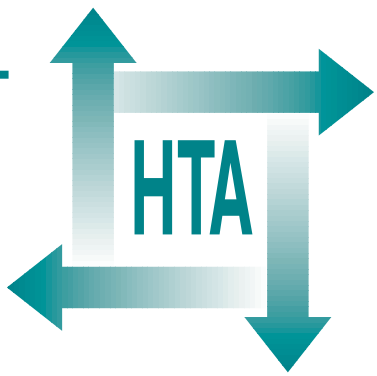
Health Technology Assessment 2001; Vol. 5: No. 6

Methodology

General health status measures for people with cognitive impairment: learning disability and acquired brain injury

RP Riemsma
CA Forbes
JM Glanville
AJ Eastwood
J Kleijnen

Health Technology Assessment
NHS R&D HTA Programme



How to obtain copies of this and other HTA Programme reports.
An electronic version of this publication, in Adobe Acrobat format, is available for downloading free of charge for personal use from the HTA website (http://www.hta.ac.uk). A fully searchable CD-ROM is also available (see below).

Printed copies of HTA monographs cost £20 each (post and packing free in the UK) to both public **and** private sector purchasers from our Despatch Agents.

Non-UK purchasers will have to pay a small fee for post and packing. For European countries the cost is £2 per monograph and for the rest of the world £3 per monograph.

You can order HTA monographs from our Despatch Agents:

– fax (with **credit card** or **official purchase order**)
– post (with **credit card** or **official purchase order** or **cheque**)
– phone during office hours (**credit card** only).

Additionally the HTA website allows you **either** to pay securely by credit card **or** to print out your order and then post or fax it.

Contact details are as follows:

HTA Despatch
c/o Direct Mail Works Ltd
4 Oakwood Business Centre
Downley, HAVANT PO9 2NP, UK

Email: orders@hta.ac.uk
Tel: 02392 492 000
Fax: 02392 478 555
Fax from outside the UK: +44 2392 478 555

NHS libraries can subscribe free of charge. Public libraries can subscribe at a very reduced cost of £100 for each volume (normally comprising 30–40 titles). The commercial subscription rate is £300 per volume. Please see our website for details. Subscriptions can only be purchased for the current or forthcoming volume.

Payment methods

Paying by cheque
If you pay by cheque, the cheque must be in **pounds sterling**, made payable to *Direct Mail Works Ltd* and drawn on a bank with a UK address.

Paying by credit card
The following cards are accepted by phone, fax, post or via the website ordering pages: Delta, Eurocard, Mastercard, Solo, Switch and Visa. We advise against sending credit card details in a plain email.

Paying by official purchase order
You can post or fax these, but they must be from public bodies (i.e. NHS or universities) within the UK. We cannot at present accept purchase orders from commercial companies or from outside the UK.

How do I get a copy of *HTA on CD*?

Please use the form on the HTA website (www.hta.ac.uk/htacd.htm). Or contact Direct Mail Works (see contact details above) by email, post, fax or phone. *HTA on CD* is currently free of charge worldwide.

The website also provides information about the HTA Programme and lists the membership of the various committees.

General health status measures for people with cognitive impairment: learning disability and acquired brain injury

RP Riemsma*
CA Forbes
JM Glanville
AJ Eastwood
J Kleijnen

NHS Centre for Reviews and Dissemination, University of York, York, UK

* Corresponding author

Competing interests: none declared

Published March 2001

This report should be referenced as follows:

Riemsma RP, Forbes CA, Glanville JM, Eastwood AJ, Kleijnen J. General health status measures for people with cognitive impairment: learning disability and acquired brain injury. *Health Technol Assess* 2001;**5**(6).

Health Technology Assessment is indexed in *Index Medicus*/MEDLINE and *Excerpta Medica*/EMBASE. Copies of the Executive Summaries are available from the NCCHTA website (see opposite).

NHS R&D HTA Programme

The NHS R&D Health Technology Assessment (HTA) Programme was set up in 1993 to ensure that high-quality research information on the costs, effectiveness and broader impact of health technologies is produced in the most efficient way for those who use, manage and provide care in the NHS.

Initially, six HTA panels (pharmaceuticals, acute sector, primary and community care, diagnostics and imaging, population screening, methodology) helped to set the research priorities for the HTA Programme. However, during the past few years there have been a number of changes in and around NHS R&D, such as the establishment of the National Institute for Clinical Excellence (NICE) and the creation of three new research programmes: Service Delivery and Organisation (SDO); New and Emerging Applications of Technology (NEAT); and the Methodology Programme.

Although the National Coordinating Centre for Health Technology Assessment (NCCHTA) commissions research on behalf of the Methodology Programme, it is the Methodology Group that now considers and advises the Methodology Programme Director on the best research projects to pursue.

The research reported in this monograph was funded as project number 97/17/99.

The views expressed in this publication are those of the authors and not necessarily those of the Methodology Programme, HTA Programme or the Department of Health. The editors wish to emphasise that funding and publication of this research by the NHS should not be taken as implicit support for any recommendations made by the authors.

Criteria for inclusion in the HTA monograph series
Reports are published in the HTA monograph series if (1) they have resulted from work commissioned for the HTA Programme, and (2) they are of a sufficiently high scientific quality as assessed by the referees 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.

The editors and publisher have tried to ensure the accuracy of this report but do not accept liability for damages or losses arising from material published in this report. They would like to thank the referees for their constructive comments on the draft document.

ISSN 1366-5278

Published by Core Research, Alton, on behalf of the NCCHTA.
Printed on acid-free paper in the UK by The Basingstoke Press, Basingstoke.

M

Contents

List of abbreviations and glossary

Technical terms and abbreviations are used throughout this report. The meaning is usually clear from the context but a glossary is provided for the non-specialist reader. In some cases usage differs in the literature but the term has a constant meaning throughout this review.

List of abbreviations

ABS	Adaptive Behaviour Scale
ADL	Activities of Daily Living*
BDI	Beck's Depression Inventory*
BICRO-39	Brain Injury Community Rehabilitation Outcome Scales (39-items)
CI	confidence interval*
CIQ	Community Integration Questionnaire*
ComQol-I5	Comprehensive Quality of Life Scale – Intellectual Disability
COOP/ WONCA	Dartmouth COOP functional health assessment charts/WONCA
CPD	chronic physical disorders
CQOL	Child Quality of Life Questionnaire
DIP	Disability and Impact Profile
EDSS	Expanded Disability Status Scale
EQ-5D	EuroQol – 5 dimensions
FAMS	Functional Assessment of Multiple Sclerosis
GOS	Glasgow Outcome Scale*
HADS	Hospital Anxiety and Depression Scale*
HRQL	Health-Related Quality of Life
HTA	Health Technology Assessment
HUI	Health Utilities Index
ICC	intraclass correlation coefficient
IQ	intelligence quotient
LD	learning disabilities
LLQ	Laman & Lankhorst Questionnaire
LSI	Life Satisfaction Index
LSS	Life Situation Survey
MMSE	Mini Mental State Examination
MOS	Medical Outcome Studies
MRC	Medical Research Council
MS	multiple sclerosis
MS-QLI	Multiple Sclerosis-Quality of Life Interview
MSQOL -54	Multiple Sclerosis Quality of Life 54-item Scale
NHP	Nottingham Health Profile
Oregon QLQ	Oregon Quality of Life Questionnaire (Interviewer Rating Version)
OR	odds ratio*
PASSING	Program Analysis Service System Implementation of Normalisation Goals
PCRS	Patient Competency Rating Scale
PD	psychiatric disorders
QALY	quality-adjusted life-year
QoL	quality of life
QOL-Q	Quality of Life – Questionnaire
Q-TWiST	Extended Quality-Adjusted Time Without Symptoms and Toxicity
QUAL-OT	Quality of Life Scale – Occupational Therapist
QUOLIS	Quality of Life Scale
QWB	Quality of Well-Being Index
RCAS	Resident Choice Assessment Scale
RCT	randomised controlled trial*

continued

continued

SD	standard deviation*
SF-36	Medical Outcome Studies-Short Form-36
SIP	Sickness Impact Profile
SSQ	Soweto Stroke Questionnaire
SWB	Subjective Well-Being
TBI	traumatic brain injury*
VAS	visual analogue scale

*Used only in tables

Glossary

***Cognitive impairment** For the purpose of this review the term 'cognitive impairment' only includes impairments due to acquired brain injury (caused by stroke, trauma, MS or other causes) or LD. LD was once described as 'mental handicap' and is also referred to as 'mental retardation' in other countries. However, in the review the term 'learning disability' refers to a particular state of functioning that begins in childhood and is characterised by limitations in both intelligence and adaptive skills.

Concurrent validity The extent to which instrument results agree with an independent external criterion, that is, an independent measure of the same variable as that which the instrument is investigating.

Construct validity Refers to whether the instrument results are in accord with present theories concerning the relevant areas of research. For example, do the resultant symptom patterns concur with physiological theories or causality? Construct validity may be assessed by comparison of instrument data with the theories expounded in the relevant medical literature.

Content validity The extent to which the instrument adequately probes the various aspects of the area it is supposed to measure. This relates to the idea of 'completeness' and is usually tested by reference to clinical experience.

Convergent validity The extent to which the instrument relates to other variables and other measures of the same construct to which it should be related. For example, if our theory states that anxious people are supposed to be more aware of autonomic nervous system activity then scores on the instrument of anxiety should correlate with scores on a measure of autonomic awareness.

Criterion validity The extent to which the instrument correlates with some other measure of the trait or disorder under study, ideally, a 'gold standard' which has been used and accepted in the field. Criterion validity is usually divided into two types: **concurrent validity** (the criterion measure is assessed at the same time) and **predictive validity** (the criterion measure will not be available until some time in the future).

Cronbach's alpha A coefficient that provides a measure of inter-item consistency (internal consistency) within a scale, by describing how well a group of items focuses on a single idea or construct. Generally used for measures where subjects respond to questions on a scale (1 to 3, 1 to 4, 1 to 5, etc.), Cronbach's alpha can range between 0 and 1. If a scale has an alpha above 0.60, it is usually considered to be internally consistent. However, Cronbach's alpha does not indicate how well the scale has covered an idea. Thus, a scale may have a high alpha (i.e. 0.70 or higher) indicating that scale items focus on one construct, but the scale may have only partially covered the breadth of the construct.

