age = 47.33 (SD = 12.25) years 45 males and 60 females	Known-group validity

TABLE 33 Bipolar disorder review studies (*continued*)

Study details	Population characteristics	Properties measured
	Comorbid anxiety disorders: current; lifetime At least one disorder: $n = 34$; $n = 43$	
	Panic disorder (with or without agoraphobia): $n = 5$; $n = 11$	
	Social phobia: $n = 3$; $n = 7$	
	OCD: $n = 13$; $n = 14$	
	GAD: $n = 16$; $n = 17$	
Arbuckle (2009)[98] USA	Bipolar disorder – type I and II $N = 225$ Type I: $n = 150$ Type II: $n = 31$ Missing data: $n = 31$ Mean age = 41.42 (SD = 11.49) years 75 males and 149 females	Known-group validity Convergent validity
Arnold (2000)[102] USA	Bipolar disorder – type I, II and NOS $N = 44$ Type I: $n = 38$ Type II: $n = 5$ NOS: $n = 1$ Mean age = 41.8 (SD = 9.8) years 15 males and 39 females	Known-group validity
Bauer (2005)[103] USA	Bipolar disorder – type I and II $n = 328$ Mean age = 46.6 (SD = 10.1) years Female/male not reported *Anxiety disorders: current; lifetime* Any: $n = 126$; $n = 142$ PTSD: $n = 85$; $n = 93$ Panic disorder: $n = 56$; $n = 65$ Panic disorder with agoraphobia: $n = 37$; $n = 44$ Panic disorder without agoraphobia: $n = 19$; $n = 21$ OCD: $n = 28$; $n = 36$ GAD: $n = 0$; $n = 0$	Known-group validity

TABLE 33 Bipolar disorder review studies (*continued*)

Study details	Population characteristics	Properties measured
Cuijpers (2009)[104] Netherlands	Bipolar disorder – type not specified *n* = 43 Age and sex not reported	Known-group validity
Depp (2006)[105] USA	Bipolar disorder – type not specified *n* = 54 bipolar individuals (30/54 had SF-36 data as this measure was introduced later in the study); mean age = 57.6 (SD = 9.2) years; 34 males and 20 females *n* = 38 healthy comparison subjects; mean age = 64.7 (SD = 12.7) years; 9 males and 29 females	Known-group validity Convergent validity
Depp (2009)[106] USA	Bipolar disorder – type I and II *n* = 30 bipolar individuals; mean age = 60.2 (SD = 11.2) years; 16 males and 14 females *n* = 31 healthy comparison subjects; mean age = 59.7 (SD = 9.1) years; 9 males and 22 females	Known-group validity Convergent validity
Gutierrez-Rojas (2008)[107] Spain	Bipolar disorder – type I, II and cyclothymic *N* = 108 Type I: *n* = 80 Type II: *n* = 22 Cyclothymic: *n* = 6 Mean age = 48 (SD = 14) years 33 males and 75 females	Known-group validity Convergent validity
Kebede (2006)[108] Ethiopia	Bipolar disorder (not type II disorder) *n* = 264 [69 recent-onset (within the last 2 years) cases and 195 long-standing cases (> 2 years)] Mean age = 29.3 (SD = 0.5) years 143 males and 121 females	Known-group validity Responsiveness
Kebede (2004)[74] Ethiopia	Bipolar disorder – type not specified *n* = 293 Mean age and age range not reported 176 males and 117 females	Known-group validity
Leidy (1998)[109] USA	Type I bipolar disorder Euthymic patients: mean age = 42.5 (SD = 12.8) years; 14 males and 20 females Depressed patients: mean age = 38.5 (SD = 9.9) years; 14 males and 14 females	Known-group validity Convergent validity Responsiveness

TABLE 33 Bipolar disorder review studies (*continued*)

Study details	Population characteristics	Properties measured
Maina (2007)[110] Italy	Bipolar disorder – type I and II $N = 142$ $n = 90$ bipolar I: mean age = 41.04 (SD = 11.80) years; 43 males and 47 females $n = 52$ bipolar II: mean age = 47.92 (SD = 10.76) years; 21 males and 31 females	Known-group validity
Namjoshi (2002)[111] USA	Bipolar disorder – type not specified $N = 139$ Olanzapine group, $n = 70$; placebo group, $n = 69$ Mean age: olanzapine group = 40.2 (SD = 11.6) years; placebo group = 38.7 (SD = 10.3) years 72 males and 67 females	Responsiveness
Shi (2004)[112] Australia, Colombia, Greece, Mexico, Portugal, Spain and USA	Type I bipolar disorder $N = 573/833$ completed the trial Baseline data for $N = 573$ who completed are provided: Olanzapine group ($n = 250$): mean age = 42.0 (SD = 12.4) years; 85 males and 165 females Olanzapine–fluoxetine group ($n = 58$): mean age = 38.4 (SD = 12.4) years; 18 males and 40 females Placebo group ($n = 265$): mean age = 40.6 (SD = 12.5) years; 98 males and 167 females	Convergent validity
Shi (2002)[113] Country not reported	Type I bipolar disorder $n = 453$ enrolled Demographics provided for 298 patients included in analyses: Olanzapine: mean age = 40.3 (SD = 12.8) years; 62 males and 99 females Haloperidol: mean age = 38.0 (13.1) years; 55 males and 82 females	Responsiveness
Sierra (2005)[114] Spain	Bipolar disorder: type I and II $N = 50$ Type I: $n = 45$ Type II: $n = 5$ Mean age = 45.14 (SD = 12.9) years 20 males and 30 females	Known-group validity

TABLE 33 Bipolar disorder review studies (*continued*)

Study details	Population characteristics	Properties measured
Simon (2007)[115] USA	Bipolar disorder: type I and II $N = 441$ Type I: $n = 336$ Type II: $n = 105$ Mean age = 44 years 137 males and 304 females	Known-group validity Responsiveness
Yatham (2004)[116] International – 15 countries	Type I bipolar disorder $n = 520/598$ completed the SF-36 Reasons for non-completion were not reported Mean age = 42 years (SD not reported) 374 males and 584 females	Known-group validity Convergent validity
Zhang (2006)[117] USA	Bipolar disorder – type not specified $n = 1999$ Mean age = 40.6 (SD = 12.8) years 842 males and 1157 females	Known-group validity
SF-12		
Vojta (2001)[101] USA	Bipolar disorder – type not specified $n = 86$ All completed Mean age = 47.3 (SD = 6.9) years (no range reported) 70 males and 14 females	Known-group validity Convergent validity

GAD, generalised anxiety disorder; NOS, not otherwise specified; OCD, obsessive compulsive disorder.

Personality disorders review studies

TABLE 34 Personality disorders review studies

Study details	Population characteristics	Properties measured
SF-36		
Hueston (1996)[121] USA	Patients 'at risk' of PDs: determined by completion of the SCID for DSM-III-R for personality disorders $N = 93$ $n = 65$ patients at high risk of one or more PDs; mean age = 44.7 (SD = 15.3) years; 17 males and 48 females $n = 28$ patients at low risk of PDs; mean age = 39.7 (SD = 15.1) years; 6 males and 22 females	Known-group validity
Narud (2005)[119] Norway	Personality disorders $n = 91$ (43 males and 48 females) Mean age = 36.6 (SD = 10.5) years (range 19–74 years)	Known-group validity Convergent validity Responsiveness
SF-12		
Jackson (2000)[122]/Jackson (2002)[123] Australia	$n = 10,641$ 4705 males and 5936 females Age data not reported $n = 704$ (319 males and 385 females) had at least one PD	Known-group validity
Sanderson (2002)[96] Australia	Sample taken from Australian National Survey of Mental Health and Well-being ($n = 10,641$) – a nationally representative household survey of mental disorders in adults 5214 males and 5427 females Participants in survey mean age = 45 years No mental disorder: $n = 9902$ Any personality disorder: $n = 564$	Known-group validity
EQ-5D		
Bartak (2011a)[124] Netherlands	DSM–IV-TR axis II cluster A personality disorders $n = 57$ 17 males and 40 females Mean age = 29.4 (SD 8.2) years **PD diagnosis** 'Pure' cluster A, $n = 9$; cluster A and B, $n = 7$; cluster A and C, $n = 18$; cluster A, B and C, $n = 23$ Paranoid PD, $n = 49$; schizoid PD, $n = 5$; schizotypal PD, $n = 4$	Responsiveness

TABLE 34 Personality disorders review studies (*continued*)

Study details	Population characteristics	Properties measured
	All received psychotherapeutic treatment in one of three settings: Outpatient, $n = 20$	
	Day hospital, $n = 19$	
	Inpatient, $n = 18$	
Bartak (2011b)[125]	DSM–IV-TR axis II cluster B personality disorders	Responsiveness
Netherlands	$n = 207$	
	70 males and 147 females	
	Mean age = 31.3 (SD = 8.5) years	
	PD diagnosis Pure cluster B, $n = 84$; cluster B and C, $n = 93$; cluster B and A, $n = 7$; cluster A, B and C, $n = 23$	
	Borderline PD, $n = 160$; narcissistic PD, $n = 47$; histrionic PD, $n = 26$; antisocial PD, $n = 18$	
	All received psychotherapeutic intervention in one of three settings: Outpatient, $n = 46$	
	Day hospital, $n = 81$	
	Inpatient, $n = 80$	
Bartak (2010)[126]	DSM–IV-TR axis II cluster C personality disorders	Responsiveness
Netherlands	$n = 371$	
	110 males and 261 females	
	Mean age = 33.5 (SD = 9.5) years	
	PD diagnosis Pure cluster C, $n = 247$; cluster C and B, $n = 88$; cluster C and A, $n = 15$; cluster C, B and A, $n = 21$	
	Avoidant PD, $n = 235$; obsessive-compulsive PD, $n = 183$; dependent PD, $n = 84$	
	All received one of five modalities of psychotherapeutic interventions: Long outpatient, $n = 68$	
	Short day hospital, $n = 77$	
	Long day hospital, $n = 74$	
	Short inpatient, $n = 59$;	
	Long inpatient, $n = 93$	
Soeteman (2005)[127]	Complex personality problems or personality disorders	Known-group validity
Netherlands	$n = 1651$	
	541 males and 1110 females	Convergent validity
	Mean age = 31.9 (SD = 9.2) years (range 18–61 years)	

TABLE 34 Personality disorders review studies (*continued*)

Study details	Population characteristics	Properties measured
Soeteman (2008)[120]	*n* = 1708	Known-group validity
Netherlands	605 males and 1103 females	
	Mean age = 33.7 (SD = 9.9) years (range 18–67 years)	
Van Asselt (2009)[128]	Borderline personality disorder	Responsiveness
Netherlands	*n* = 48	
	5 males and 43 females	
	Mean age of completers was 31 (SD = 8.55) years	

DSM-IV-TR, *Diagnostic and Statistical Manual of Mental Disorders* – Fourth Edition-Text Revision; PD, personality disorder; SCID, Structured Clinical Interview for DSM Disorders.

Anxiety and depression review studies

TABLE 35 Anxiety and depression review studies

Study details	Population characteristics	Properties measured
Validity and responsiveness – EQ-5D and SF-6D		
Aydemir (2009)[129] Turkey	74 patients, aged 18–65 years, diagnosed major depressive episode according to DSM-IV criteria	Known-group validity
	Exclusions: other psychiatric disorder, comorbid condition	Convergent validity
	Mean age = 39.6 years, 63.5% female	
	32.4% recurrent depression	
Bosmans (2008)[130] Netherlands	Patients with minor or mild-major depression in primary care	Responsiveness
	Exclusions: currently receiving antidepressants or psychological therapy	
	n = 44 usual care no antidepressants; mean age = 48 years; 73% female	
	n = 45 usual care antidepressant; mean age = 46 years; 76% female	
Caruso (2010)[131] Italy	*n* = 513 patients in primary care, with clinically diagnosed episode of depression requiring pharmacological treatment	Responsiveness
	Mean age = 49.2 years, 72.9% female	
	Data recorded at baseline and at 3 and 6 months	
Ergun (2007)[132] Turkey	74 patients with MDD in RCT	Convergent validity Responsiveness
Fernandez (2005)[133] UK	293 outpatients (aged 18–85 years) fulfilling DSM-IV criteria for severe MDD, without suicidal tendencies	Responsiveness
	Exclusions: history mania, bipolar, schizophrenia or psychotic disorder, evidence of obsessive compulsive disorder, eating disorder, mental retardation, pervasive development disorder	
	n = 126 assigned to escitalopram; mean age = 48.4 years; 75.4% female	
	n = 125 assigned to venlafaxine; mean age = 46.5 years; 71.2% female	
Gunther (2008)[134] Germany	Patients with a depressive episode according to the ICD-10	Convergent validity Responsiveness
Konig (2009)[135] Germany	389 patients with anxiety disorder from 46 GP practices	Responsiveness
	23 practices allocated to intervention (training) group (*n* = 201), and 23 to control group (usual care) (*n* = 213)	
Konig (2010)[136] Germany	389 patients with anxiety disorder	Known-group validity Convergent validity Responsiveness

TABLE 35 Anxiety and depression review studies (*continued*)

Study details	Population characteristics	Properties measured
Lamers (2006)[50] Netherlands	Patients aged 18–65 years with a diagnosis of a major depressive disorder, dysthymic disorder, panic disorder, social phobia, generalised anxiety (classified by DSM-IV)	Known-group validity Responsiveness
Mann (2009)[137] UK	114 patients with depression participating in RCT on collaborative care (MDD according to SCID) Mean age = 42.5 years, 77% female	Known-group validity Convergent validity Responsiveness
Mychaskiw[138] (2008) USA	374 non-depressed patients with GAD treated with pregabalin, venlafaxine-XR or placebo *n* at endpoint ranged from 323 to 360	Known-group validity Responsiveness
Petrou (2009)[139] UK	623 women in a RCT of postnatal support. 1046 declined. 312 allocated to control group receiving usual care. 311 offered 10 visits by a community postnatal support worker. 493 had complete data Patients aged ≥ 17 years who had given birth to a live baby	Known-group validity
Peveler (2005)[140] UK	Of 388 patients with a new episode of depression referred to the study, 67.3% female, mean age 42.5 years *n* = 327 randomised	Responsiveness
Pyne (2010)[141] USA	395 primary care patients screened positive for depression using PHQ-9 360 completed 6-month follow-up, 335 completed 12 months Excluded schizophrenia, suicide intention, pregnancy, substance dependence, bipolar	Responsiveness
Reed (2009)[142] UK	Patients (aged ≥ 18 years) with clinical depression enrolled prior to commencing antidepressant treatment Of 3468 at baseline, 343 had no follow-up data; 271 had data at 3 months only; 2854 had data at 3 and 6 months Age and sex not given	Known-group validity
Revicki (2008)[143] USA	297 patients with GAD; 72% female, mean age 47.6 years HAM-A score: ≤ 9 = asymptomatic; 10–15 = mild; 16–24 = moderate; ≥ 25 = severe	Known-group validity Convergent validity
Saarni (2007)[144] Finland	Population survey, aged ≥ 30 years Included assessment of 12-month prevalence of depressive anxiety or alcohol disorders (DSM-IV)	Known-group validity
Sapin (2004)[145] France	Outpatient population consulting at GP for new episode of MDD according to DSM-IV Aged ≥ 18 years, not treated with any antidepressants prior to inclusion Exclusion symptoms suggest schizophrenia or psychotic symptoms	Known-group validity Convergent validity Responsiveness

TABLE 35 Anxiety and depression review studies (*continued*)

Study details	Population characteristics	Properties measured
Sobocki (2007)[146] Sweden	Subjects attending GP at 56 centres in Sweden $N = 447$; baseline data on $n = 394$ Mean age = 47 years, 67% female Aged $\geq$ 18 years with a diagnosis of depression (according to centres' practice), initiating new treatment with antidepressants	Known-group validity Responsiveness
Stein (2005)[147] USA	480 outpatients with anxiety disorder, 63% female	Known-group validity
Supina (2007)[148] Canada	Alberta Mental Health Survey, stratified random sample Sample size $n = 5410$ (77% return), $n = 5383$ successful data Mean age = 40.8 years, 61.2% female	Known-group validity
Zivin (2008)[149] USA	$N = 87,797$ veterans, $n = 58,442$ with depression Identified from VA depression registry and VA outpatients Mean age = 60 years, 10% female	Known-group validity

GAD, generalised anxiety disorder; GP, general practitioner; HAM-A, Hamilton Anxiety Rating Scale; MDD, major depressive disorder; SCID, Structured Clinical Interview for DSM Disorders; VA, Veterans Administration.

Appendix 3 *Chapter 4* appendices

Preliminary analysis

TABLE 36 Correlation analysis: EQ-5D, SF-6D and HADS (AHEAD)

	HADS-D	HADS-A	HADS-T
HADS-D	1		
HADS-A	0.2721*	1	
HADS-T	0.8184*	0.7533*	1
EQ-5D	−0.3606*	−0.2839*	−0.4089*
Mobility	0.0814	0.0224	0.0673
Self-care	0.1856*	0.1245	0.1988*
Usual activities	0.2815*	0.1727*	0.3011*
Pain/discomfort	0.0354	0.0897	0.0640
Anxiety/depression	0.4033*	0.3943*	0.5031*
SF-6D	−0.5346*	−0.3368*	−0.5618*

* Significant at 1% level or less.

TABLE 37 Correlation analysis: EQ-5D and HADS (POPMACT)

	HADS-D	HADS-A	HADS-T
HADS-D	1		
HADS-A	0.5242*	1	
HADS-T	0.8927*	0.8387*	1
EQ-5D	−0.4533*	−0.3947*	−0.4858*
Mobility	0.1148	0.0073	0.0785
Self-care	0.1623*	0.1098	0.1609*
Usual activities	0.4052*	0.2601*	0.3861*
Pain/discomfort	0.2079*	0.1571*	0.2115*
Anxiety/depression	0.4524*	0.4363*	0.5045*

* Significant at 1% level or less.

TABLE 38 Correlation analysis: SF-6D, PHQ-9, GAD-7 and CORE-OM totals/dimensions (IAPT)

	PHQ-9	GAD-7	CORE T	CORE W	CORE F	CORE S	CORE R
SF-6D	−0.6174*	−0.5272*	−0.6209*	−0.5618*	−0.5273*	−0.6480*	−0.3457*
Physical functioning	0.3877*	0.2806*	0.3449*	0.3069*	0.2902*	0.3895*	0.1449*
Role limitation	0.4632*	0.4049*	0.4755*	0.4359*	0.4340*	0.4718*	0.2464*
Social functioning	0.5032*	0.4316*	0.5392*	0.4806*	0.4802*	0.5466*	0.2920*
Pain	0.2572*	0.2224*	0.2810*	0.1958*	0.1856*	0.3505*	0.1650*
Mental health	0.6846*	0.6453*	0.6888*	0.6638*	0.6121*	0.6681*	0.4305*
Vitality	0.5233*	0.3600*	0.4721*	0.4942*	0.4352*	0.4715*	0.2370*

F, function; R, risk; S, symptoms; T, total; W, well-being.
* Significant at 1% level or less.

TABLE 39 Correlation analysis: SF-6D and PHQ-9 items (IAPT)

	PHQ1	PHQ2	PHQ3	PHQ4	PHQ5	PHQ6	PHQ7	PHQ8	PHQ9	PHQ10
PHQ1	1									
PHQ2	0.7727*	1								
PHQ3	0.6195*	0.5720*	1							
PHQ4	0.6216*	0.5350*	0.6095*	1						
PHQ5	0.5273*	0.5154*	0.5107*	0.4963*	1					
PHQ6	0.5744*	0.6569*	0.4709*	0.4387*	0.4913*	1				
PHQ7	0.6263*	0.5921*	0.5752*	0.5471*	0.5592*	0.5843*	1			
PHQ8	0.5198*	0.4736*	0.4393*	0.3757*	0.4238*	0.4747*	0.5777*	1		
PHQ9	0.4654*	0.5022*	0.3294*	0.3445*	0.3805*	0.5151*	0.4080*	0.3743*	1	
PHQ10	0.6120*	0.5780*	0.4561*	0.5428*	0.4621*	0.5147*	0.5449*	0.5053*	0.4250*	1
SF-6D	−0.5827*	−0.4971*	−0.4486*	−0.5657*	−0.4019*	−0.4022*	−0.5287*	−0.4145*	−0.3434*	−0.5949*
Physical functioning	0.4132*	0.2870*	0.3086*	0.4360*	0.2253*	0.1665*	0.3571*	0.2455*	0.1927*	0.3900*
Role limitation	0.4217*	0.3779*	0.3611*	0.3925*	0.3180*	0.3514*	0.4330*	0.3289*	0.2231*	0.5140*
Social functioning	0.5057*	0.3740*	0.2998*	0.4286*	0.3143*	0.3380*	0.4442*	0.3950*	0.2848*	0.5814*
Pain	0.2217*	0.1898*	0.2041*	0.2850*	0.1763*	0.1055	0.2170*	0.1849*	0.1309*	0.2414*
Mental health	0.6128*	0.6311*	0.4635*	0.4903*	0.4706*	0.5574*	0.5503*	0.4414*	0.4321*	0.6183*
Vitality	0.5195*	0.3893*	0.4154*	0.6311*	0.3495*	0.3477*	0.4288*	0.2765*	0.2366*	0.5247*

* Significant at 1% level or less.

TABLE 40 Correlation analysis: SF-6D and GAD-7 items (IAPT)

	GAD1	GAD2	GAD3	GAD4	GAD5	GAD6	GAD7	GAD8
GAD1	1							
GAD2	0.7353*	1						
GAD3	0.7170*	0.8924*	1					
GAD4	0.6338*	0.6554*	0.6784*	1				
GAD5	0.5638*	0.5204*	0.5443*	0.6517*	1			
GAD6	0.5176*	0.4775*	0.4980*	0.5327*	0.4873*	1		
GAD7	0.5729*	0.5951*	0.5786*	0.4976*	0.4534*	0.4431*	1	
GAD8	0.5833*	0.5461*	0.5293*	0.5274*	0.4308*	0.4455*	0.5061*	1
SF-6D	−0.4896*	−0.4593*	−0.4362*	−0.4403*	−0.3659*	−0.3630*	−0.3787*	−0.5143*
Physical functioning	0.2864*	0.2453*	0.2170*	0.2391*	0.2227*	0.1513*	0.1703*	0.2678*
Role limitation	0.4116*	0.3680*	0.3468*	0.3590*	0.2399*	0.2916*	0.2712*	0.4600*
Social functioning	0.4343*	0.3993*	0.3583*	0.3729*	0.2961*	0.2705*	0.3114*	0.4964*
Pain	0.1862*	0.1532*	0.1531*	0.1860*	0.1768*	0.1513*	0.1751*	0.2070*
Mental health	0.6001*	0.6056*	0.5763*	0.5186*	0.4263*	0.4802*	0.4759*	0.5831*
Vitality	0.3216*	0.3567*	0.2956*	0.3231*	0.2168*	0.2579*	0.2415*	0.4037*

* Significant at 1% level or less.