Delphi method A method in which the opinion of experts in the field is sought. The Delphi method is dependent upon the judgement of knowledgeable experts and utilises repeated rounds of questioning, including feedback of earlier-round responses, to take advantage of group input while avoiding the biasing effects possible in face-to-face panel deliberations. In the Delphi method, panellists respond anonymously, preventing the identification of a specific opinion with any individual or company. This anonymity also provides the comfort of confidentiality, allowing panellists to freely express their opinions, and enabling previous responses to be revised in subsequent iterations after reviewing new information submitted by other panellists. All participants are encouraged to comment on their own forecasts and on the combined panel results.

continued

Continued

This procedure reduces the effects of personal agendas or biases and assists the panellists in remaining focused on the questions, issues and comments at hand.

Discriminant validity The extent to which the instrument does not correlate with dissimilar, unrelated constructs. If our theory states that anxiety is independent of intelligence, then we should not find a strong correlation between the two.

Ease of usage Refers to the time necessary to complete the instrument and the response rates, with and without missing values.

Face validity The extent to which the instrument looks as though it measures what it is supposed to measure.

*__General health status__ In this review these are defined as instruments that determine an individual's perceived level of physical and psychological functioning in relation to their lifestyle. General health status can include various domains of life, such as physical, psychological, economical and social functioning, but for the purpose of this review, measures should include at least the domains of physical functioning and psychological functioning. Other domains of life (e.g. occupational or interpersonal) may be included in the measure.

General health status measures Instruments that determine an individual's perceived level of physical and psychological functioning in relation to their lifestyle.

Homogeneity (synonym: internal consistency) Refers to the consistency of answers to items in the instrument and often reported using Cronbach's alpha.

*__Instrument and measure__ In this review, these terms are used interchangeably to describe the general health status measures evaluated.

Internal consistency (synonym: homogeneity) See homogeneity.

Inter-rater reliability The extent to which an instrument provides the same results between different observers or users of the instrument. Different raters interviewing the same respondents with a very short gap between interviews ideally measure this. The most appropriate measure of inter-rater reliability is the Kappa coefficient of agreement. Simple correlation analysis makes no allowance for chance agreement. In contrast, Kappa takes chance agreement into consideration and produces a coefficient between –1 and +1, negative values indicating levels of agreement worse than chance.

Learning disabilities (synonyms: mentally retarded, mentally disabled) A particular state of functioning that begins in childhood and is characterised by limitations in both intelligence and adaptive skills.

Patient–proxy agreement The extent to which agreement exists between measures completed by the patients themselves and proxies such as family members, carers, or health professionals.

Predictive validity The extent to which agreement exists between instrument results and an independent external criterion, which will not be available until some time in the future. For instance, a college admission test should correlate with the person's performance on graduation.

Reliability The extent to which a variable apparently fluctuates as a result of errors of measurement as opposed to real changes in the object of measurement itself. Basically there are two types of reliability: **internal consistency** and **repeatability** (test–retest reliability).

Repeatability (synonym: Test–retest reliability) See Test–retest reliability.

*__Scale and subscale__ In this review, these terms are only used to refer to the dimensions of domains within a measure.

Skewness The extent to which a scale does not have a symmetrical distribution. If the scale is skewed to the right (high values are over-represented) there is a 'ceiling effect', and if the scale is skewed to the left (low values are over-represented) there is a 'floor effect'.

Test–retest reliability The extent to which the instrument provides the same results on the same subjects on two or more occasions, with the same or another observer, the subject of the test being in the same state of health.

* Definitions of terms used in this report

Definitions of terms in the glossary have been based on the following publications: Feinstein AR. Clinimetrics. London: Yale University Press, 1987; Bennet AE, Ritchie K. Questionnaires in Medicine. Oxford: Nuffield Provincial Hospitals Trust, 1975; Wilkin D, Hallam L, Dogget M. Measures of need and outcome for primary health care. New York: Oxford Medical Publications, 1992; Streiner DL, Norman GR. Health measurement scales: a practical guide to their development and use. Oxford: Oxford University Press, 1995.

Executive summary

Background

Currently there is a wide range of health status measures that aim to assess general health status in people with cognitive impairment. However, the validity and/or applicability to this patient group are largely unknown. This has implications for the assessment of treatment outcomes and rehabilitation, for prognostic purposes, for planning services, and for determining the benefits and adverse effects of health technologies targeted at these patient groups.

Objectives

- To identify the general health status measures that have been validated in patients with cognitive impairment.
- To assess the extent to which these measures have been validated.
- To draw out the implications of the findings for the use of existing measures and for future primary research in this area.

Methods

Selection criteria

Studies that assessed general health status in people with cognitive impairment due to acquired brain injury (traumatic brain injury, cerebro-vascular accident or multiple sclerosis (MS)) or learning disability (LD) were included in the review. Studies that used general health status instruments measuring only one general health dimension, and studies that only featured participants with cognitive impairment due to dementia were excluded.

Search strategy

A wide range of relevant databases were searched for studies on cognitive impairment, general health status measures, and validation of health status measures. A handsearch of general health status bibliographies was also conducted.

Data were collected on the general health status measure used, the population characteristics, aims of the study, validity details, and conclusions.

Results

The review includes data from 71 studies, reported in 83 separate publications. In total 34 different general health status measures were described in the 83 publications, with the Sickness Impact Profile (SIP) and the Short Form-36 (SF-36) the most frequently used measures (20 and 19 studies, respectively). These studies included a total of 98 instrument validations, 52 of which definitely or probably included people with cognitive impairment. Six measures were extensively validated (quality scores ranged from 0.25 to 0.5, on a scale from 0 to 1) in studies in which more than 50% of the respondents were people with cognitive impairment. A further three measures were also validated in studies in which more than 50% of the respondents were people with cognitive impairment, but their level of validation was more limited (quality scores ranged from 0.1 to 0.2).

Five measures were validated in studies in which 20–50% of the respondents were cognitively impaired, which may limit their relevance to participants with cognitive impairment (quality scores ranged from 0.1 to 0.6). The SF-36 was also validated in two studies in which 20–50% of the respondents were cognitively impaired and the quality score was 0.3.

Finally, nine of the measures were only validated in studies in which less than 20% of the respondents were cognitively impaired. For these measures it was unclear whether the findings applied to people with cognitive impairment.

Conclusions

Very few measures have been validated specifically for cognitively impaired respondents. Studies where at least 50% of the respondents were cognitively impaired generally showed poorer validity results compared with studies with fewer cognitively impaired persons, indicating that general health status measures designed for the general population are not automatically suitable for people with cognitive impairment. The few measures that were specifically developed for people with cognitive impairment also reported

poor validity results. **Therefore, there are no validated instruments available for use in cognitively impaired respondents; existing measures, specifically designed for use in these populations, should be used with caution.**

The most promising measure is the MS-Quality of Life Interview (MS-QLI) for MS patients. The MS-QLI was thoroughly validated in 300 MS patients and the results were good, except for the 'social function' subscale. However, only 20–50% of the respondents in this study had cognitive impairment.

Most information on the validity of general health status measures was found in studies among people with LD. For these patients, six measures were found that have been validated in a populations where more than 50% of the respondents were cognitively impaired LD patients.

Implications for practice

- Existing general health status measures should be used with caution in individuals with cognitive impairments.
- There is no evidence to indicate the most suitable general health status measure for use in economic evaluations of cognitive impairment.
- There is little evidence to support the validity of proxy assessments in cognitively impaired populations.

Recommendations for further research

- There is a need for the development of new general health status measures for cognitively impaired people, particularly for people with acquired brain injury due to stroke, MS or trauma.
- Existing general health status measures, such as the SIP, SF-36, the EuroQol – 5 dimensions (EQ-5D) and the Nottingham Health Profile need to be validated for people with cognitive impairment. More research is needed into how these existing measures could be modified so that they are more suitable for people with cognitive impairment.
- Currently there are no general health status measures for cognitively impaired populations available for use in economic evaluations. In general, the EQ-5D and Health Utilities Index were found to be superior, compared with other preference-based measures of health. The validity of these instruments needs to be assessed in cognitively impaired populations, as well as the feasibility of using choice-based techniques in people with cognitive impairments.
- Health status measures need to be validated for use by proxies in certain populations.
- Validity assessment of general health status measurements for people with cognitive impairment should be addressed in studies specifically designed for this patient population.
- Objective validated psychometric tests or a neurologist's diagnosis should assess the level of cognitive impairment. Separate analyses should be performed to assess the validity of the instrument for different levels of cognitive impairment.
- Validity assessment of general health status measures should include information on the choice of component items, sensibility, consistency, accuracy and suitability. When there is need for proxy assessments the instrument should be assessed for patient–proxy agreement and inter-rater agreement.
- Studies should include a large number of respondents with different levels of cognitive impairment, so that differences in the measure's validity for different groups of people with cognitive impairment can be assessed.

Chapter 1
Background

The need for general health status measures

The measurement of general health status is important because it indicates the extent to which interventions really make a difference to a patient's overall life and helps to quantify the relative effects of different interventions for different classes of patients.[1] Physiological measures provide information to clinicians but are of limited interest to patients; they often correlate poorly with functional capacity and well-being, the areas which are most important to patients. Another reason to measure general health status is the commonly observed phenomenon that two patients with the same clinical criteria often have dramatically different responses to treatment.[2]

General health status measures, such as the Short Form-36 (SF-36), multidimensional indices of quality of life (QoL), and broad approaches to QoL in terms of quality-adjusted life-years (QALYs) are important for comparative studies aimed at informing health policy planning decisions regarding the allocation of scarce medical resources to different conditions and interventions.[1] The assessment of general health status in patients with cognitive impairment is difficult.

Traditionally, outcome assessment in this group has focused on the assessment of levels of impairment and disability, rather than on general health status or general well-being. The presence of cognitive and communication problems may have meant that other outcomes which may be of importance to these patients have been overlooked. Self-reported outcomes in particular may be difficult to obtain; as a result, relatively few studies in patients with cognitive impairment have assessed general health status. In trials which have done so, the diminution of general health status may be underestimated, either for the reasons given above or because existing measures that have been used are simply inappropriate, or are insensitive to changes in health-related functioning. In addition, studies assessing general health status in certain relevant patient or age groups have simply excluded patients with cognitive impairment, perhaps because they have difficulties in communicating. One way to measure general health status in such patients is to use proxies.