TABLE 41 Correlation analysis: CORE-OM items (IAPT)

	CORE1	CORE2	CORE3	CORE4	CORE5	CORE6	CORE7	CORE8
CORE1	1							
CORE2	0.6054*	1						
CORE3	0.3192*	0.1954*	1					
CORE4	0.5982*	0.5103*	0.3145*	1				
CORE5	0.5303*	0.5319*	0.0666	0.4527*	1			
CORE6	0.1920*	0.0852	0.1403*	0.1174	0.1345*	1		
CORE7	0.4937*	0.4551*	0.2019*	0.5607*	0.4099*	0.0998	1	
CORE8	0.1621*	0.2598*	0.0161	0.1760*	0.2996*	0.1097	0.1276	1
CORE9	0.4893*	0.3362*	0.1803*	0.4222*	0.2707*	0.2263*	0.3508*	0.1305*
CORE10	0.5537*	0.5119*	0.1897*	0.4353**	0.4222*	0.0992	0.4418*	0.1568*
CORE11	0.5948*	0.6685*	0.1763*	0.4839*	0.5250*	0.1087	0.4470*	0.3079*
CORE12	0.5249*	0.4398*	0.2860*	0.6151*	0.4029*	0.1490*	0.5021*	0.1118
CORE13	0.5332*	0.4939*	0.1809*	0.4346*	0.3533*	0.1646*	0.3869*	0.2290*
CORE14	0.5604*	0.4991*	0.1677*	0.4640*	0.4691*	0.1429*	0.4156*	0.1543*
CORE15	0.5332*	0.6971*	0.1321*	0.4285*	0.4502*	0.1307*	0.3849*	0.2461*
CORE16	0.4916*	0.3521*	0.1592*	0.3419*	0.2782*	0.2315*	0.3052*	0.1226
CORE17	0.7112*	0.6587*	0.2336*	0.5686*	0.5387*	0.1797*	0.4985*	0.2089*
CORE18	0.4190*	0.4509*	0.1025	0.3281*	0.3775*	0.1509*	0.2716*	0.2870*
CORE19	0.3661*	0.2019*	0.2774*	0.3364*	0.2021*	0.1666*	0.2921*	0.0737
CORE20	0.6075*	0.6192*	0.1524*	0.5420*	0.5185*	0.1227	0.4398*	0.2774*
CORE21	0.5095*	0.4400*	0.2097*	0.5070*	0.4480*	0.1446*	0.4958*	0.1915*
CORE22	0.2474*	0.1798*	0.1070	0.1338*	0.1070	0.4875*	0.2205*	0.1466*
CORE23	0.7157*	0.5894*	0.2667*	0.6455*	0.5163*	0.2399*	0.4882*	0.2157*
CORE24	0.5592*	0.3784*	0.2396*	0.4778*	0.3597*	0.2123*	0.3654*	0.1289
CORE25	0.4759*	0.3578*	0.2879*	0.3862*	0.3009*	0.2256*	0.3050*	0.0873
CORE26	0.6172*	0.4567*	0.4082*	0.4330*	0.3567*	0.1801*	0.3872*	0.1371*
CORE27	0.7316*	0.6392*	0.2359*	0.6233*	0.5605*	0.1534*	0.5210*	0.1938*
CORE28	0.5361*	0.4703*	0.1980*	0.4329*	0.3139*	0.1251	0.3782*	0.1960*
CORE29	0.4510*	0.4613*	0.1714*	0.4332*	0.3323*	0.2055*	0.3356*	0.2245*
CORE30	0.5152*	0.4039*	0.2464*	0.4320*	0.2956*	0.1536*	0.3212*	−0.004
CORE31	0.3829*	0.3451*	0.1535*	0.4554*	0.3910*	0.0078	0.4745*	0.0478
CORE32	0.4529*	0.3071*	0.2565*	0.4918*	0.3267*	0.1027	0.4432*	0.0737
CORE33	0.4561*	0.3431*	0.2227*	0.3012*	0.2879*	0.2064*	0.2673*	0.0935
CORE34	0.3507*	0.2773*	0.1902*	0.2776*	0.2213*	0.2714*	0.2562*	0.1728*

* Significant at 1% level or less.

CORE9	CORE10	CORE11	CORE12	CORE13	CORE14	CORE15	CORE16	CORE17
1								
0.3919*	1							
0.3326*	0.5784*	1						
0.4499*	0.4528*	0.4886*	1					
0.4128*	0.4123*	0.4516*	0.3438*	1				
0.4076*	0.3633*	0.3851*	0.4373*	0.4152*	1			
0.3691*	0.4161*	0.6286*	0.4143*	0.4886*	0.4988*	1		
0.7115*	0.3555*	0.3490*	0.3750*	0.4109*	0.3974*	0.4044*	1	
0.4514*	0.5128*	0.6561*	0.5171*	0.5091*	0.5570*	0.6271*	0.4262*	1
0.2043*	0.3534*	0.4228*	0.2891*	0.3204*	0.3368*	0.4184*	0.2603*	0.4269*
0.2051*	0.2340*	0.2107*	0.3378*	0.1548*	0.1610*	0.2161*	0.1609*	0.2307*
0.3367*	0.4701*	0.6136*	0.4278*	0.5368*	0.4915*	0.5741*	0.3674*	0.6487*
0.3138*	0.4072*	0.5927*	0.5365*	0.3326*	0.3366*	0.3828*	0.2659*	0.4967*
0.2537*	0.2144*	0.1493*	0.1613*	0.2077*	0.1682*	0.1925*	0.2778*	0.2135*
0.5195*	0.5655*	0.5884*	0.5733*	0.5834*	0.5616*	0.5511*	0.4844*	0.7383*
0.7602*	0.4114*	0.3978*	0.4602*	0.4920*	0.4477*	0.3571*	0.6761*	0.5200*
0.3467*	0.4447*	0.3383*	0.3320*	0.3179*	0.3347*	0.3018*	0.3553*	0.4691*
0.4371*	0.5103*	0.4523*	0.4631*	0.3780*	0.3929*	0.4074*	0.4343*	0.5158*
0.4446*	0.5301*	0.5766*	0.5500*	0.5243*	0.6579*	0.5124*	0.4488*	0.7090*
0.4034*	0.4078*	0.4527*	0.3925*	0.6349*	0.4522*	0.5115*	0.4079*	0.5132*
0.2840*	0.4438*	0.4228*	0.4387*	0.3616*	0.3949*	0.3503*	0.2862*	0.4744*
0.4231*	0.4622*	0.3969*	0.3727*	0.3703*	0.4140*	0.3381*	0.3975*	0.5102*
0.2798*	0.3379*	0.3698*	0.4860*	0.2085*	0.2666*	0.2745*	0.2166*	0.3921*
0.3793*	0.2989*	0.3850*	0.5403*	0.2774*	0.4085*	0.2760*	0.3146*	0.3988*
0.3039*	0.3761*	0.3214*	0.2884*	0.3080*	0.3131*	0.3600*	0.3435*	0.4075*
0.6004*	0.3117*	0.3050*	0.2971*	0.3607*	0.2902*	0.3374*	0.5932*	0.3513*

TABLE 41 Correlation analysis: CORE-OM items (IAPT) (*continued*)

	CORE18	CORE19	CORE20	CORE21	CORE22	CORE23	CORE24	CORE25
CORE18	1							
CORE19	0.2199*	1						
CORE20	0.4430*	0.1711*	1					
CORE21	0.2861*	0.3249*	0.4708*	1				
CORE22	0.1612*	0.1701*	0.2236*	0.1565*	1			
CORE23	0.4538*	0.2942*	0.6327*	0.4855*	0.2794*	1		
CORE24	0.3092*	0.2144*	0.4206*	0.3990*	0.2826*	0.6498*	1	
CORE25	0.2837*	0.1604*	0.3757*	0.2912*	0.2675*	0.4866*	0.4177*	1
CORE26	0.3598*	0.2343*	0.4709*	0.3833*	0.2485*	0.5559*	0.4648*	0.5708*
CORE27	0.4827*	0.2526*	0.6466*	0.4965*	0.2058*	0.7647*	0.5310*	0.4741*
CORE28	0.3887*	0.2122*	0.5046*	0.3764*	0.2245*	0.5487*	0.4436*	0.3229*
CORE29	0.3415*	0.1607*	0.4521*	0.3464*	0.3448*	0.4980*	0.3640*	0.4985*
CORE30	0.2484*	0.1294	0.4448*	0.3391*	0.2287*	0.5400*	0.4832*	0.4852*
CORE31	0.1962*	0.2624*	0.3098*	0.3527*	0.0298	0.3844*	0.3130*	0.2648*
CORE32	0.2190*	0.3737*	0.3640*	0.4741*	0.2116*	0.4800*	0.4221*	0.2712*
CORE33	0.2070*	0.1621*	0.3005*	0.2844*	0.2970*	0.3942*	0.3920*	0.6118*
CORE34	0.2296*	0.1444*	0.3386*	0.2798*	0.3467*	0.4478*	0.5295*	0.2902*

* Significant at 1% level or less.

CORE26	CORE27	CORE28	CORE29	CORE30	CORE31	CORE32	CORE33	CORE34
1								
0.5750*	1							
0.4769*	0.5422*	1						
0.4513*	0.5511*	0.4015*	1					
0.5270*	0.5196*	0.4411*	0.4524*	1				
0.2498*	0.4380*	0.1864*	0.2921*	0.2563*	1			
0.4006*	0.4900*	0.2740*	0.3437*	0.3648*	0.4227*	1		
0.5283*	0.3898*	0.3766*	0.4000*	0.4115*	0.2139*	0.2116*	1	
0.3464*	0.3301*	0.3848*	0.2802*	0.3269*	0.0746	0.2348*	0.3134*	1

TABLE 42 Correlation analysis: SF-6D and CORE-OM items (IAPT)

	CORE1	CORE2	CORE3	CORE4	CORE5	CORE6	CORE7	CORE8
SF-6D	–0.4949*	–0.4812*	–0.1438*	–0.5205*	–0.4724*	–0.1390*	–0.4257*	–0.5498*
sfphysical	0.2498*	0.2379*	0.1178	0.3253*	0.3470*	0.0118	0.2812*	0.4231*
sfrole	0.4244*	0.4045*	0.1359*	0.3957*	0.3272*	0.0828	0.3287*	0.2186*
sfsocial	0.4453*	0.4538*	0.1161	0.4682*	0.3959*	0.0631	0.4125*	0.3085*
sfpain	0.1690*	0.1571*	0.003	0.2271*	0.2085*	0.1411*	0.1376*	0.6958*
sfmental	0.5891*	0.5936*	0.1700*	0.5450*	0.4639*	0.1418*	0.5074*	0.2303*
sfvital	0.3894*	0.3326*	0.1069	0.4663*	0.4933*	0.1046	0.3710*	0.2796*
	CORE18	**CORE19**	**CORE20**	**CORE21**	**CORE22**	**CORE23**	**CORE24**	**CORE25**
SF-6D	–0.4254*	–0.2323*	–0.4989*	–0.4330*	–0.0853	–0.4989*	–0.3533*	–0.2833*
sfphysical	0.2701*	0.1730*	0.2725*	0.2805*	–0.0106	0.2805*	0.1795*	0.064
sfrole	0.2751*	0.1422*	0.4092*	0.3557*	0.1075	0.3643*	0.2077*	0.2276*
sfsocial	0.2517*	0.2426*	0.4611*	0.4438*	0.1154	0.4446*	0.2849*	0.2387*
sfpain	0.2730*	0.056	0.2514*	0.1943*	0.0647	0.1922*	0.1523*	0.095
sfmental	0.4099*	0.2773*	0.5405*	0.4124*	0.1239	0.6109*	0.4238*	0.3708*
sfvital	0.3413*	0.2037*	0.3761*	0.4145*	0.0172	0.3931*	0.2549*	0.1981*

sfmental, mental health; sfpain, pain; sfphysical, physical functioning; sfrole, role limitation; sfsocial, social functioning; sfvital, vitality.
* Significant at 1% level or less.