A relative, friend, the investigator, or any other individual such as a carer, can fill the role of proxy. Obviously proxy assessments may not accurately reflect the individual's own assessment of their general health status, but where individuals are too severely impaired to complete the assessment themselves, proxy assessments may be useful. The issue of proxy assessment is therefore important but still remains controversial, requiring further investigation.

In summary, although it is widely accepted that it is important to assess general health status in cognitively impaired patients, there are many unresolved issues, leading to questions about how assessment can be best achieved and what additional primary research is required.

Existing health status measures in people with cognitive impairment

While there is a range of general health assessment tools for assessing health status, and other measures that aim to assess this outcome in specific disease conditions, the extent to which they can provide valid indicators of outcome when applied to cognitively impaired populations is unclear. Cognitively impaired patients may differ in the domains of functioning that they see as important, and may differ in their outcome assessments. Often their health and overall functioning are difficult to assess because of the cognitive and communication problems from which they suffer. Consequently there are a number of studies that question the issue of whether cognitively impaired individuals can accurately report their own health.[3,4] Therefore, such assessments are more likely to be made on their behalf by others, for example using measures involving ratings by carers and health professionals. However, these ratings also pose problems through inaccurate reporting (i.e. due to a lack of carer introspection or bias in the carers perspectives arising from their own high levels of depression and anxiety).[5,6]

Previous research

In general, measures that have been used to date have tended to be directed at disability, such as

impairment of psychological and physical functioning. However, general health status and other domains of functioning (e.g. social, sexual, employment, activities of daily living and general well-being) have remained relatively under-investigated. In patients with cognitive impairment, the assessments of clinicians and caregivers have often been given greater prominence. In addition, in many patients with cognitive impairment other physical and mental health problems may co-exist, such as anxiety and depression, and pain. These may further affect an individual's health status and increase their overall disability.

The issue of whether general health status measures should be used in economic evaluations and whether that information could be used in making healthcare decisions are also important. A recent systematic review on the issue made a distinction between instruments that generate scores based on people's preferences (e.g. EuroQol – 5 dimensions (EQ-5D)) and instruments that use arbitrary scoring procedures (e.g. SF-36), identifying the former as more suitable for use in economic evaluations.[7] Such instruments can be used to generate health state utilities, which can be used to calculate QALYs, for use in cost–utility analyses.

Learning disability

To date the extent to which health status has been investigated for individuals with learning disability (LD) is somewhat variable. Some literature is available on the assessment of physical health and mental health (e.g. using the Psychiatric Assessment Schedule for Adults with Developmental Disability), but there has been little investigation of general health status. However, recently the Health of the Nation Outcome Scale was adapted for use in these patients. There have also been several recent studies in patients with LD in the UK, which have attempted to measure changes in general health status associated with moving from institutions into the community.[8–11] One of these was carried out in the context of an economic analysis.[11] These studies used a variety of methods to assess general health status, including life-experience checklists, questionnaires and interviews.

Acquired brain injury

There are a number of instruments that attempt to measure emotional and behavioural components of health following acquired brain injury. In the field of stroke, for example, there have been several recent studies assessing the use of the Health Utilities Index (HUI),[12] the EQ-5D[13] and the Sickness Impact Profile (SIP).[14] However, there are a number of problems with these studies, as the measures are often rather vaguely defined and not always validated for use within cognitively impaired populations.[15]

The need for a review of health status measures in patient with cognitive impairment

In summary there are a range of health status measures in current use, which aim to assess general health status in people with cognitive impairment. Although some studies have investigated these in some detail, the validity and/or applicability of most measures to patients with cognitive impairment appear to be largely unknown. However, this is an important issue for the assessment of outcomes of treatment and rehabilitation in these patients, for prognostic purposes, for planning services, and for determining the benefits and adverse effects of health technologies targeted at these groups. This review therefore responds to a need identified by the National Coordinating Centre for Health Technology Assessment (NCCHTA) for more information on existing health status measures. The review aims to assess the current state of knowledge in this area, to provide guidance regarding the current use of health status measures and to direct future research in this area.

According to the 1988 Survey of Disability (Office of Population Censuses and Surveys), 1,475,000 persons in Great Britain have disabilities in intellectual functioning, which accounts for approximately 3.4% of the total adult population. Data from a more recent longitudinal population study, showed that 11% of men and 19% of women in England and Wales, aged 65 years and over were disabled (1.3 million people in total); 38% of these were cognitively impaired.[16]

Cognitive impairment is most prominent in conditions like dementia, LD and acquired brain injuries. The Medical Research Council/Health Technology Assessment (MRC/HTA) Dementia Trials Development Group is currently undertaking a study into the use of general health status measures among people with dementia. Therefore studies that only looked at dementia are excluded from the current review. However, it is likely that recent reports of the use of health status measures in dementia will be of general relevance to this review. Such reports include those from the International Working Group on the Harmonization of Guidelines for Trials in Dementia and the MRC/HTA Dementia Trials Development

Group.[17] These are examined for issues of relevance to the assessment of health status in LD, and acquired brain injury.

Objectives

- To identify the general health status measures that have been used in cognitively impaired patients with LD and acquired brain injury.
- To assess the extent to which these measures have been validated for these patient groups.
- To look at the purposes for which these measures have been used, and how valid they are for these purposes.
- To identify general health status measures that have been used in cognitively impaired patients for use in economic evaluations, and to assess the extent to which these measures have been validated for these patient groups.
- To draw out the implications of the findings for the use of existing measures and for future primary research in this area.

Expert panel

A panel of experts was formed (see acknowledgements) to provide guidance on the scope of the review and to advise on both clinical and methodological issues. Panel members were chosen for their expertise in the fields of cognitive impairment, stroke, multiple sclerosis (MS), health status measurements, and clinimetrics.

Throughout the project the expert panel provided definitions for terms, such as 'cognitive impairment' and 'general health status measurement', in addition to helping devise the protocol, search strategy and the framework for the assessment of study validity.

Chapter 2
Methods

Included studies

Two reviewers independently assessed all of the studies for inclusion in the review. Any disagreements were resolved by discussion. In order to be included in the review the studies had to fulfil inclusion criteria for study type, study population and outcome measures.

Study type

- **Studies assessing general health status in people with cognitive impairment** In this review, general health status was defined as an individual's perceived level of physical and psychological functioning in relation to their lifestyle. General health status can include various domains of life, such as physical, psychological, economic and social functioning. Measures should include at least the domains of physical functioning and psychological functioning. Other domains of life (e.g. occupational or interpersonal) may be included in the measure. As pain includes both a physical and psychological dimension, measures were included that contain measures of pain and either physical or psychological functioning. Instruments measuring only one dimension of overall well-being or QoL were also included.

Disease-specific general health status measures are included, but only if the measure was aimed at measuring general health status, as described above. Finally, as non-validated measures are of little interest for this review, we only included measures that have been validated for at least one of the relevant outcomes (see below).

The following types of study were also included:

- **Studies mentioning third party reports on general health status in people with cognitive impairment** (e.g. proxy assessments by relatives or health professionals)
- **Studies mentioning economic perspectives on general health status in people with cognitive impairment** (e.g. valuation of health states)
- **Dementia studies if they also included the populations of interest** (i.e. acquired brain injury and LD).

Study population

Studies that assessed the validity of general health status measures in people with cognitive impairment due to acquired brain injury (caused by stroke, trauma, MS or other causes) and LD, were included in the review. LD was once described as 'mental handicap' and is described in other countries as 'mental retardation'. In this review LD is a particular state of functioning that begins in childhood and is characterised by limitations in both intelligence and adaptive skills. The search strategy takes account of the differences in terminology.

Outcome measures

The review only includes papers that assess or report issues relating to the reliability and validity of general health status measures. Feinstein identified five steps in evaluating clinimetric tests.[18] Within each of these steps different terms can be used to describe the validity assessment. The review only includes papers if they mention at least one of the following terms (for definitions of the validity terms see the Glossary, page ii):

- **choice of component variables**: instrument construction, item selection, Delphi method
- **sensibility**: content validity, face validity, feasibility, ease of usage
- **consistency**: reliability, reproducibility, test–retest validity, generalisability, internal consistency, homogeneity, Cronbach's alpha
- **accuracy**: criterion validity, concurrent validity
- **suitability**: construct validity.

The study participants' level of cognitive impairment and to what extent the instrument's validity assessment had taken into account the level of cognitive impairment were also considered.

Excluded studies

The following studies were excluded from the review:

- studies that used general health status instruments measuring only one general health dimension (i.e. only physical functioning, psychological status, or pain)
- studies that only featured participants with cognitive impairment due to dementia or

conditions strongly related to dementia (Parkinson's disease, Alzheimer's disease or Progressive Supranuclear Palsy). These studies are currently included in a separate research project conducted by Dr Subarjee and co-workers, which examines the validity of health status measures in people with dementia

Search strategy

The following databases were searched:

- MEDLINE (1966–1999/8)
- PsycLIT (1966–1999/6)
- EMBASE (1974–1999/8)
- CAB HEALTH (1983–1999/8)
- Dissertation Abstracts Online (1861–1999/8)
- CINAHL (1982–1999/6)
- Sociological Abstracts (1963–1999/6)
- Linguistics and language behaviour abstracts (1973–1999/8)
- Health and Psychosocial Instruments (HAPI) (1985–1998/12)
- Ageline (1965–1999/8)
- Japan Science & Technology – JICST-Eplus (1985–1999/8)
- Pascal (1973–1999/8)
- Conference Papers Index (1973–1999/8)
- NTIS (1964–1999/8)
- Mental Health abstracts (1969–1999/8)
- The MAPI Institute's International Health-related Quality of Life Outcomes Database (print-out received 19-5-1999)
- Nuffield Institute for Health, Outcomes Activities database (searched)
- Econlit (1969–1999/7)
- HealthSTAR (1975–1999/8)
- Social Science Citation Index (BIDS) (1981–1999/9)
- Science Citation Index (BIDS) (1981–1999/9)
- BibEc (Printed working papers in economics) (searched 8-6-1999)
- Economics Working Paper Archive (1993–1999/6)
- Internet Documents in Economics Access Service (IDEAS) (searched 30-6-1999).

A range of general health status bibliographies were also handsearched and these are recorded in appendix 1. Reference lists of all the studies included were manually searched to identify additional trials, and members of the expert panel were asked for any studies relevant for this systematic review.

The search process followed a typical pattern of iterative testing. Each iteration produced studies of interest, and further terms to enrich the strategy (see appendix 2).

Data extraction

Data extraction was completed independently by two reviewers and double-checked for accuracy. Information on the following items was collected:

- instrument description
- sampling and population characteristics (including the number of participants with cognitive impairment)
- study design
- aims of the study
- results of the validity assessment (choice of component variables, sensibility, consistency, accuracy and suitability)
- conclusions about the measure and overall study conclusions.

All other measurements concerned with the validation of measures were also collected and described, including:

- outcome measures relating to instrument construction (e.g. factor analysis)
- consistency (e.g. Cronbach's alpha)
- accuracy (e.g. Pearson's correlation between the measure and some 'gold standard'), and
- suitability (e.g. Pearson's correlation between certain domains of the measure and validated concurrent measurements of that domain).

A copy of the data extraction form can be found in appendix 3.

Assessment of study quality

The quality of included studies was assessed in order to make judgements about the extent to which an instrument has been validated in the studies, and to undertake recommendations about the usefulness and appropriateness of measures in specific populations. The quality assessment was not used to include or exclude studies. Two reviewers independently assessed the quality of each study and any disagreements were resolved by discussion.

Level of cognitive impairment within the study population

As an initial part of the quality assessment, studies were first rated on the number of participants with cognitive impairment. On this basis the studies were divided into the following three groups:

- more than 50% of respondents had cognitive impairment, or separate analyses were included for people with cognitive impairment
- 20–50% of respondents had cognitive impairment
- less than 20% of respondents had cognitive impairment, number not mentioned or unclear (even after contacting the authors).

Level of validation within the study

Assessment was carried out using a self-developed instrument with six dimensions, five of which were derived from Feinstein's guidelines on evaluating clinimetric tests.[18] The five quality dimensions derived from Feinstein's guidelines were as follows.

- **Choice of component variables** This dimension was concerned with questions like:
  - Where do the items come from?
  - Are the questions based on existing measures, clinical observations, expert opinion, or patients' reports?
  - Is the measure based on a theory?
  - Were items checked for high or low endorsement frequency, restrictions in the range of answers, comprehensibility and ambiguity of phrases or possible offensive content?
- **Sensibility** Feinstein formulated criteria for what he calls the sensibility of a measure, including the applicability of the instrument, its clarity and simplicity, likelihood of bias, comprehensiveness, and whether redundant items have been included. These criteria facilitate quantitative rating of an instrument's face and content validity. **Face validity** examines whether an instrument appears to be measuring what it is intended to measure, and **content validity** examines the extent to which the domain of interest is comprehensively sampled by the items, or questions in the instrument.
- **Consistency** For general health status measures, consistency refers to having a high signal-to-noise ratio. The goals of general health status measures include differentiating between people who have better general health status and those who have worse general health status (a discriminative instrument) as well as measuring how much the general health status has changed (evaluative instrument). For discriminative instruments, the way of quantifying the signal-to-noise ratio is called **reliability**. If the variability in scores between patients (the signal) is much greater than the variability within patients (the noise), an instrument will be deemed reliable. Reliable instruments will generally show that stable patients have more or less the same results after repeated administration (**test–retest reliability**). For evaluative instruments, the method of determining the signal-to-noise ratio is called **responsiveness**. Responsiveness refers to an instrument's ability to detect change.
- **Accuracy** Accuracy refers to whether the general health status measure is really measuring what it is intended to measure. Although no overall gold standard for general health status measures exists, instances occur where a specific target for a particular measure can be treated as a criterion or gold standard. Under these circumstances, **criterion validity** is used to describe whether the results correspond to those of the criterion standard. Criterion validity is usually divided into two types: **concurrent validity** (where the criterion measure is assessed at the same time) and **predictive validity** (where the criterion measure will not be available until some time in the future).
- **Suitability** Feinstein refers to suitability as construct validity, that is the extent to which the general health status measure is suitable for the job for which it has been designed.[18] A construct is a theoretically derived notion of the domain(s) to be measured. An understanding of the construct will lead to expectations as to how the instrument should behave if it is valid. **Construct validity** compares the measure to other measures and examines the logical relations that should exist between a measure and characteristics of patients and patient groups.

As the assessment of general health status in cognitively impaired people often involves assessments by third parties or proxies, it was decided to add one further dimension to those described above.

- **Patient–proxy or inter-rater agreement** Patient–proxy agreement refers to the extent to which patient and proxy (family member, carer or health professional) ratings of general health status agree. The extent to which agreement exists between different observers or users of the instrument is referred to as **inter-rater reliability**. Ideally different raters interviewing the same respondents within a very short time period and comparing their responses assess the latter. The most appropriate measure of inter-rater reliability is the Kappa coefficient of agreement. Simple correlation analysis makes no allowance for chance agreement. In contrast, Kappa takes chance agreement into consideration and produces a coefficient between –1 and +1, negative values indicating a consistent disagreement between raters.

The dimensions of patient–proxy agreement and inter-rater agreement have been added to reflect

the importance of these issues in the measurement of general health status in cognitively impaired individuals. It may be assumed that agreement between patients and proxies and between different raters shows the strength of the measure. However it should be noted that agreement between two people does not automatically make their judgements more accurate. Therefore the outcomes of patient–proxy and inter-rater agreement should be interpreted with caution. The scoring for each of these dimensions was as follows and reflects likelihood of bias:

- item not mentioned (0 points)
- item mentioned, but insufficiently described (1 point)
- item mentioned and described in full (2 points).

The total score for each study was derived from the sum of each of the scores from the six quality criteria (i.e. 0 to 2 points per criterion), divided by the maximum score possible (i.e. 10 points for established measures and 12 points for new measures not previously validated). Where more than one study used the same measure a total score for the measure was calculated by assessing the maximum total score for each quality criteria (i.e. 0 to 2 points), combining those scores and dividing the total by the maximum score possible (as above).

Quality scores were used in this report to reflect the level of psychometric validation of the instrument under investigation (i.e. how much of the different facets of an instrument's reliability and validity have been reported) and its applicability to cognitively impaired individuals. Therefore the quality scores are not used in the traditional sense to reflect the methodological validity of a study. The quality scores are purely arbitrary. However, where possible the data used to assign scores have been clearly stated either in the text or in the appendices, so the reader can check the decisions and use their own judgement if they wish.

Analysis

Descriptive data on the construction of the health instruments, information on the choice of component variables, and information concerning the sensibility of the measure were sometimes not quantifiable; in these cases the data were analysed in a qualitative manner.

The studies were initially categorised into the three groups according to the number of cognitively impaired individuals participating in the study. Those studies with a greater number of cognitively impaired participants (i.e. at least 20% of respondents with cognitive impairment) and describing an instrument with a more extensive level of validation (i.e. a quality score of 0.25 or more based on all relevant studies) were discussed in detail. These data were presented in two ways: first in terms of the individual measures themselves and second, according to the disease status of the study population (i.e. stroke, MS, LD and brain injury). Both sections present the same data, but this duplication enables the reader to see what evidence is available for a specific measure and which measures are available for a specific population (see appendix 4 for details of included studies).

The remaining studies which had fewer cognitively impaired participants (i.e. < 20%) or a limited level of psychometric validation (i.e. a quality score of < 0.25) were summarised in tabular form (see appendix 5).

Chapter 3
Results

The results of the review are described below in the form of a narrative. An overview of the included studies is presented along with a general description of the quality of the studies and the level of validation according to the criteria outlined in the chapter 2. The data from eight measures with quality scores greater than 0.25, which were used in populations where at least 20% of the individuals were cognitively impaired, are then presented in further detail. These data are presented in two different ways – the eight measures are discussed individually and then the same eight measures are grouped together and discussed in terms of the four disease groups (stroke, MS, LD and brain injury). In both cases the same data are presented so the reader need only look at the format most suited to their needs.

Throughout this chapter and chapter 4, 'validity' and 'validation' refer to the psychometric evaluation of the measures. The quality of studies and the quality scores awarded reflect the level of validation in terms of how many of the six psychometric components identified in chapter 2 were evaluated.

- A quality score of less than 0.25 indicates that the level of validation was limited.
- A quality score between 0.25 and 0.40 indicates that the level of validation was moderate.
- A quality score of more than 0.40 indicates that the measure was extensively validated.

The terms limited, moderate and extensive are used throughout the following sections. This grading system and the quality score itself are arbitrary, but they do provide an indication of the extent to which the validity of the measure has been examined. This is important when trying to determine which data and general health status measures are most likely to be applicable to individuals with cognitive impairment.

Included studies

The search strategies identified 5987 candidate publications. Two reviewers examined the titles and abstracts of these publications independently. Full reports of 665 references were then examined in further detail. Eighty-three of the 665 references met the inclusion criteria and data were extracted from these reports. The 83 publications reported data from 71 different studies. A total of 34 different general health status measures were described in the 83 publications, with the SIP and the SF-36 being the most frequently used measures (20 and 19 studies, respectively).

None of the studies identified in this review specifically looked at the use of general health status measures for economic evaluations. Although 98 validity assessments of a general health status measure were reported in people with acquired brain injuries and LD, only 52 of the validity assessments were performed in populations definitely or probably including people with cognitive impairment, these are described in further detail in the following section (see also appendices 4 and 6). The remaining 46 validity assessments, performed in populations that probably or definitely did not include people with cognitive impairments, were not included in the main analyses but have been described in appendix 5.

To reflect their applicability to people with cognitive impairment, studies were ranked according to the possibility that they included people with cognitive impairment, this yielded the following five categories of studies.