CORE9	CORE10	CORE11	CORE12	CORE13	CORE14	CORE15	CORE16	CORE17
−0.3102*	−0.4313*	−0.5809*	−0.4618*	−0.3789*	−0.4018*	−0.4148*	−0.3247*	−0.5203*
0.1475*	0.2781*	0.3913*	0.2831*	0.2354*	0.1919*	0.2532*	0.1463*	0.2541*
0.2227*	0.3744*	0.4677*	0.3548*	0.2843*	0.2662*	0.3109*	0.1955*	0.4290*
0.2795*	0.4187*	0.5628*	0.3930*	0.2944*	0.3117*	0.3864*	0.2867*	0.4674*
0.1444*	0.1478*	0.2564*	0.1776*	0.2264*	0.1649*	0.1708*	0.1661*	0.1943*
0.3959*	0.4927*	0.5679*	0.4947*	0.4198*	0.4767*	0.4952*	0.3662*	0.6194*
0.2040*	0.3452*	0.3980*	0.4126*	0.2539*	0.3345*	0.3005*	0.1929*	0.3986*

CORE26	CORE27	CORE28	CORE29	CORE30	CORE31	CORE32	CORE33	CORE34
−0.3975*	−0.5234*	−0.3771*	−0.3498*	−0.2722*	−0.3244*	−0.3067*	−0.2838*	−0.2307*
0.1921*	0.2826*	0.2313*	0.1525*	0.0719	0.2192*	0.1621*	0.1241	0.0892
0.3269*	0.4033*	0.2373*	0.3001*	0.2738*	0.2439*	0.2452*	0.2085*	0.1972*
0.3572*	0.4839*	0.2878*	0.2446*	0.3186*	0.2745*	0.3084*	0.2892*	0.2074*
0.1422*	0.1695*	0.2319*	0.1910*	0.0308	0.0535	0.0741	0.102	0.1539*
0.4538*	0.6532*	0.4117*	0.4456*	0.4026*	0.4301*	0.4039*	0.3317*	0.2755*
0.2825*	0.4413*	0.2596*	0.2795*	0.1908*	0.3407*	0.3221*	0.1961*	0.1288

TABLE 43 Correlation analysis: SF-6D and GHQ-12 score and items (EMPIRIC)

	GHQ-12 total	GHQ1	GHQ2	GHQ3	GHQ4	GHQ5	GHQ6	GHQ7	GHQ8	GHQ9	GHQ10	GHQ11	GHQ12
GHQ-12 total	1												
GHQ1	0.6177*	1											
GHQ2	0.7001*	0.4239*	1										
GHQ3	0.5000*	0.3230*	0.2150*	1									
GHQ4	0.5874*	0.3291*	0.3321*	0.4482*	1								
GHQ5	0.7653*	0.4983*	0.5495*	0.2417*	0.4080*	1							
GHQ6	0.7475*	0.4171*	0.4690*	0.3255*	0.3501*	0.6113*	1						
GHQ7	0.7471*	0.5352*	0.4896*	0.4609*	0.4094*	0.5532*	0.4693*	1					
GHQ8	0.6686*	0.3629*	0.3703*	0.2572*	0.4096*	0.5083*	0.5726*	0.4691*	1				
GHQ9	0.8489*	0.4695*	0.5855*	0.3071*	0.4963*	0.6727*	0.6233*	0.6164*	0.5682*	1			
GHQ10	0.7879*	0.3924*	0.4899*	0.2733*	0.3990*	0.5154*	0.5888*	0.4873*	0.5366*	0.6539*	1		
GHQ11	0.7334*	0.3683*	0.4224*	0.3824*	0.4066*	0.4435*	0.5046*	0.4559*	0.5183*	0.5291*	0.7034*	1	
GHQ12	0.7648*	0.3834*	0.4755*	0.4534*	0.4813*	0.5136*	0.5093*	0.5831*	0.5271*	0.6484*	0.5317*	0.5340*	1
SF-6D	−0.2593*	−0.2038*	−0.1396	−0.2456*	−0.1999*	−0.2712*	−0.1738	−0.1707	−0.2146*	−0.2783*	−0.0888	−0.1197	−0.2386*
Physical functioning	0.0875	0.081	0.0163	0.2052*	0.0951	0.0878	0.0734	0.0564	0.0562	0.0388	0.0222	0.0472	0.0419
Role limitation	0.2433*	0.2016*	0.1524	0.2572*	0.1990*	0.2325*	0.2011*	0.1316	0.1883*	0.2121*	0.0972	0.1625	0.1915*
Social functioning	0.1777*	0.1689	0.1001	0.127	0.147	0.2144*	0.0691	0.1396	0.1246	0.2174*	0.0567	0.068	0.1573
Pain	0.1063	0.1169	0.0923	0.1758	0.1262	0.1686	0.0599	0.0657	0.1217	0.0947	−0.0466	0.0096	0.1143
Mental health	0.1829*	0.1476	0.1251	0.0398	0.1828*	0.166	0.0627	0.0892	0.1668	0.2612*	0.037	0.0764	0.2076*
Vitality	0.2970*	0.2850*	0.1601	0.1900*	0.1571	0.2212*	0.2030*	0.2198*	0.2198*	0.2427*	0.2261*	0.2150*	0.2247*

GHQ: 1, concentrate; 2, problem sleeping; 3, felt useful; 4, felt capable of decision-making; 5, strain; 6, could not overcome difficulties; 7, enjoyed normal activities; 8, able to face problems; 9, unhappy/depressed; 10, lost confidence; 11, felt worthless; 12, happy all things considered.
* Significant at 1% level or less.

Summary performance statistics

TABLE 44 Summary performance statistics: best-fitting OLS, Tobit and response-level mapping models – EQ-5D and HADS (AHEAD)

Independent variables	Observed EQ-5D	Predicted EQ-5D		
		OLS	Tobit	RLM
		HADS-T + age	HADS-T + age	HADS-T
Mean	0.6118	0.6118	0.6058	0.7480
SD	0.262	0.115	0.110	0.023
Minimum	−0.126	0.333	0.327	0.697
Maximum	1	0.917	0.866	0.852
R^2		0.193	–	–
Pseudo-R^2		–	0.932	–
RMSE		0.188	0.189	0.201
BIC		1	33	–

EQ-5D range	n	EQ-5D	OLS	RMSE	Tobit	RMSE	RLM	RMSE
< 0.3	59	0.1816	0.5495	0.368	0.5453	0.364	0.7372	0.556
0.3–0.599	34	0.4186	0.5254	0.124	0.5219	0.122	0.7349	0.316
0.6–0.699	52	0.6709	0.6002	0.099	0.5959	0.099	0.7483	0.077
0.7–0.799	34	0.7402	0.655	0.105	0.6474	0.103	0.7564	0.025
0.8–1	107	0.8408	0.6655	0.179	0.6575	0.184	0.7554	0.086

HADS-D	n	EQ-5D	OLS	RMSE	Tobit	RMSE	RLM	RMSE
Non-case/possible ≤ 10	147	0.6931	0.6816	0.160	0.6731	0.161	0.7621	0.147
Probable case > 10	139	0.5258	0.5379	0.218	0.5347	0.218	0.7332	0.259

RLM, response-level mapping.

TABLE 45 Summary performance statistics: best-fitting OLS, Tobit and response-level mapping models – EQ-5D and HADS (POPMACT)

| Independent variables | Observed EQ-5D | Predicted EQ-5D | | |
		OLS HADS-D and -A	Tobit HADS-D and -A	RLM HADS-D and -A
Mean	0.5039	0.5039	0.5008	0.5480
SD	0.319	0.157	0.156	0.129
Minimum	−0.594	0.149	0.129	0.261
Maximum	1	1.016	0.915	0.795
R^2		0.242	–	
Pseudo-R^2		–	0.405	–
RMSE		0.227	0.227	0.230
BIC		150	230	–

EQ-5D range	n	EQ-5D	OLS	RMSE	Tobit	RMSE	RLM	RMSE
≤ 0	33	−0.1245	0.4021	0.527	0.3977	0.522	0.4660	0.591
0.001–0.199	60	0.1295	0.4423	0.313	0.4395	0.310	0.4975	0.368
0.2–0.299	64	0.263	0.4324	0.178	0.4296	0.178	0.4879	0.225
0.3–0.599	90	0.4242	0.4449	0.103	0.4421	0.106	0.4973	0.111
0.6–0.699	57	0.6699	0.532	0.159	0.5319	0.157	0.5797	0.116
0.7–0.799	45	0.7374	0.5421	0.203	0.5395	0.202	0.5813	0.157
0.8–0.899	103	0.8382	0.586	0.259	0.5829	0.257	0.6172	0.221
0.9–1	23	1	0.7295	0.272	0.7157	0.284	0.7100	0.290

HADS-D	n	EQ-5D	OLS	RMSE	Tobit	RMSE	RLM	RMSE
Non-case/possible ≤ 10	210	0.622	0.6375	0.220	0.6341	0.220	0.6564	0.218
Probable case > 10	265	0.4104	0.3981	0.233	0.3951	0.233	0.4622	0.238

RLM, response-level mapping.

TABLE 46 Summary performance statistics: Tobit models (IAPT)

Independent variables	Observed SF-6D	Predicted SF-6D					
		PHQ-9		GAD-7		CORE items	
Mean	0.6193	0.6191		0.6191		0.6191	
SD	0.126	0.091		0.078		0.107	
Minimum	0.301	0.425		0.482		0.367	
Maximum	0.929	0.857		0.832		0.888	
RMSE		0.070		0.079		0.054	
SF-6D range	**SF-6D**	**PHQ-9**	**RMSE**	**GAD-7**	**RMSE**	**CORE items**	**RMSE**
< 0.4	0.3610	0.492	0.131	0.5313	0.170	0.4582	0.101
0.4–0.499	0.4690	0.5589	0.093	0.5695	0.101	0.497	0.046
0.5–0.599	0.5587	0.5691	0.053	0.5869	0.056	0.5682	0.047
0.6–0.699	0.6386	0.6354	0.051	0.625	0.054	0.6442	0.047
0.7–0.799	0.7444	0.687	0.066	0.6745	0.079	0.7081	0.053
0.8–1	0.8552	0.7377	0.119	0.7098	0.145	0.7789	0.080
PHQ-9 cut-offs	**SF-6D**	**PHQ-9**	**RMSE**	**GAD-7**	**RMSE**	**CORE items**	**RMSE**
None/mild	0.7272	0.7272	0.073	0.6932	0.085	0.7155	0.061
Moderate	0.6501	0.6475	0.071	0.6481	0.079	0.6507	0.053
Severe	0.5634	0.564	0.068	0.5777	0.076	0.5676	0.051

TABLE 47 Summary performance statistics: response-level mapping models (IAPT)

Independent variables	Observed SF-6D	Predicted SF-6D					
		PHQ-9		GAD-7		CORE items	
Mean	0.6193	0.6631		0.5672		0.6783	
SD	0.126	0.073		0.009		0.067	
Minimum	0.301	0.558		0.559		0.555	
Maximum	0.929	0.813		0.587		0.888	
RMSE		0.084		0.103		0.089	
SF-6D range	**SF-6D**	**PHQ-9**	**RMSE**	**GAD- 7**	**RMSE**	**CORE items**	**RMSE**
< 0.4	0.3610	0.5858	0.225	0.5593	0.198	0.6054	0.244
0.4–0.499	0.4690	0.6165	0.148	0.5626	0.094	0.6145	0.146
0.5–0.599	0.5587	0.6217	0.066	0.5634	0.023	0.6431	0.086
0.6–0.699	0.6386	0.6768	0.057	0.5676	0.071	0.6955	0.063
0.7–0.799	0.7444	0.7173	0.053	0.5736	0.171	0.7281	0.044
0.8–1	0.8552	0.7443	0.111	0.5766	0.279	0.7594	0.096
PHQ-9 cut-offs	**SF-6D**	**PHQ-9**	**RMSE**	**GAD- 7**	**RMSE**	**CORE items**	**RMSE**
None/mild	0.7272	0.7567	0.087	0.5765	0.161	0.7416	0.077
Moderate	0.6501	0.6929	0.083	0.5701	0.105	0.7029	0.083
Severe	0.5634	0.6133	0.083	0.5623	0.080	0.6429	0.097

Ordinary least squares regression results

TABLE 48 Regression analysis results: EQ-5D and HADS (AHEAD)

OLS variables	EQ-5D HADS-D and -A (SE)	EQ-5D HADS-D (SE)	EQ-5D HADS-A (SE)	EQ-5D HADS-T (SE)	EQ-5D HADS-T + age (SE)
HADS-D	−0.0213* (0.004)	−0.0243* (0.004)			
HADS-A	−0.0131* (0.004)		−0.0193* (0.004)		
HADS-T				−0.0176* (0.002)	−0.0185* (0.002)
Age					−0.0033* (0.001)
Constant	1.0059* (0.061)	0.8659* (0.041)	0.8638* (0.059)	1.0264* (0.059)	1.1872* (0.073)
Observations	286	286	286	286	286
R^2	0.161	0.133	0.065	0.156	0.193
Adjusted R^2	0.155	0.130	0.062	0.153	0.187
RMSE	0.194	0.198	0.212	0.194	0.188

SE, standard error.
* $p < 0.01$.