- **Definitely included** Studies explicitly stated that people with cognitive impairment were included, for instance by reporting a score for cognitively impaired respondents (e.g. Mini Mental State Examination (MMSE) < 24), or studies reported separate scores for cognitively impaired respondents.
- **Probably included** The level of cognitive impairment was not explicitly stated, but either the sampling procedure (<10% did not complete the instrument) or the study aims (e.g. to examine the response rate of a measure) suggested that people with cognitive impairment were likely to be included.
- **Unclear whether included** Studies did not state how many cognitively impaired persons were included, or the non-response rate was between 10% and 20%.
- **Unlikely to be included** Studies did not state how many cognitively impaired persons were

included, but the high non-response (> 20%) suggested it was unlikely that people with cognitive impairment were included.
- **Definitely not included** People with cognitive impairment were explicitly excluded from the studies.

Table 1 shows the likelihood that measures have been evaluated in people with cognitive impairment.

The SIP was included in the greatest number of studies (n = 10), followed by the SF-36 (n = 8). The NHP and EQ-5D measures were each included in four studies, while the MSQOL-54, QoL-Index, Life Satisfaction (Viitanen) scale, VAS-QoL, ComQol-I5, QoL-Interview and LSS were included in two studies each. The remaining measures (COOP/WONCA charts, Life Satisfaction (Cantril's ladder), SWB, MS-QLI, PCRS, Oregon QLQ, LSI, CQOL, QUOLIS, BICRO-39, QoL for MS and Ho's QOL-scale) were each included in one study.

A large number of the measures were validated in stroke patients (19 studies), 11 measures were validated in MS patients and ten measures were validated in patients with LD and acquired brain injury. One of the studies that assessed both the SWB scale and the Life Satisfaction (Cantril's Ladder) scale, included patients with stroke, MS and cognitive impairment.[19]

Most studies assessed the validity of measures in populations based in the USA (n = 19) and the UK (n = 14). The other studies were carried out in The Netherlands (n = 5), Canada (n = 4), Australia (n = 2), Finland (n = 2), Hong Kong (n = 2), Italy (n = 2) and Sweden (n = 2).

Six instrument validations were published in 1999, eight in 1998 and 15 in 1997, therefore more than half were published during the past 3 years. The number of publications in earlier years were as follows: 1996 (n = 2), 1995 (n = 7), 1994 (n = 5), 1992 (n = 1), 1990 (n = 1), 1989 (n = 1), 1988 (n = 3), 1987 (n = 1), 1986 (n = 1), and 1984 (n = 1).

None of the studies specifically examined the use of general health status measures in economic evaluations. The two most important instruments identified in the review by Brazier and co-workers were the EQ-5D and the HUI.[7] However, none of the studies in this review evaluated the HUI. Two studies using the EQ-5D were identified, but both included populations with only limited numbers of cognitively impaired persons (i.e. < 20%), and so little can be concluded from these studies (see appendix 6 for further details).

The number of patients in which measures were validated varied between studies from 15[20] to 5279.[19] Twenty-seven validity assessments included fewer than 100 participants, while 13 included between 100 and 299 participants. Twelve validity assessments were performed in 300 or more participants.

Quality of included studies

The quality scores for each individual study are described in appendix 7. Most studies had very low quality scores. In most cases this was due to the fact that the studies were not specifically designed to assess the validity of the general health status measure(s) used, but rather some items of validity were assessed in the process of the study. In total, 52 validity assessments were identified. In seven cases the same validity assessment was described in multiple publications. In such cases the quality score was based on all of the relevant publications.

The studies reported in appendix 7 are first grouped according to the likelihood that they include respondents with cognitive impairment and then the measure that they evaluate. Quality scores are assigned to each individual study, and where a measure has been validated in more than one study a cumulative quality score across all studies for that measure is presented.

Scores on the quality assessment ranged from 0 to 1. The total number of points was divided by ten for existing measures and 12 for new measures (where points were added for the description of the choice of component variables). A large number of measures (nine out of 23) scored 0.25 or less, and eight measures scored 0.5 or more. These included: the ComQol-I5[21,22] where more than 50% of the respondents were cognitively impaired; the MS-QLI[23] and the SIP[14,23,24] where 20–50% of respondents were cognitively impaired; and the BICRO-39,[25] SF-36,[15,26–29] SIP,[30–36] EQ-5D,[13,26,27,37] MSQOL-54[38,39] and LSS[40] validated in populations where less than 20% were cognitively impaired. Four of these measures were new (ComQol-I5, MS-QLI, BICRO-39 and MSQOL-54).

Level of validation of general health status measures

Studies including 50% or more of respondents with cognitive impairment

Nine (out of 23) measures were validated in studies where 50% or more of the respondents

TABLE 1 *Number of instruments found in 83 publications presented in order of the number of studies and the certainty that population includes people with cognitive impairment (number of participants included)*

	Total no. of instruments	Population includes people with cognitive impairment				
		Definitely	Probably	Unclear	Unlikely	Not
SIP	20	7 (1365)	3 (177)	9		1
SF-36	19	4 (628)	4 (2514)	7	3	1
NHP	7	1 (96)	3 (293)	1	2	
EQ-5D	5		4 (2399)	1		
MSQOL-54	5	2 (204)		3		
QoL-Index MS	3	2 (61)				1
Life Sat. (Viitanen)	3	2 (19)				1
VAS-QoL	3	1 (92)	1 (96)	1		
DIP	3			3		
ComQol-I5	2	2 (430)				
QoL-Interview	2	2 (232)				
LSS	2	1 (120)	1 (153)			
COOP/WONCA	2	1 (84)		1		
FAMS	2			1		1
Life Sat. (Cantril)	1	1 (5279)				
SWB	1	1 (5279)				
MS-QLI	1	1 (300)				
PCRS	1	1 (53)				
Oregon QLQ	1	1 (50)				
LSI	1	1 (48)				
CQOL	1	1 (26)				
QUOLIS	1	1 (10)				
BICRO-39	1		1 (235)			
QoL for MS	1		1 (171)			
Ho's QoL	1		1 (109)			
Farmer QoL Index	1			1		
HUI	1			1		
Q-TWiST	1			1		
QOL-Q	1			1		
QWB	1			1		
SSQ	1			1		
Danish LLQ	1				1	
QUAL-OT	1				1	
Utility Assessment	1				1	
TOTAL	98	33	19	33	8	5

NHP, Nottingham Health Profile; MSQOL-54, Multiple Sclerosis Quality of Life 54-item Scale;VAS, visual analogue scale; DIP, Disability and Impact Profile; ComQol-I5, Comprehensive Quality of Life Scale – Intellectual Disability; LSS, Life Situation Survey; COOP/WONCA, Dartmouth COOP Functional Health Assessment charts/WONCA; FAMS, Functional Assessment of Multiple Sclerosis; SWB, Subjective Well-Being scale; MS-QLI, Multiple Sclerosis – Quality of Life Interview; PCRS, Patient Competency Rating Scale; Oregon QLQ, Oregon Quality of Life Questionnaire; LSI, Life Satisfaction Index; CQOL, Child Quality of Life Questionnaire; QUOLIS, Quality of Life Scale; BICRO-39, Brain Injury Community Rehabilitation Outcome Scales (39-items); Q-TWiST, Extended Quality-Adjusted Time Without Symptoms and Toxicity; QOL-Q, Quality of Life – Questionnaire; QWB, Quality of Well-Being index; SSQ, Soweto Stroke Questionnaire; LLQ, Laman & Lankhorst Questionnaire; QUAL-OT, Quality of Life – Occupational Therapist

were cognitively impaired. The SF-36 was validated in 71 LD patients, 85% of whom were illiterate and 48% of whom had no or only partial verbal communication skills.[41] The ComQol-I5 was validated in 430 LD patients; their Slosson age equivalent was 5.4 years (range, 2.7–9.3).[21,22] The LSS was validated in 120 LD patients, 50% of whom were severely or profoundly affected.[42] The SWB and Life Satisfaction (Cantril's ladder) were validated in 5279 patients with MS, stroke and cognitive impairment; patients with an MMSE score of less than 17 were excluded, but data on a subgroup of patients with an MMSE score of less than 24 were described.[19] The VAS-QoL was validated in 92 brain injury patients, whose cognitive status was assessed by a neurologist and self-report.[43] The LSI was validated in 48 LD patients, whose mean overall age equivalent was 7.92 years.[44] The CQOL was validated in 26 LD patients, whose intelligence quotients (IQ) were lower than 70 (most <50).[45] The QUOLIS was validated in ten LD patients, four of whom had mild, two moderate and four severe disability.[46]

Studies including 20–50% of respondents with cognitive impairment

Five other measures were validated in studies where 20–50% of the respondents were cognitively impaired. The MS-QLI was validated in 300 MS patients, one-third of whom had severe neurological impairment.[23] In the same study the SIP and SF-36 were used as comparisons for the MS-QLI. The NHP was validated in two studies. One included 96 stroke patients, 28% of whom had an MMSE score of 24 or less. The SF-36 was also validated in this study.[47] The other study included 82 stroke patients, 10% of whom had severe problems with orientation, 15% had considerable problems and 23% had moderate problems, as assessed using the London Handicap Scale.[48] The SIP was validated in three studies including 20–50% cognitively impaired respondents. The first study has been mentioned above.[23] The second study included 437 stroke patients, 108 of whom had cognitive impairment, and proxies (family members) completed their SIP. These data are presented with proxy assessments of 228 stroke patients without cognitive impairment.[14] The third study included 20 brain injury patients, who were described as having cognitive and memory difficulties, but there were 25 (out of 45) drop-outs.[24] The COOP/WONCA charts were validated in 84 stroke patients, 80% of whom had normal speech.[49] The Oregon QLQ was validated in 50 LD patients, 29% of whom were severely or profoundly disabled.[50]

Studies including less than 20% of respondents with cognitive impairment

The SIP,[30–36] SF-36,[15,26–29] NHP,[34,51] EQ-5D,[13,26,27,37] MSQOL-54,[38,39] QoL-Index MS,[52,53] Life Satisfaction (Viitanen),[20,54] VAS-QoL,[55] QoL-Interview,[56,57] LSS,[40] PCRS,[58] BICRO-39,[25] QoL for MS[29] and Ho's QoL[59] were validated in 22 studies of populations in which less than 20% of the respondents were cognitively impaired. Most studies did not report any information on the level of cognitive impairment of the respondents. In other studies the level of cognitive impairment was clearly stated. For instance, the MSQOL-54 was validated in 204 MS patients – the mean MMSE score was 26.9 (range, 14.6–30.3), and 9% (n = 18) had an MMSE-score of 23.8 or less;[38,39] and the QoL-Index MS was validated in 61 MS patients – 41% had no cognitive disturbances and 43% had cognitive disturbances that did not intervene with their daily activities, as assessed using the Incapacity Status Scale.[52,53]

The BICRO-39 was validated in 235 brain injury patients but the level of cognitive impairment was not reported.[25] As the level of validation was quite extensive in this study (quality score, 0.67) the authors were contacted for further information. The authors replied that many, if not most, of the participants had some degree of cognitive impairment, but it would be impossible to say how many. Therefore this study was rated as having less than 20% of respondents with cognitive impairment.

The MS-QLI (validated in a population where 20–50% of the respondents were cognitively impaired) was thoroughly assessed for the **choice of component items**, as was the BICRO-39 to some extent. However, none of the other new measures acquired any points on this dimension.

Consistency was assessed in all of the measures (except for the VAS-QoL) validated in studies that included more than 50% of respondents with cognitive impairment, and in 70% of the measures overall. Consistency was nearly always assessed using internal consistency and sometimes by using test–retest validity, or both.

None of the studies in which more than 50% of respondents were cognitively impaired assessed the **sensibility** and **accuracy** of the measures. Overall, sensibility and accuracy were assessed in 26% and 13% of studies, respectively. Accuracy was not reported mainly because there was no gold standard available for general health status measurement, therefore comparisons with other

measures were judged as convergent or discriminant validity (suitability).

Five of the nine measures evaluated in studies where more than 50% of the respondents were cognitively impaired, were to some extend validated for **suitability**. Overall, 72% of measures were assessed on suitability. Twenty-eight per cent of these measures, and three out of the nine measures evaluated in studies in which more than 50% of respondents were cognitively impaired, were validated for patient–proxy agreement or inter-rater agreement.

As mentioned previously, the goals of general health status measures include differentiating between people who have better general health status and those who have a worse general health status (a discriminative instrument), as well as measuring how much the general health status has changed (evaluative instrument). No assessments on the responsiveness of measures were found and therefore there were no indications as to the instrument's ability to detect change.

Detailed results of validity assessments for general health status measures

The instruments described below (ComQol-I5,[21,22] SF-36,[23,41,47] LSI,[44] LSS,[42] CQOL,[45] QUOLIS,[46] SIP,[14,23,24] and MS-QLI[23]) had a quality score of 0.25 or higher (i.e. they were validated using at least two of the specified criteria), and at least 20% of their respondents were cognitively impaired. This scoring system is purely arbitrary and reflects the level of validation, so these eight measures may not necessarily be the best or most appropriate measures for cognitively impaired individuals. The actual suitability of the measures is dependent on the outcome of the validity assessments, which are reported below. However, from the data retrieved, these eight measures are most likely to reflect the current status of general health status measurement in cognitively impaired people (see appendix 5 for details of the remaining studies).

As previously stated, identical data are presented in both of the following sections (*Individual general health status measures* and *Health status measures within specific patient groups*) and so readers need only refer to the section which is most appropriate to their requirements.

Further details of how to obtain more information about these eight instruments are given in appendix 8.

Individual general health status measures

ComQoL-I5

The ComQol-I5 is a multidimensional instrument which includes seven domains:

- material well-being
- health
- productivity
- intimacy
- safety
- place in community, and
- emotional well-being.

Each domain has an objective and a subjective component. The measurement of each objective domain is achieved by obtaining an aggregate score based on the measurement of three objective indices relevant to that domain, for example 'material well-being' is measured by an aggregate score of income, type of accommodation and personal possessions. The measurement of each subjective domain is achieved by obtaining a satisfaction score of that domain which is weighted by the perceived importance of the domain for the individual.

The measure was validated in three studies that included a total of 430 patients with LD.[21,22] None of the studies assessed the instrument in terms of choice of component variables and sensibility.

Consistency

Consistency was assessed by internal consistency and test–retest reliability. Internal consistency of the objective scales, as measured by Cronbach's alpha, including 430 patients with LD was as follows:

- material well being = 0.14
- health = 0.64
- productivity = 0.35
- intimacy = 0.45
- safety = 0.56
- community = 0.15
- emotional well-being = 0.42
- total objective scale (21 items) = 0.56.

Subjective data showed that for the seven components combined, Cronbach's alpha was 0.48 for importance, 0.65 for satisfaction, and 0.68 for 'importance × satisfaction'.

Test–retest reliabilities were generally high at 1–2 weeks, with the exception of the 'importance of intimacy' and 'satisfaction with safety' components. Beyond the 2-week interval, the retest correlations became 'somewhat erratic' according to the authors.[21,22]

Inter-domain correlations between the objective domains were all low, with 'material well-being × intimacy' having the highest correlation of 0.25. In general there was found to be no significant correlation between objective and subjective domain scores, with the exception of objective health, which correlated positively with both health importance ($r = 0.25$; $p < 0.05$) and health satisfaction ($r = 0.33$; $p < 0.001$).

Patient–proxy agreement

Relationships were found between objective scores and the third-party subjective estimates provided by carers. Significant correlations were found in three domains for importance:

- health = 0.25 ($p < 0.05$)
- safety = 0.23 ($p < 0.05$)
- community = 0.33 ($p < 0.01$)

and for satisfaction:

- material well-being = 0.29 ($p < 0.05$)
- health = 0.28 ($p < 0.05$)
- intimacy = 0.33 ($p < 0.01$).

These correlations did not conform to those of the clients themselves and may indicate that carer third-party estimations were influenced by objective cues.

In terms of intra-domain data the degree of correspondence was generally slight. For ratings of importance, only the domain of 'intimacy' (0.23; $p < 0.05$) reached significance. For ratings of satisfaction, significant agreement was recorded for the domains of 'health' (0.22; $p < 0.05$), 'productivity' (0.33; $p < 0.01$) and 'safety' (0.31; $p < 0.01$). In terms of the relative inter-domain hierarchy of importance and satisfaction, only the former achieved a significant degree of congruence between the client and carer estimations ($r = 0.96$, $n = 7$, $p < 0.01$ and $r = 0.55$, $n = 7$, p = not significant, respectively). The lack of correspondence for estimations of satisfaction was most marked for the domain of 'community'.

Summary

The same study group validated the ComQoL-I5 in three studies, which included 430 patients with LD. All studies included more than 50% of respondents with cognitive impairment, and the quality score was 0.50. Internal consistency was under 0.60 for the total objective scale and for all subscales except 'health'. Test–retest reliability was generally high within 2 weeks, and reported by the authors as 'somewhat erratic' beyond 2 weeks.

Patient–proxy assessments showed generally little correspondence, carer estimations correlated more strongly with objective cues than with patient estimates.

SF-36

The SF-36 comprises eight health scales:

- physical functioning (ten items)
- role limitations-physical (four items)
- bodily pain (two items)
- general health (five items)
- vitality (four items)
- social functioning (two items)
- role limitations-emotional (three items)
- mental health (five items).

Two core dimensions of health: 'physical' and 'mental' can be derived from these eight scales.

Seven studies assessed the validity of this measure and included a total of 3142 participants. Four of the studies included 2600 stroke patients,[15,26–28,47] two studies included 471 patients with MS,[23,29] and one study included 71 patients with LD.[41] The results from one study,[41] in which more than 50% of the respondents were cognitively impaired, and two studies,[23,47] in which 20–50% of the respondents were cognitively impaired, will be described.

Consistency

Reliability of the SF-36 was assessed by Cronbach's alpha in two studies.[23,41] In both studies alpha values were above 0.70 for all subscales except 'social functioning', which equalled 0.33 in the study with 71 patients with LD. Factor analysis showed that the SF-36 could be divided into two factors, one representing physical health and one representing mental health. However 'social functioning' and 'vitality' had high factor loadings for both factors.[41]

Suitability

Construct validity was assessed by correlating the SF-36 subscales with the Barthel Index.[47] As expected the 'physical functioning' (0.81) and 'vitality' (0.50) subscales showed the highest correlations, while the 'role limitations-emotional' (0.22) and 'mental health' (0.33) subscales showed the lowest correlations.

In a study among patients with MS, 'physical functioning' correlated significantly with the SIP-'physical' subscale (–0.62), and 'mental health' correlated significantly with the SIP-'psychosocial' subscale (–0.51), whereas correlations between

dissimilar constructs were weak (<0.30).[23] In the same study, correlations between SF-36 subscales and impairment measures were weak, with the exception of 'physical function' which correlated strongly with the Expanded Disability Status Scale (EDSS) and moderately with quantitative measures (walk test and 9-hole peg test).

Inter-rater agreement

This was assessed in the study with 71 LD patients.[41] Moderate reliability was achieved (Pearson's *r* ranged from 0.54 to 0.89), however, the 'role-physical' (0.44) and the 'role-emotional' (0.31) subscales were considerably less reliable. The authors stated that the level of agreement is such that it would not be useful for comparison across individuals but is within the range required to make group comparisons.

Summary

The SF-36 was validated in seven studies that included 3142 patients. In most of the studies less than 20% of respondents were cognitively impaired. Two of the studies included 20–50% of respondents with cognitive impairment, but these studies had low-to-moderate quality scores (0.1[47] and 0.3[23]). Only one study included over 50% of respondents with cognitive impairment and this had a quality score of 0.3.[41] Results from two of the studies showed good-to-excellent internal consistency for the total instrument and for all of the subscales with the exception of the 'social function' subscale.[23,47] Correlations between SF-36 subscales and impairment measures were weak, with one exception (SF-'physical' correlated strongly with the EDSS and moderately with quantitative measures).

Results from one of the studies showed that the internal consistency and item total biserial correlations for each subscale were largely acceptable with the exception of the 'social functioning' scale where both the mean item-total correlation and Cronbach's alpha (0.33) were low.[41] Through factor analysis on the subscales, a two-factor solution was obtained and the factors could clearly be interpreted as a physical health factor and a mental health factor.

Inter-rater reliability between patients with LD and their carers was moderate; however, the 'role-physical' scale and the 'role-emotional' scale were considerably less reliable.

LSI

The instrument consists of 36 items, included in five dimensions:

- living arrangement
- relationships and social activities
- job status
- health
- general happiness.