TABLE 49 Regression analysis: EQ-5D and HADS (POPMACT)

OLS variables	EQ-5D HADS-D and -A (SE)	EQ-5D HADS-D (SE)	EQ-5D HADS-A (SE)	EQ-5D HADS-T (SE)
HADS-D	−0.0245* (0.003)	−0.0323* (0.003)		
HADS-A	−0.0168* (0.004)		−0.0319* (0.003)	
HADS-T				−0.0210* (0.002)
Constant	1.0161* (0.048)	0.8663* (0.035)	0.9554* (0.050)	1.0366* (0.045)
Observations	475	475	475	475
R^2	0.242	0.212	0.155	0.240
Adjusted R^2	0.239	0.210	0.154	0.238
RMSE	0.227	0.233	0.245	0.228

SE, standard error.
* $p < 0.01$.

TABLE 50 Regression analysis: SF-6D and HADS (AHEAD)

OLS variables	SF-6D HADS-D and -A (SE)	SF-6D HADS-D (SE)	SF-6D HADS-A (SE)	SF-6D HADS-T (SE)	SF-6D HADS-D and -A + age (SE)
HADS-D	−0.0107* (0.001)	−0.0118* (0.001)			−0.0110* (0.001)
HADS-A	−0.0048* (0.001)		−0.0079* (0.001)		−0.0050* (0.001)
HADS-T				−0.0081* (0.001)	
Age					−0.0008* (0.000)
Constant	0.7945* (0.019)	0.7437* (0.013)	0.7229* (0.020)	0.8095* (0.018)	0.8355* (0.023)
Observations	286	286	286	286	286
R^2	0.305	0.273	0.094	0.283	0.326
Adjusted R^2	0.301	0.271	0.0909	0.280	0.319
RMSE	0.059	0.061	0.067	0.059	0.057

SE, standard error.
* $p < 0.01$.

TABLE 51 Regression analysis: SF-6D and PHQ-9 (IAPT)

Variables	SF-6D PHQ-9 total (SE)	SF-6D PHQ-9 items (SE)	SF-6D PHQ-9 items (significant variables) (SE)
PHQ-9 total	−0.0108*** (0.001)		
Little interest		−0.0583*** (0.021)	−0.0546*** (0.019)
		−0.0638** (0.025)	−0.0536** (0.021)
		−0.0974*** (0.028)	−0.0832*** (0.023)
Feeling down		0.0003 (0.027)	
		0.0155 (0.029)	
		0.0248 (0.031)	
Sleep problems		−0.0190 (0.022)	
		−0.0358 (0.023)	
		−0.0169 (0.023)	
Tired		−0.0034 (0.024)	−0.0101 (0.023)
		−0.0343 (0.025)	−0.0392* (0.024)
		−0.0681*** (0.025)	−0.0702*** (0.024)
Eating problems		0.0082 (0.017)	
		0.0083 (0.017)	
		0.0134 (0.017)	
Felt bad about self		0.0477*** (0.017)	0.0483*** (0.016)
		0.0105 (0.020)	0.0184 (0.019)
		0.0263 (0.020)	0.0364* (0.019)
Trouble concentrating		−0.0243 (0.018)	−0.0271 (0.017)
		−0.0476** (0.020)	−0.0469** (0.019)
		−0.0604*** (0.021)	−0.0572*** (0.019)
Slow movement/speech		0.0018 (0.013)	
		0.0038 (0.016)	
		−0.0047 (0.017)	
Suicidal/self-harm thoughts		0.0230* (0.013)	0.0219* (0.013)
		0.0183 (0.015)	0.0186 (0.015)
		−0.0209 (0.019)	−0.0188 (0.019)
Affected activities		−0.0506** (0.020)	−0.0536*** (0.018)
		−0.0951*** (0.023)	−0.0975*** (0.022)
		−0.1530*** (0.026)	−0.1527*** (0.024)
Constant	0.7878*** (0.012)	0.8026*** (0.026)	0.7967*** (0.023)
Observations	394	394	394
R^2	0.364	0.516	0.508
Adjusted R^2	0.362	0.476	0.484
RMSE	0.078	0.070	0.070

SE, standard error.
***$p < 0.01$; **$p < 0.05$; *$p < 0.1$.

TABLE 52 Regression analysis: SF-6D and GAD-7 (IAPT)

Variables	SF-6D GAD-7 total (SE)	SF-6D GAD-7 items (SE)	SF-6D GAD-7 items (significant variables) (SE)
GAD-7 total	−0.0118*** (0.001)		
Anxious/nervous		−0.0782*** (0.028)	−0.0790*** (0.027)
		−0.0925*** (0.031)	−0.0957*** (0.030)
		−0.1057*** (0.032)	−0.1115*** (0.031)
Uncontrollable worry		−0.0075 (0.030)	0.0013 (0.025)
		0.0130 (0.035)	0.0187 (0.027)
		−0.0390 (0.040)	−0.0286 (0.028)
Worry too much		0.0224 (0.037)	
		0.0181 (0.041)	
		0.0358 (0.045)	
Trouble relaxing		−0.0014 (0.030)	
		−0.0112 (0.033)	
		−0.0310 (0.034)	
Restless		−0.0173 (0.016)	
		−0.0229 (0.018)	
		−0.0088 (0.020)	
Irritable		−0.0416* (0.024)	−0.0341 (0.023)
		−0.0608** (0.026)	−0.0587** (0.024)
		−0.0541** (0.026)	−0.0563** (0.024)
Afraid		−0.0006 (0.017)	
		0.0076 (0.019)	
		−0.0048 (0.020)	
Affected activities		0.0033 (0.025)	−0.0034 (0.024)
		−0.0563** (0.028)	−0.0664** (0.026)
		−0.0907*** (0.031)	−0.1032*** (0.029)
Constant	0.7777*** (0.014)	0.8025*** (0.031)	0.8008*** (0.026)
Observations	394	394	394
R^2	0.273	0.383	0.371
Adjusted R^2	0.271	0.342	0.351
RMSE	0.085	0.079	0.080

SE, standard error.
***$p < 0.01$; **$p < 0.05$; *$p < 0.1$.

TABLE 53 Regression analysis: SF-6D and PHQ-9 and GAD-7 (IAPT)

Variables	SF-6D PHQ-9 and GAD-7 total (SE)	SF-6D PHQ-9 and GAD-7 items (SE)	SF-6D PHQ-9 and GAD-7 items (significant variables) (SE)
PHQ-9 total	−0.0086*** (0.001)		
GAD-7 total	−0.0038*** (0.001)		
Little interest		−0.0549*** (0.020)	−0.0490*** (0.018)
		−0.0559** (0.023)	−0.0483** (0.021)
		−0.0824*** (0.025)	−0.0746*** (0.023)
Sleep problems		−0.0193 (0.022)	
		−0.0374 (0.023)	
		−0.0149 (0.023)	
Tired		−0.0012 (0.024)	−0.0112 (0.022)
		−0.0284 (0.025)	−0.0397* (0.023)
		−0.0638** (0.025)	−0.0704*** (0.024)
Eating problems		0.0027 (0.017)	
		0.0101 (0.017)	
		0.0103 (0.017)	
Felt bad about self		0.0425** (0.017)	0.0446*** (0.016)
		0.0132 (0.019)	0.0182 (0.018)
		0.0312 (0.019)	0.0373** (0.019)
Trouble concentrating		−0.0180 (0.018)	−0.0238 (0.017)
		−0.0330 (0.020)	−0.0377** (0.018)
		−0.0513** (0.021)	−0.0478** (0.019)
Slow movement/speech		0.0035 (0.013)	
		0.0147 (0.016)	
		0.0051 (0.019)	
Suicidal/self-harm thoughts		0.0235* (0.013)	0.0229* (0.013)
		0.0171 (0.015)	0.0186 (0.015)
		−0.0184 (0.019)	−0.0144 (0.018)
Affected activities		−0.0382* (0.020)	−0.0417** (0.018)
		−0.0790*** (0.023)	−0.0801*** (0.022)
		−0.1317 *** (0.026)	−0.1341*** (0.025)
Anxious/nervous		−0.0517** (0.022)	−0.0493** (0.021)
		−0.0583** (0.024)	−0.0554** (0.023)
		−0.0853*** (0.025)	−0.0771*** (0.023)
Restless		−0.0101 (0.014)	
		−0.0166 (0.017)	
		−0.0003 (0.019)	
Irritable		0.0050 (0.022)	
		−0.0115 (0.024)	

TABLE 53 Regression analysis: SF-6D and PHQ-9 and GAD-7 (IAPT) (*continued*)

Variables	SF-6D PHQ-9 and GAD-7 total (SE)	SF-6D PHQ-9 and GAD-7 items (SE)	SF-6D PHQ-9 and GAD-7 items (significant variables) (SE)
		0.0104 (0.024)	
Afraid		0.0101 (0.015)	
		0.0185 (0.016)	
		0.0068 (0.017)	
Constant	0.8037*** (0.014)	0.8364*** (0.029)	0.8303*** (0.026)
Observations	394	394	394
R^2	0.376	0.538	0.524
Adjusted R^2	0.373	0.487	0.497
RMSE	0.077	0.067	0.068

SE, standard error.
***$p < 0.01$; **$p < 0.05$; *$p < 0.1$.

TABLE 54 Regression analysis: SF-6D and CORE-OM dimensions (IAPT)

	SF-6D CORE mean (SE)	SF-6D CORE dimensions (SE)	CORE dimensions (significant variables) (SE)
CORE-OM mean	−0.0099* (0.001)		
CORE-OM functioning		−0.0086 (0.011)	
CORE-OM risk		0.0101 (0.008)	
CORE-OM symptom		−0.0753* (0.010)	−0.0874* (0.005)
CORE-OM well-being		−0.0140 (0.010)	
Constant	0.8165* (0.014)	0.8453* (0.016)	0.8263* (0.014)
Observations	394	394	394
R^2	0.373	0.418	0.409
Adjusted R^2	0.372	0.412	0.407
RMSE	0.077	0.075	0.075

SE, standard error.
*$p < 0.01$.