The validity of this instrument was assessed in one publication that included 48 patients with LD.[44]

Consistency

Internal consistency as measured by Cronbach's alpha was acceptable to good for most scales:

- living arrangement = 0.79
- job status = 0.64
- health = 0.75
- general happiness = 0.84
- total instrument = 0.87,

but poor for 'relationships and social activities' (0.48).

Suitability

Convergent and discriminant validity was assessed using a multitrait–multimethod matrix. There was a moderate-to-high level of convergence between corresponding subscales from the LSI and the Life Satisfaction Scale-Modified (a larger version of the LSI). Low-to-moderate relations between all other subscales of the LSI and modified version supported discriminant validity.

The LSI showed significant negative correlations as predicted with 'leisure-interests' ($r = -0.38$) and 'leisure-constraints' ($r = -0.45$) and no significant relation with broad cognitive functioning ($r = -0.30$). Contrary to predictions, the LSI showed no significant negative correlations with 'leisure-activity participation' ($r = 0.02$) and 'leisure-preference' ($r = -0.06$).

Summary

The LSI was validated in one study, which included 48 patients with LD.[44] More than 50% of respondents had cognitive impairment, and the quality score for the study was 0.3. Internal consistency was acceptable to good for most scales, but poor for 'relationships and social activities'. Construct validity, as measured using the multitrait–multimethod matrix, was good. Convergent validity results with other measures were somewhat unclear.

LSS

The LSS is a 20-item Likert-type rating instrument developed by Chubon for use in assessing perceived life quality by persons with disabilities and chronic illness, as well as by the general population. The items are clear and concise, with

readability estimated to be in the fourth grade to fifth grade range. Although designed to be self-administered, provisions are included for oral administration to persons who have reading or other limitations that prohibit them from completing the written self-report form. The measure was developed around ten commonly accepted QoL domains, including:

- work
- leisure
- health
- love/affection
- self-esteem

and ten additional areas, including:

- stress
- mobility
- autonomy
- energy level
- social support
- mood/affect
- public support.

These were derived through a critical incidents-based study of adults with various chronic illnesses and/or disability including MS.[40] Items were rated on a six-point scale, disagree very strongly (1) to agree very strongly (6). A one- to seven-point scale has also been used.[42]

Validity was described in two studies, which included a total of 153 MS patients and 120 patients with LD and other conditions.[40,42] Only the second study is described here, because this study included more than 50% cognitively impaired respondents;[42] the other study included less than 20%.[40]

Consistency

In patients with LD, Cronbach's alpha was 0.74 for the high-IQ group and 0.61 for the low-IQ group.[42] The Cronbach's alpha scores for the high- and low-IQ groups were lower than those reported for non-cognitive disabilities, but were sufficient to render the results usable in interpreting the group data.

Suitability

Low correlations were found with the Resident Choice Assessment Scale (RCAS) ($r = 0.35$; $p < 0.01$), the Mental Deficiency Adaptive Behaviour Scale-Part I (ABS) ($r = 0.35$; $p < 0.01$), and the Mental Deficiency Adaptive Behaviour Scale-Part II ($r = 0.04$; p = not significant). However the LSS correlated significantly with the Program Analysis of Service System Implementation of Normalisation Goals (PASSING) scores ($r = 0.91$; $p = 0.01$).

Summary

The LSS was validated in two studies, which included 153 MS patients and 120 patients with LD and other conditions. Less than 20% of the MS patients had cognitive impairment, but more than 50% of the patients with LD had cognitive impairment. The latter study had a quality score of 0.3.[42]

Internal consistency was 0.74 for the high-IQ group and 0.61 for the low-IQ group, showing that internal consistency was lower but sufficient for people with cognitive impairment, according to the authors. Convergent validity showed that the LSS is sensitive to living conditions (PASSING), and less sensitive to mental deficiencies (ABS) and independence (RCAS).

CQOL

The CQOL consists of 15 domains:

- getting about and using hands
- doing things for self
- soiling or wetting
- school
- out of school activities
- friends
- family relationships
- physical discomfort
- worries
- depression
- seeing
- communication
- eating
- sleep
- appearance.

One study, which included 26 mothers of patients with LD, assessed the validity of this instrument.[45] This study did not assess the instrument in terms of choice of component variables and sensibility.

Consistency

Cronbach's alpha for combined scales measuring function, satisfaction and upset for all mothers ($n = 75$), including mothers of children with chronic physical disorders (CPD) and psychiatric disorders (PD), were 0.81, 0.86 and 0.86, respectively.

Test–retest reliability showed that the correlation of the combined function scores was 0.83. Individual correlations for the mothers' scores ranged from 0.11 to 1.00 (based on 19 mothers of patients with LD and 15 mothers of children with CPD), with the majority between 0.4 and 0.7. The mothers of children with LD obtained the highest level of reliability.

Suitability

The correlation between the mothers' total function score and the Child's Global Adjustment Score was 0.64 (p = 0.01). Mothers of children with LD perceived their children as functioning less well than either of the other two groups (CPD and PD) on 12 of the 15 items and significantly worse than both the other two groups on six of these items. Only on 'family relationships', 'depression' and 'physical discomfort' did mothers of LD children not provide the highest mean scores.

Summary

The CQOL was validated in one study that included 26 mothers of patients with LD.[45] More than 50% of the patients were cognitively impaired, and the quality score of the study was 0.25. Results showed that internal consistency for combined scales was good, and test–retest reliability was acceptable. Convergent and discriminant validity results were as predicted.

QUOLIS

The QUOLIS consists of 78 statements or indicators, grouped according to 12 domains:

- health services
- family or guardianship
- income maintenance
- education or employment
- housing and safety
- transportation
- social or recreational
- religious or cultural
- case management
- advocacy
- counselling
- aesthetics.

Each of the indicators is rated on four counts, according to the four dimensions of support, access, participation and contentment. Validity of this instrument was assessed in one study that included ten patients with LD.[46] This study did not assess the instrument in terms of choice of component variables and sensibility.

Consistency

Cronbach's alpha for the contentment dimension ranged from 0.79 to 0.99.

Inter-rater agreement

Inter-rater agreement for the contentment dimension ranged from 0.48 to 0.95. Intra-rater agreement showed that 73% (35/48) of the scores were acceptable. For the 13 scores lacking agreement, refinements are suggested. Inter-rater agreement showed that 65% (31/48) of the scores were acceptable.

Summary

The QUOLIS was validated in one study that included only ten patients with LD.[46] More than 50% of the patients were cognitively impaired, and the quality score was 0.25. Results showed that the internal consistency of the instrument was good, and the inter-rater agreement was acceptable for 73% of the scores.

SIP

The SIP consists of 136 items describing the impact of ill health on behaviour in terms of 12 dimensions:

- sleep and rest
- emotional behaviour
- body care and movement
- household management
- mobility
- social interaction
- ambulation
- alertness behaviour
- communication
- work
- recreation and pastimes
- eating.

Weighted sum scores are obtained for the overall profile, physical and psychosocial subtotals, and separately for each of the 12 dimensions.

Validity was assessed in ten studies, which included a total of 1542 participants. Five of the studies included 576 patients with brain injury,[24,30,32,33,35] three studies included 585 stroke patients[14,31,34] and two studies included 381 patients with MS.[23,36]

The results from three studies in which 20–50% of the respondents were cognitively impaired will be described.[14,23,24]

Consistency

Cronbach's alpha for the total SIP-score ranged from 0.89 to 0.95.[23] For SIP subscales, alphas were lower, ranging from 0.70 to 0.85, and the internal consistency of the 'sleep and rest' and 'eating' subscales was relatively poor, based on patients and proxies in one study[14] and based on patients alone in another study.[23]

Suitability

Construct validity was assessed in a study among 300 MS patients.[23] SIP subscales correlated at least moderately with one or more objective measures,

and the pattern of correlations between SIP-subscales and quantitative measures of function (walk-test and 9-hole peg test) was not entirely according to predictions.

Patient–proxy agreement

In a study of 437 stroke patients, it was found that the individual patient level, chance-corrected agreement between patient and proxy scores ranged from moderate for the 'eating' subscale (intraclass correlation coefficient (ICC) = 0.47) to excellent for 'ambulation' (ICC = 0.80) and 'body care and movement' (ICC = 0.82).[14] Good to excellent was noted for physical (ICC = 0.85), psychosocial (ICC = 0.61) and total SIP (ICC = 0.77). ICCs for the other subscales ranged between 0.52 and 0.69.

The mean proxy-rated SIP scores of the physical and psychosocial dimensions and total SIP were found to be significantly associated with the patients' Rankin grade.[14] However, proxy scores were systematically higher than patient scores.

The correlation between patients and 'close others' was assessed in a study among 20 brain injury patients.[24] Correlations were high for all categories (except for 'emotion' (0.395), 'sleep and rest' (0.529) and 'household management' (0.583)):

- ambulation = 0.859
- body care and movement = 0.720
- mobility = 0.825
- recreation = 0.660
- social interaction = 0.646
- alertness = 0.745
- eating = 0.654
- communication = 0.970
- work = 0.965
- overall percentage of dysfunction = 0.905.

It was suggested that 'emotion' and 'sleep and rest' were the most difficult categories for relatives to assess. In the same study, a comparison of mean scores showed that the difference between patients' and relatives' reports was not significant on any category except 'alertness'. In this category patients reported a mean of 33% dysfunction as opposed to 24% reported by relatives. This difference was significant at the 5% level.

Summary

The SIP was validated in ten studies that included 1542 respondents. Very limited validation in people with cognitive impairment was found. Only three studies had 20–50% of respondents with cognitive impairment.[14,23,24] Quality scores in these studies ranged between 0.3 and 0.4, showing that the level of validation was moderate. The remaining studies all included populations where less than 20% of respondents were cognitively impaired.

Results from these three studies showed that the internal consistency for the total SIP score was generally good to excellent. For SIP subscales Cronbach's alphas were lower, the internal consistency of the 'sleep and rest' and 'eating' subscales was relatively poor. Correlations between SIP subscales and quantitative measures of function were not entirely predictable.

Agreement between patient and proxy scores ranged from moderate to excellent. However, proxy scores were systematically higher than patient scores. The 'emotion' and 'sleep and rest' subscales were the most difficult categories for relatives to assess.