TABLE 55 Regression analysis results: SF-6D and CORE-OM items (IAPT)

Dimensions	Items (item number below)	SF-6D CORE items (SE)	SF-6D CORE significant items (SE)
Well-being	Felt OK	0.0415** (0.017)	0.0385*** (0.013)
	4	0.0603*** (0.021)	0.0558*** (0.016)
		0.0705*** (0.027)	0.0643*** (0.021)
		0.0368 (0.034)	0.0490* (0.027)
	Felt like crying	0.0260 (0.025)	0.0439** (0.022)
	14	0.0255 (0.022)	0.0337* (0.020)
		−0.0070 (0.023)	0.0045 (0.020)
		−0.0111 (0.025)	0.0023 (0.020)
	Overwhelmed by problems	0.0326 (0.022)	
	17	0.0057 (0.025)	
		0.0100 (0.027)	
		0.0056 (0.030)	
	Optimistic about future	−0.0060 (0.015)	
	31	0.0060 (0.018)	
		−0.0062 (0.022)	
		0.0104 (0.025)	
Symptoms	Tense/anxious	−0.0251 (0.020)	−0.0312* (0.016)
	11	−0.0434** (0.020)	−0.0429*** (0.016)
		−0.0816*** (0.023)	−0.0713*** (0.016)
		−0.1089*** (0.026)	−0.0988*** (0.018)
	Panic or terror	−0.0009 (0.015)	
	15	0.0039 (0.017)	
		0.0311 (0.020)	
		0.0520** (0.024)	
	Problems persistent	−0.0100 (0.029)	
	20	−0.0254 (0.031)	
		−0.0160 (0.031)	
		−0.0071 (0.032)	
	Lack energy/enthusiasm	−0.0196 (0.026)	
	5	−0.0224 (0.025)	
		−0.0211 (0.025)	
		−0.0149 (0.025)	
	Hopeless	0.0019 (0.020)	
	23	0.0070 (0.022)	
		0.0058 (0.024)	
		0.0083 (0.029)	

TABLE 55 Regression analysis results: SF-6D and CORE-OM items (IAPT) (*continued*)

Dimensions	Items (item number below)	SF-6D CORE items (SE)	SF-6D CORE significant items (SE)
	Unhappy	−0.0072 (0.044)	
	27	0.0103 (0.049)	
		0.0061 (0.050)	
		−0.0098 (0.052)	
	Blamed self	−0.0012 (0.021)	
	30	0.0188 (0.020)	
		−0.0155 (0.020)	
		−0.0056 (0.021)	
	Physical problems (pains/aches)	−0.0383** (0.017)	−0.0454*** (0.014)
	8	−0.0694*** (0.017)	−0.0591*** (0.014)
		−0.0873*** (0.016)	−0.0882*** (0.014)
		−0.1370*** (0.016)	−0.1373*** (0.013)
	Problems sleeping	0.0046 (0.021)	
	18	−0.0052 (0.021)	
		−0.0019 (0.019)	
		−0.0239 (0.019)	
	Unwanted thoughts	−0.0195 (0.019)	
	13	−0.0014 (0.018)	
		−0.0230 (0.019)	
		−0.0276 (0.022)	
	Distressing images/memories	0.0249 (0.018)	
	28	0.0423** (0.017)	
		0.0343* (0.018)	
		0.0259 (0.020)	
Functioning	Felt alone	−0.0395* (0.020)	
	1	−0.0278 (0.024)	
		−0.0146 (0.027)	
		−0.0249 (0.029)	
	Felt have someone to turn to	−0.0283 (0.017)	
	3	−0.0177 (0.017)	
		−0.0323* (0.019)	
		−0.0205 (0.019)	
	Felt warmth/affection	−0.0011 (0.018)	
	19	−0.0003 (0.017)	
		0.0170 (0.018)	
		0.0024 (0.018)	

TABLE 55 Regression analysis results: SF-6D and CORE-OM items (IAPT) (*continued*)

Dimensions	Items (item number below)	SF-6D CORE items (SE)	SF-6D CORE significant items (SE)
	Thought have no friends	0.0039 (0.017)	
	26	−0.0141 (0.019)	
		−0.0174 (0.019)	
		−0.0335 (0.022)	
	Felt able to cope	0.0211 (0.016)	0.0148 (0.014)
	7	0.0006 (0.017)	0.0005 (0.015)
		0.0007 (0.021)	0.0081 (0.019)
		0.0711*** (0.025)	0.0704*** (0.020)
	Happy with actions	0.0061 (0.018)	0.0045 (0.014)
	12	0.0522** (0.020)	0.0551*** (0.016)
		0.0173 (0.025)	0.0213 (0.020)
		0.0437 (0.028)	0.0272 (0.023)
	Accomplished all needed	0.0211 (0.020)	
	21	0.0063 (0.021)	
		0.0201 (0.023)	
		−0.0022 (0.026)	
	Achieved things	−0.0003 (0.014)	
	32	0.0007 (0.017)	
		−0.0393 (0.024)	
		−0.0331 (0.035)	
	Felt humiliated	−0.0373** (0.015)	
	33	−0.0143 (0.017)	
		−0.0200 (0.020)	
		−0.0138 (0.022)	
	Difficulty talking to people	−0.0280 (0.018)	
	10	−0.0297 (0.018)	
		−0.0294 (0.020)	
		−0.0135 (0.023)	
	Felt criticised	−0.0042 (0.017)	
	25	0.0110 (0.019)	
		0.0266 (0.020)	
		0.0196 (0.022)	
	Been irritable	−0.0292 (0.024)	−0.0307 (0.020)
	29	−0.0580** (0.024)	−0.0624*** (0.020)
		−0.0363 (0.024)	−0.0434** (0.020)
		−0.0124 (0.027)	−0.0301 (0.022)

TABLE 55 Regression analysis results: SF-6D and CORE-OM items (IAPT) (*continued*)

Dimensions	Items (item number below)	SF-6D CORE items (SE)	SF-6D CORE significant items (SE)
Risk	Suicidal 1	0.0092 (0.019)	
	16	−0.0427* (0.023)	
		0.0147 (0.035)	
		0.0158 (0.050)	
	Suicidal 2	0.0038 (0.016)	
	24	0.0128 (0.019)	
		0.0205 (0.024)	
		0.0113 (0.027)	
	Self-harm	0.0087 (0.023)	
	34	0.0079 (0.025)	
		−0.0069 (0.033)	
		−0.0016 (0.031)	
	Been physically violent to others	−0.0240 (0.021)	
	6	−0.0337 (0.026)	
		0.0161 (0.047)	
		−0.0602 (0.043)	
	Felt threatened	0.0438** (0.018)	0.0399*** (0.015)
	22	0.0363 (0.023)	0.0390** (0.018)
		0.0223 (0.026)	0.0112 (0.020)
		0.0084 (0.036)	−0.0045 (0.028)
	Constant	0.7671*** (0.051)	0.6885*** (0.030)
	Observations	394	394
	R^2	0.714	0.609
	Adjusted R^2	0.576	0.574
	RMSE	0.054	0.062

SE, standard error.
***$p < 0.01$; **$p < 0.05$; *$p < 0.1$.

Appendix 4 *Chapter 5* appendices

TABLE 56 Keyword search strategies

First iteration searches

MEDLINE searches – QoL terms, October 2009

1. mental health.ti.
2. mental$ ill$.ti.
3. mental$ disorder$.ti.
4. 1 or 3 or 2
5. quality of life.ti.
6. 4 and 5

1. mental health.ti.
2. mental$ disorder$.ab,ti.
3. mental$ ill$.ab,ti.
4. 1 or 3 or 2
5. exp Mental Disorders/
6. 4 or 5
7. quality of life.ab,ti.
8. "Quality of Life"/
9. 8 or 7
10. 6 and 9
11. qualitative research/
12. qualitative.tw.
13. 11 or 12
14. 10 and 13

MEDLINE searches – concepts of possible relevance to QoL, October 2009

1. coping.ti.
2. control.ti.
3. positive attitude$.ti.
4. confidence.ti.
5. anxiety.ti.
6. self-esteem.ti.
7. self-respect.ti.
8. morale.ti.
9. resilience.ti.
10. humo?r.ti.
11. hope$.ti.
12. stress$.ti.
13. (worry or worries).ti.
14. fear$.ti.
15. self-determin$.ti.
16. or/1-15
17. mental$.ti.
18. 16 and 17
19. interview$.mp.
20. experience$.mp.
21. qualitative.tw.
22. qualitative research/
23. or/19-22
24. 18 and 23

1. energy.ti.
2. tired$.ti.
3. physical health.ti.
4. symptom$.ti.
5. hallucinat$.ti.
6. agitat$.ti.
7. paranoi$.ti.
8. (depression or depressive).ti.
9. disinhibit$.ti.
10. ill health.ti.
11. side effect$.ti.
12. sex drive.ti.
13. weight gain.ti.
14. dry mouth.ti.
15. sleep difficult$.ti.
16. substance misuse.ti.
17. substance abuse.ti.
18. gambl$.ti.
19. self harm$.ti.
20. or/1-19
21. mental$.ti.
22. 21 and 20
23. interview$.mp.
24. experience$.mp.
25. qualitative.tw.
26. qualitative research/
27. or/23-26
28. 27 and 22

1. opportunit$.ti.
2. (employment or employed or unemployed or unemployment).ti.
3. occupation$.ti.
4. social activit$.ti.
5. leisure activit$.ti.
6. structur$.ti.
7. routine$.ti.
8. social network$.ti.
9. satisfact$ service$.ti.
10. choice$.ti.
11. autonom$.ti.
12. freedom.ti.
13. responsibilit$.ti.
14. season$.ti.
15. individual$ care.ti.
16. education.ti.
17. money.ti.
18. financ$.ti.
19. income.ti.
20. living condition$.ti.
21. neighbour$.ti.
22. plan$ ahead.ti.
23. uncertain$.ti.
24. or/1-23
25. mental$.ti.
26. 25 and 24
27. interview$.mp.
28. experience$.mp.
29. qualitative.tw.
30. qualitative research/
31. or/27-30
32. 26 and 31

1. stigma$.ti.
2. attitude$.ti.
3. useful$.ti.
4. helpful$.ti.
5. role$.ti.
6. isolat$.ti.
7. alienat$.ti.
8. contact$.ti.
9. avoid$.ti.
10. enjoy$.ti.
11. support$.ti.
12. relationship$.ti.
13. belong$.ti.
14. camaraderie.ti.
15. connect$.ti.
16. (companion$ or company).ti
17. or/1-16
18. mental$.ti.
19. 18 and 17
20. interview$.mp.
21. experience$.mp.
22. qualitative.tw.
23. qualitative research/
24. 22 or 21 or 23 or 2024 and 19

TABLE 56 Keyword search strategies (*continued*)

Second iteration searches

Multiple database searches (ASSIA, CINAHL, PsycINFO, Web of Science) – QoL terms, January 2010

ASSIA

((TI=(mental health)) or(TI=(mental$ ill$)) or (TI=(mental$ disorder$))) and(TI=(quality of life))

(((TI=(mental health)) or(TI=(mental$ ill$) or AB=(mental$ ill$)) or(TI=(mental$ disorder$) or AB=(mental$ disorder$)) or (DE=("psychiatric disorders" or "adjustment disorder" or "affective disorders" or "organic mood syndrome" or "restlessness" or "seasonal affective disorders" or "akathisia" or "alexithymia" or "anxiety disorders" or "acute stress disorder" or "combat disorders" or "generalized anxiety disorders" or "panic disorders" or "nocturnal panic disorder" or "stage fright" or "phobias" or "acrophobia" or "agoraphobia" or "animal phobias" or "anthropophobia" or "claustrophobia" or "dental phobia" or "dysmorphophobia" or "erotophobia" or "school phobia" or "snake phobia" or "social phobia" or "spider phobia" or "weight phobia" or "posttraumatic stress disorder" or "chronic posttraumatic stress disorder" or "combat related posttraumatic stress disorder" or "postabortion syndrome" or "separation anxiety" or "childhood separation anxiety" or "attachment disorders" or "behaviour disorders" or "attention deficit disorder" or "attention deficit hyperactivity disorder" or "compulsive buying" or "compulsive foraging behaviour" or "conduct disorders" or "disruptive behaviour disorders" or "head banging" or "oppositional defiant disorder" or "cenesthopathy" or "character disorders" or "chronic psychiatric disorders" or "communication disorders" or "autism" or "infantile autism" or "selective mutism" or "confusional states" or "conversion disorder" or "delusional disorders" or "capgras syndrome" or "cotard s syndrome" or "fregoli syndrome" or "litigious delusional disorders" or "demonomania" or "cacodemonomania" or "emotional disorders" or "impulse control disorders" or "insanity" or "koro" or "mania" or "hypomania" or "mass psychogenic illness" or "mental illness" or "neuroticism" or "psychoticism" or "movement disorders" or "neuroleptic malignant syndrome" or "neuroses" or "depersonalization disorder" or "dissociative disorders" or "hypochondriasis" or "neurasthenia" or "obsessive compulsive neuroses" or "transference neuroses" or "personality disorders" or "antisocial personality disorder" or "avoidant personality disorders" or "borderline personality disorder" or "dependent personality" or "depressive personality disorders" or "gender identity disorder" or "histrionic personality disorder" or "identity crisis" or "kleptomania" or "multi impulsive personality disorder" or "multiple personality disorder" or "narcissistic personality disorder" or "passive aggressive personality disorder" or "sadistic personality disorder" or "schizotypal personality disorders" or "selfdefeating personality disorder" or "pervasive developmental disorders" or "asperger s syndrome" or "autistic spectrum disorders" or "childhood disintegrative disorder" or "heller s syndrome" or "rett syndrome" or "pica" or "coprophagia" or "psychogenic aspects" or "psychogenic polydipsia" or "psychoses" or "affective psychoses" or "anhedonia" or "bipolar affective disorder" or "cycloid psychosis" or "depression" or "childhood depression" or "death depression" or "delusional depression" or "maternal depression" or "melancholia" or "parental depression" or "paternal depression" or "postnatal depression" or "refractory depression" or "vascular depression" or "paranoid states" or "alcoholic psychoses" or "korsakoff s syndrome" or "mood incongruent psychoses" or "paranoia" or "querulous paranoia" or "shared paranoid disorder" or "folie a deux" or "puerperal psychosis" or "schizophrenia" or "catatonia" or "chronic schizophrenia" or "paranoid schizophrenia" or "paraphrenia" or "schizophreniform disorder" or "unipolar disorders" or "psychotic mood disorders" or "schizoaffective disorder" or "somatoform disorders" or "body dysmorphic disorder" or "briquet s syndrome" or "somatization disorders" or "thought disorder"))) and((TI=(quality of life) or AB=(quality of life)) or(DE="quality of life"))) and((TI=qualitative or AB=qualitative) or(DE="qualitative research") or(DE="qualitative methods") or(DE="qualitative data") or(DE="qualitative analysis"))

TABLE 56 Keyword search strategies (*continued*)

Second iteration searches

CINAHL

S6 S4 and S5
S5 TI quality of life
S4 S1 or S2 or S3
S3 TI mental* disorder*
S2 TI mental* ill*
S1 TI mental health

S13 S9 and S12
S12 S10 or S11
S11 (MH "Qualitative Studies+")
S10 TI qualitative or AB qualitative
S9 S5 and S8
S8 S6 or S7
S7 (MH "Quality of Life+")
S6 TI quality of life or AB quality of life
S5 S1 or S2 or S3 or S4
S4 (MH "Mental Disorders+")
S3 TI mental* disorder* or AB mental* disorder*
S2 TI mental* ill* or AB mental* ill*
S1 TI mental health

PsycINFO

1. mental health.ti.
2. mental$ ill$.ti.
3. mental$ disorder$.ti.
4. 1 or 2 or 3
5. quality of life.ti.
6. 4 and 5

1. mental health.ti.
2. mental$ ill$.tw.
3. mental$ disorder$.tw.
4. exp mental disorders/
5. 1 or 2 or 3 or 4
6. quality of life.tw.
7. exp "quality of life"/
8. 6 or 7
9. 5 and 8
10. qualitative.tw.
11. qualitative research/
12. 10 or 11
13. 9 and 12

Web of Science

\# 6 284 #5 AND #4
\# 5 36,217 ti="quality of life"
\# 4 44,097 #3 OR #2 OR #1
\# 3 5,600 ti="mental* disorder*"
\# 2 10,629 ti="mental* ill*"
\# 1 28,355 ti="mental health"

\# 20 #19 OR #18 Timespan=All Years
\# 19 #16 AND #14 Timespan=All Years
\# 18 #15 AND #14 Timespan=All Years
\# 17 ti=qualitative Timespan=All Years
\# 16 ts=qualitative Timespan=1900-2000
\# 15 ts=qualitative Timespan=2001-2010
\# 14 #13 OR #12 Timespan=All Years
\# 13 #11 AND #4 Timespan=All Years
\# 12 #5 AND #4 Timespan=All Years
\# 11 #9 OR #8 OR #7 OR #6 Timespan=All Years
\# 10 ti="mental health" Timespan=All Years
\# 9 ts="quality of life" Timespan=1900-1970
\# 8 ts="quality of life" Timespan=1971-1980
\# 7 ts="quality of life" Timespan=1981-1990
\# 6 ts="quality of life" Timespan=1991-2000
\# 5 ts="quality of life" Timespan=2001-2010
\# 4 #3 OR #2 OR #1 Timespan=All Years
\# 3 ts="mental* disorder*" Timespan=All Years
\# 2 ts="mental* ill*" Timespan=All Years
\# 1 Title=(mental health) Timespan=All Years

Hand-searching and citation searching using key references identified as being of potential relevance through experts and through first iteration searches.

TABLE 56 Keyword search strategies (*continued*)

Multiple database searches (ASSIA, CINAHL, MEDLINE, PsycINFO, Web of Science) – quasi or indirectly related QoL terms, April 2010

ASSIA

(TI=(recovery or (lived experience) or (subjective experience)) or TI=(coping or adaptation or (life functioning)) or TI=((life changes) or (life satisfaction) or wellbeing) or TI=(well being)) and (TI=((mental health) or (mental$ ill$) or (mental$ disorder$)))

((TI=(recovery or (lived experience) or (subjective experience)) or TI= (coping or adaptation or (life functioning)) or TI=((life changes) or (life satisfaction) or wellbeing) or TI=((well being) or lifestyle)) or(AB= (recovery or (lived experience) or (subjective experience)) or AB= (coping or adaptation or (life functioning)) or AB=((life changes) or (life satisfaction) or wellbeing) or AB=((well being) or lifestyle)) or ((DE="recovery") or(DE="personal experiences") or (DE="experiences") or(DE="life experiences") or(DE="coping") or (DE="coping strategies") or(DE=("adaptation" or "cognitive adaptation")) or(DE="adaptability") or(DE="life changes") or(DE="life changes") or(DE=("wellbeing" or "emotional wellbeing" or "social wellbeing" or "spiritual wellbeing" or "subjective wellbeing")) or(DE= ("psychological wellbeing" or "sense of coherence")) or (DE="lifestyle"))) and((TI=(mental health)) or(TI=(mental$ ill$) or AB= (mental$ ill$)) or(TI=(mental$ disorder$) or AB=(mental$ disorder$)) or (DE=("psychiatric disorders" or "adjustment disorder" or "affective disorders" or "organic mood syndrome" or "restlessness" or "seasonal affective disorders" or "akathisia" or "alexithymia" or "anxiety disorders" or "acute stress disorder" or "combat disorders" or "generalized anxiety disorders" or "panic disorders" or "nocturnal panic disorder" or "stage fright" or "phobias" or "acrophobia" or "agoraphobia" or "animal phobias" or "anthropophobia" or "claustrophobia" or "dental phobia" or "dysmorphophobia" or "erotophobia" or "school phobia" or "snake phobia" or "social phobia" or "spider phobia" or "weight phobia" or "posttraumatic stress disorder" or "chronic posttraumatic stress disorder" or "combat related posttraumatic stress disorder" or "postabortion syndrome" or "separation anxiety" or "childhood separation anxiety" or "attachment disorders" or "behaviour disorders" or "attention deficit disorder" or "attention deficit hyperactivity disorder" or "compulsive buying" or "compulsive foraging behaviour" or "conduct disorders" or "disruptive behaviour disorders" or "head banging" or "oppositional defiant disorder" or "cenesthopathy" or "character disorders" or "chronic psychiatric disorders" or "communication disorders" or "autism" or "infantile autism" or "selective mutism" or "confusional states" or "conversion disorder" or "delusional disorders" or "capgras syndrome" or "cotard s syndrome" or "fregoli syndrome" or "litigious delusional disorders" or "demonomania" or "cacodemonomania" or "emotional disorders" or "impulse control disorders" or "insanity" or "koro" or "mania" or "hypomania" or "mass psychogenic illness" or "mental illness" or "neuroticism" or "psychoticism" or "movement disorders" or "neuroleptic malignant syndrome" or "neuroses" or "depersonalization disorder" or "dissociative disorders" or "hypochondriasis" or "neurasthenia" or "obsessive compulsive neuroses" or "transference neuroses" or "personality disorders" or "antisocial personality disorder" or "avoidant personality disorders" or "borderline personality disorder" or "dependent personality" or "depressive personality disorders" or "gender identity disorder" or "histrionic personality disorder" or "identity crisis" or "kleptomania" or "multi impulsive personality disorder" or "multiple personality disorder" or "narcissistic personality disorder" or "passive aggressive personality disorder" or "sadistic personality disorder" or "schizotypal personality disorders" or "selfdefeating personality disorder" or "pervasive developmental disorders" or "asperger s syndrome" or "autistic spectrum disorders" or "childhood disintegrative disorder" or "heller s syndrome" or "rett syndrome" or "pica" or "coprophagia" or "psychogenic aspects" or "psychogenic polydipsia" or "psychoses" or "affective psychoses" or "anhedonia" or "bipolar affective disorder" or "cycloid psychosis" or "depression" or "childhood depression" or "death depression" or

TABLE 56 Keyword search strategies (*continued*)

Third iteration searches

| | "delusional depression" or "maternal depression" or "melancholia" or "parental depression" or "paternal depression" or "postnatal depression" or "refractory depression" or "vascular depression" or "paranoid states" or "alcoholic psychoses" or "korsakoff s syndrome" or "mood incongruent psychoses" or "paranoia" or "querulous paranoia" or "shared paranoid disorder" or "folie a deux" or "puerperal psychosis" or "schizophrenia" or "catatonia" or "chronic schizophrenia" or "paranoid schizophrenia" or "paraphrenia" or "schizophreniform disorder" or "unipolar disorders" or "psychotic mood disorders" or "schizoaffective disorder" or "somatoform disorders" or "body dysmorphic disorder" or "briquet s syndrome" or "somatization disorders" or "thought disorder"))) and((TI=qualitative or AB=qualitative) or(DE="qualitative research") or(DE="qualitative methods") or(DE="qualitative data") or(DE="qualitative analysis")) |

CINAHL

S17 S12 and S16
S16 S13 or S14 or S15
S15 TI mental* disorder*
S14 TI mental* ill*
S13 TI mental health
S12 S1 or S2 or S3 or S4 or S5 or S6 or S7 or S8 or S9 or S10 or S11
S11 TI lifestyle
S10 TI well being
S9 TI wellbeing
S8 TI life satisfaction
S7 TI life change*
S6 TI life functioning
S5 TI adaptation
S4 TI coping
S3 TI subjective experience*
S2 TI lived experience
S1 TI recovery

S30 S21 and S26 and S29
S29 S27 or S28
S28 (MH "Qualitative Studies+")
S27 TI qualitative or AB qualitative
S26 S22 or S23 or S24 or S25
S25 (MH "Mental Disorders+")
S24 TI mental* disorder* or AB mental* disorder*
S23 TI mental* ill* or AB mental* ill*
S22 TI mental health
S21 S1 or S2 or S3 or S4 or S5 or S6 or S7 or S8 or S9 or S10 or S11 or S12 or S13 or S14 or S15 or S16 or S17 or S18 or S19 or S20
S20 (MH "Life Style Changes")
S19 (MH "Life Style")
S18 (MH "Psychological Well-Being")
S17 (MH "Personal Satisfaction")
S16 (MH "Attitude to Life")
S15 (MH "Life Change Events")
S14 (MH "Adaptation, Psychological")
S13 (MH "Coping")
S12 (MH "Life Experiences")
S11 (MH "Recovery")
S10 TI lifestyle or AB lifestyle
S9 TI wellbeing or AB wellbeing or TI well being or AB well being
S8 TI life satisfaction or AB life satisfaction
S7 TI life change* or AB life change*
S6 TI life functioning or AB life functioning
S5 TI adaptation or AB adaptation
S4 TI coping or AB coping
S3 TI subjective experience* or AB subjective experience*
S2 TI lived experience or AB lived experience
S1 TI recovery or AB recovery

MEDLINE

1. recovery.ti.
2. lived experience.ti.
3. subjective experience.ti.
4. coping.ti.
5. adaptation.ti.
6. life functioning.ti.
7. life change$.ti.
8. life satisfaction.ti.
9. well being.ti.
10. wellbeing.ti.
11. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10
12. mental health.ti.

1. recovery.tw.
2. lived experience.tw.
3. subjective experience.tw.
4. coping.tw.
5. adaptation.tw.
6. life functioning.tw.
7. life change$.tw.
8. life satisfaction.tw.
9. well being.tw.
10. wellbeing.tw.
11. Adaptation, Psychological/
12. Life Change Events/
13. personal satisfaction/

TABLE 56 Keyword search strategies (*continued*)

Third iteration searches

13. mental$ ill$.ti.
14. mental$ disorder$.ti.
15. 12 or 14 or 13
16. 11 and 15

14. or/1-13
15. mental$ disorder$.ab,ti.
16. mental$ ill$.ab,ti.
17. exp Mental Disorders/
18. mental health.ti.
19. 18 or 16 or 15
20. 19 or 17
21. qualitative research/
22. qualitative.tw.
23. 21 or 22
24. 14 and 20 and 23

PsycINFO

1. recovery.ti.
2. lived experience.ti.
3. subjective experience.ti.
4. coping.ti.
5. adaptation.ti.
6. life functioning.ti.
7. life change$.ti.
8. life satisfaction.ti.
9. well being.ti.
10. wellbeing.ti.
11. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10
12. mental health.ti.
13. mental$ ill$.ti.
14. mental$ disorder$.ti.
15. 12 or 13 or 14
16. 11 and 15

1. recovery.tw.
2. lived experience.tw.
3. subjective experience.tw.
4. coping.tw.
5. adaptation.tw.
6. life functioning.tw.
7. life change$.tw.
8. life satisfaction.tw.
9. well being.tw.
10. wellbeing.tw.
11. "recovery (disorders)"/
12. "experiences (events)"/
13. coping behavior/
14. adjustment/
15. life changes/
16. life satisfaction/
17. well being/
18. or/1-17
19. mental health.ti.
20. mental$ ill$.tw.
21. mental$ disorder$.tw.
22. exp mental disorders/
23. 19 or 20 or 21 or 22
24. qualitative.tw.
25. qualitative research/
26. 24 or 25
27. 18 and 23 and 26

Web of Science

25 #24 OR #23 OR #22 OR #21 OR #20 OR #19 OR #18 OR #17 OR #16 OR #15
24 #14 AND #10
23 #14 AND #9
22 #14 AND #8
21 #14 AND #7
20 #14 AND #6
19 #14 AND #5
18 #14 AND #4
17 #14 AND #3
16 #14 AND #2
15 #14 AND #1
14 #13 OR #12 OR #11
13 ti="mental* disorder*"
12 ti="mental* ill*"
11 ti="mental health"
10 ti="well being"
9 ti=wellbeing

17 #16 AND #15 AND #11 Timespan=2009-2010
16 ts=qualitative Timespan=2009-2010
15 #14 OR #13 OR #12 Timespan=2009-2010
14 ts="mental* disorder*" Timespan=2009-2010
13 ts="mental* ill*" Timespan=2009-2010
12 ti="mental health" Timespan=2009-2010
11 #10 OR #9 OR #8 OR #7 OR #6 OR #5 OR #4 OR #3 OR #2 OR #1 Timespan=2009-2010
10 ts=wellbeing Timespan=2009-2010
9 ts="life satisfaction" Timespan=2009-2010
8 ts="life change*" Timespan=2009-2010
7 ts="life functioning" Timespan=2009-2010
6 ts=adaptation Timespan=2009-2010
5 ts=coping Timespan=2009-2010
4 ts="subjective experience*" Timespan=2009-2010
3 ts="subjective experience" Timespan=2009-2010
2 ts="lived experience" Timespan=2009-2010
1 ts=recovery Timespan=2009-2010

TABLE 56 Keyword search strategies (*continued*)

Third iteration searches
8 ti="life satisfaction"
7 ti="life change*"
6 ti="life functioning"
5 ti=adaptation
4 ti=coping
3 ti="subjective experience*"
2 ti="lived experience"
1 ti=recovery

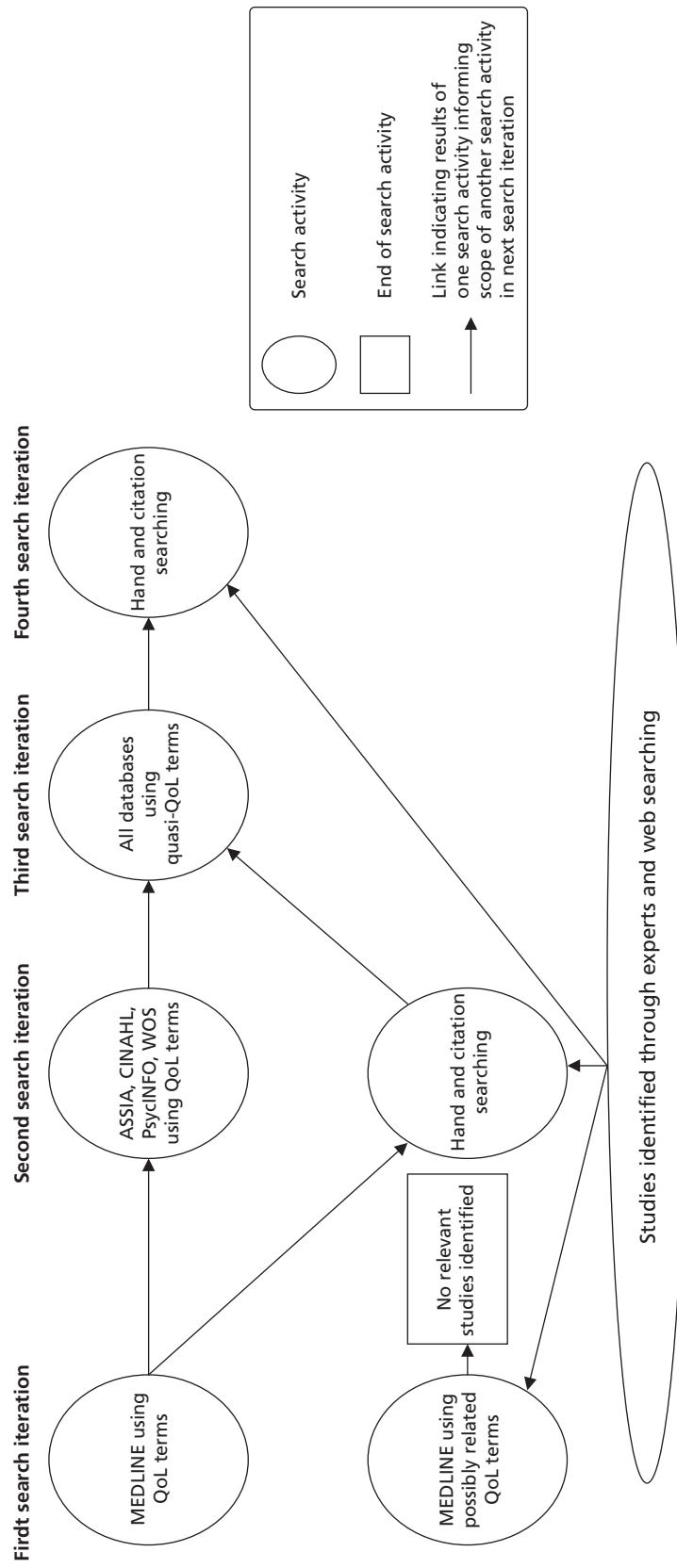| Fourth iteration searches |
| --- |
| **Hand-searching and citation searching using key references identified as being of potential relevance through experts and through third iteration searches.** |

QoL, quality of life.

FIGURE 12 Summary of search iterations. QoL, quality of life; WOS, Web of Science.

TABLE 57 Summary of qualitative review studies

First author and year	Affiliation of first author	Country	Participants	Recruited from	Diagnosis/mental health problem
Cook (2009)[Chambers 2012,208]	Occupational therapy	UK	16 males 8 females	CMHTs	Schizophrenia $n=17$; bipolar disorder $n=4$; other $n=4$
Corring (2007)[206,209]	Occupational therapy	Canada	25 males 28 females	Mental Health Association programmes and Consumer Survivor agency	Schizophrenia $n=27$; schizoaffective disorder $n=7$; bipolar disorder $n=4$; depression $n=9$; PTSD $n=4$; anxiety disorder $n=4$; borderline personality disorder $n=4$; don't know/no response $n=5$
Fisher (1998)[210]	Nursing	Canada	$n=24$	Acute inpatients	Diagnoses included major depression; bipolar disorder; affective disorder; depression/anxiety; schizoaffective disorder; psychosis; sociopathic personality disorder
Gee (2003)[211]	Psychology	UK	3 males 3 females	CMHTs/acute inpatients	Schizophrenia
Gould (2005)[212]	Occupational therapy	Canada	4 males	Poster campaign	Schizophrenia or other psychotic disorder
Hamer (2009)[213]	Nursing – general and psychiatric	New Zealand	10 females	Private psychology and psychotherapy clinics	Panic disorder
Hedberg (2009)[214]	Nursing/ psychiatry	Sweden	6 males 6 females	Outpatients	No diagnosis – 'psychiatric disability'
Laliberte-Rudman (2000)[215]	Occupational therapy	Canada	25 males 10 females	Outpatients	Schizophrenia – DSM-IV diagnostic criteria
Lim (2004)[216]	Psychology	Australia	$n=18$	Outpatients	Bipolar disorder
Mayers (2000)[217]	Occupational therapy	UK	10 males 1 female	Via occupational therapists	No diagnosis – 'enduring mental health problems'
Michalak (2006)[207,218]	Psychiatry	Canada	12 males 23 females	Outpatients and inpatients	Bipolar disorder
Rusner (2010)[219]	Social work	Sweden	6 males 4 females	Outpatients and advocacy group	Bipolar disorder
Vellenga (1994)[220]	Nursing	USA	15 males	Outpatients	Schizophrenia $n=8$; bipolar $n=4$; schizoaffective disorder $n=3$

Appendix 5 *Chapter 6* appendices

DRAFT TOPIC GUIDE 4 MARCH 2010

Objectives

- How/if quality of life is affected by mental health difficulties
- To establish the domains of quality of life that are important to people with mental health problems

Introduction

- Introduce self
- Aims: What affects quality of life for people with mental health difficulties?
- Proposed structure of interview (start with own perceptions – warn questions will be deliberately vague, give reason – will move on to areas highlighted in previous research if not already discussed – complete and give opinions on current QoL questionnaires)
- Stress that we are interested in what is important to them (not about what is important to quality of life more generally); there are no right or wrong answers, it is their perspective that we are interested in
- They are not obliged to answer; if they don't want to talk about a subject, to please say so, and will move on
- They can stop the recording of the interview at any time
- They can stop the interview at any time if they do not wish to continue
- Interview confidential and anything used will be completely anonymous

Interview

SECTION A
Explore quality of life from their own perspective without giving any prompts

Background
Tell me a bit about yourself

- How do they spend their time?
- How is their health (mental and physical)

Quality of Life
What are the things that make life good/bad for you?

- **Prompts if primary question does not work** (example probing questions, don't expect to ask them all)

  - What is important/unimportant to you?
  - What would you change to make your life better?
  - What would make your life worse?
  - Has anything stopped/enabled you doing the things you want to do?
  - Has there been a time in your life when you feel your life was better/worse than it is now? What is it about these times that has made a difference?
  - What do you think are the most important of the things you have mentioned and why?

What do you/don't you enjoy?

Mental Health (pick up on anything said in response to above questions – use their terminology when referring to mental health difficulties)

- Does (their mental health) affect the different areas of their life
- Directly (e.g. symptoms/medication) and/or indirectly (relationships/activities/stigma etc.)

SECTION B

Domains from literature review and EQ-5D and SF-36

Once their own perceptions have been exhausted introduce concepts that research has indicated might be important to people with mental health difficulties that have **not** already been explored. Explore how each helps and hinders. Explore what it is about this aspect of their life that is important to them; give suggestions if person unable to explain further e.g. Work: being active, meeting people, money, self-esteem, purpose, structure, routine etc.

Relationships – being with people, having a partner/friends, socialising, support

Work/occupation/study (include voluntary etc.)

Leisure, spare time, hobbies

Mental health symptoms/experience

Medication/side effects

Physical health/pain/mobility

Self-esteem/confidence

Stigma

Mental health services/practitioners (how have helped/hindered)

Finances/accommodation/basic needs

Spirituality

**EME
HS&DR
HTA
PGfAR
PHR**

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

Published by the NIHR Journals Library