MS-QLI

The MS-QLI is a modular MS-specific health-related quality of life (HRQL) instrument consisting of a widely used generic measure, the Health Status Questionnaire (SF-36), supplemented by nine symptom-specific measures, covering:

- fatigue
- pain
- bladder function
- bowel function
- emotional status
- perceived cognitive function
- visual function
- sexual satisfaction
- social relationships.

Validity was assessed in one study that included 300 MS patients.[23]

Choice of component variables

Three expert panels (neurologists specialising in MS, allied health professionals skilled in assessing and treating MS, and MS patients and their families) reviewed a set of candidate measures for the MS-QLI. This resulted in the addition of two items: 'bowel and bladder function', and 'caregiving', in addition to a new domain of 'sexual function'.

Sensibility

The total instrument consisting of 137 items, 80 with abbreviated scales, takes 45 minutes to complete.

Consistency

Cronbach's alpha for all the symptom-specific scales (including 'fatigue impact', 'perceived deficits', 'mental health' and 'social support') was good to excellent (0.77–0.97); for generic HRQL, summary scales were high (0.89–0.95), and SF-36 subscales good to excellent (0.75–0.94) with one exception (SF-36-'social function').

Suitability

Construct validity of generic HRQL (SF-36) was supported at both summary and subscale level. For example SF-'physical' and SIP-'physical' were strongly correlated ($r = -0.62$), as were SF-'mental' and SIP-'psychosocial' ($r = -0.51$); whereas correlations between dissimilar constructs were weak (<0.30). Construct validity of most symptom-specific measures was also supported by modest (<0.40) correlations of the different constructs of the MS-QLI with 'pure' measures, including measurements on 'bladder', 'bowel', 'sexual' and 'visual functioning', Medical Outcome Studies (MOS)-'pain', and MOS-'social support'. Correlations between SF-36 subscales and impairment measures were weak, with one exception – SF-'physical' correlated strongly with the EDSS and moderately with quantitative measures (walk-test and 9-hole peg test).

Summary

The MS-QLI was validated in one study that included 300 MS patients.[23] Twenty to 50% of the respondents were cognitively impaired, and the quality score for the study was 0.5, showing that the level of validation was extensive, however the results may not be fully applicable to people with cognitive impairment. The results showed that the internal consistency of the total MS-QLI and the subscales was good, with the exception of 'social function', and the instrument had good construct validity.

Health status measures within specific patient groups

Stroke

Nine measures were validated in stroke patients:

- SIP
- NHP
- SF-36
- SWB
- Life Satisfaction (Cantril's ladder)
- COOP/WONCA charts
- EQ-5D
- VAS-QoL
- Ho's QoL scale.

The SIP was validated in three studies that included 585 stroke patients. Two studies included less than 20% of respondents with cognitive impairment,[31,34] and one study included 20–50% of respondents with cognitive impairment.[14] The latter study had a moderate level of validation, and showed that the internal consistency (Cronbach's alpha) surpassed or approached 0.70 for nine of 11 SIP subscales. The internal consistency of the 'sleep and rest' and 'eating' subscales was relatively poor, based on patients and proxies. The mean proxy-rated SIP scores on the physical and psychosocial dimensions and total SIP were significantly associated with the patients' Rankin grade. Proxy scores were systematically higher than patient scores. At the individual patient level, chance-corrected agreement between patient and proxy scores ranged from moderate to excellent.

The NHP was validated in four studies, which included a total of 389 stroke patients. Two studies included less than 20% of respondents with cognitive impairment,[34,51] and two studies included 20–50%.[47,48] The latter two studies had limited levels of validation, which showed strong correlations between the London Handicap Scale and the NHP extended activities of daily living score, and the NHP Part 1 overall score, 'energy', 'social isolation' and 'pain' subscales. Weak correlations were shown between the London Handicap Scale and the NHP Part 2 overall score and the Part 1 subscales 'emotion' and 'sleep'. All the dimensions of NHP apart from that for 'sleep' had a significant correlation with the Barthel Index score.

The SF-36 was validated in four studies, which included a total of 2600 stroke patients. Three studies included less than 20% of respondents with cognitive impairment,[15,26–28] and one study included 20–50% of respondents who were cognitively impaired.[47] The latter study had a limited level of validation, which indicated that all the various dimensions for the SF-36 showed a significant correlation with the individual subject's scores and their Barthel Index score.

The SWB and Life Satisfaction (Cantril's ladder) were both validated in one study that included 5279 patients with stroke ($n = 129$) and other conditions.[19] The study included more than 50% of respondents with cognitive impairment, but the level of validation was limited, showing that both measures had internal consistencies ranging from satisfactory (Cronbach's alpha > 0.75) to very good (0.91).

The COOP/WONCA charts were validated in one study that included 84 stroke patients.[49] The study

included 20–50% of respondents who were cognitively impaired, but the level of validation was limited, showing that the COOP/WONCA charts were easy to use and well accepted by patients and doctors.

The EQ-5D was validated in two studies, which included 2399 stroke patients.[13,26,27,37] Ho's QoL scale was validated in one study that included 109 stroke patients[59] and the VAS-QoL was validated in one study that included 96 stroke patients.[55] All these studies included less than 20% of respondents with cognitive impairment.

Summary

In stroke patients, only two measures (the SWB[19] and Life Satisfaction (Cantril's ladder)[19]) were found which had been validated in a population that included more than 50% of stroke patients with cognitive impairment. The level of validation was limited, although both measures had good internal consistency. The SIP, NHP, SF-36 and COOP/WONCA charts were all validated in populations with 20–50% cognitively impaired respondents. For most instruments the level of validation was limited, although the SIP was more extensively validated.

MS

Nine measures were validated in MS patients:

- MS-QLI
- SF-36
- SIP
- SWB
- Life Satisfaction (Cantril's ladder)
- MSQOL-54
- QoL Questionnaire for MS
- LSS
- QoL Index for MS.

The MS-QLI was validated in one study that included 300 MS patients. The study included 20–50% of respondents with cognitive impairment,[23] and the level of validation was extensive, showing that the internal consistency of the total MS-QLI and the subscales was good, with one exception ('social function'), and that the instrument has good construct validity.

The SF-36 was validated in two studies including 471 MS patients. One study included less than 20% of respondents with cognitive impairment,[29] and one study included 20–50% of respondents with cognitive impairment.[23] The latter study was limited in its level of validation; however, the Cronbach's alpha values were above 0.70 for all subscales except for 'social functioning'. Correlations between SF-36 subscales and impairment measures were weak, with one exception (SF-'physical' correlated strongly with the EDSS and moderately with quantitative measures).

The SIP was validated in two studies that included 381 MS patients. One study included less than 20% of respondents with cognitive impairments,[36] and one study included 20–50% of respondents who were cognitively impaired.[23] The latter study was limited in its level of validation, but showed that internal consistency for the total SIP score was generally good to excellent. For SIP subscales, Cronbach's alphas were lower; the internal consistency of the 'sleep and rest' and 'eating' subscales was relatively poor. Correlations between SIP subscales and quantitative measures of function were not entirely predictable.

The SWB and Life Satisfaction (Cantril's ladder) were both validated in one study that included 5279 patients with MS ($n = 10$) and other conditions.[19] The study included more than 50% of respondents with cognitive impairment, but there was only a limited level of validation of the measures. However, the internal consistencies ranged from satisfactory (Cronbach's alpha > 0.75) to very good (0.91).

The MS-QOL-54,[38,39] the QoL Questionnaire for MS,[29] LSS[40] and QoL Index for MS[52,53] were each validated in one study including 61 to 204 MS patients. The study included less than 20% of respondents with cognitive impairment.

Summary

In MS patients only two measures (the SWB[19] and Life Satisfaction (Cantril's ladder)[19]) were found that have been validated in a population including more than 50% cognitively impaired MS patients. However, this study included only ten MS patients. The SIP, SF-36 MS-QLI have all been validated in populations that included 20–50% cognitively impaired respondents. The level of validation for the SIP and SF-36 were moderate, while the MS-QLI was more extensively validated.

LD

Eight measures have been validated in patients with LD:

- ComQol-I5
- LSS
- LSI
- SF-36
- CQOL
- QUOLIS
- Oregon QLQ
- QoL Interview.

The ComQol-I5 was validated in three studies, which included a total of 430 LD patients.[21,22] The studies included more than 50% of respondents with cognitive impairment. In each case the level of validation was extensive and showed internal consistency scores under 0.60 for the total objective scale and for all of the subscales except 'health'. Test–retest reliability was generally high within 2 weeks, and reported by the authors as 'somewhat erratic' beyond 2 weeks. Patient–proxy assessments showed generally little correspondence, carer estimations correlated more strongly with objective cues than with patient estimates.

The LSS was validated in one study that included 120 LD patients.[42] The study included more than 50% of respondents who were cognitively impaired. The level of validation was moderate and showed that the internal consistency of the measure was lower, but sufficient for people with cognitive impairment. The convergent validity showed that the LSS was sensitive to living conditions (PASSING), but less sensitive to mental deficiencies (ABS) and independence (RCAS).

The LSI was validated in one study that included 48 LD patients.[44] More than 50% of the respondents were cognitively impaired, and the study reported a moderate level of validation for the measure. The internal consistency was acceptable to good for most of the subscales, but poor for 'relationships and social activities'. Construct validity, as measured with the multitrait–multimethod matrix, was good. Convergent validity with other measures was somewhat unclear.

The SF-36 was validated in one study that included 71 LD patients.[41] The study included more than 50% of respondents with cognitive impairment. A moderate level of validation was reported, which showed that the Cronbach's alpha scores and item total biserial correlations for each subscale were largely acceptable, with the exception of the 'social functioning' subscale. Both the mean item-total correlation and Cronbach's alpha (0.33) for this subscale were low. Using factor analysis, a two-factor solution (physical health and mental health) was obtained for the instrument. Inter-rater reliability between patients with LD and their carers was moderate, and the 'role-physical' and 'role-emotional' subscales were considerably less reliable.

The CQOL was validated in one study that included 26 mothers of LD patients.[45] More than 50% of the LD patients were cognitively impaired, and the study reported a moderate level of validation. The internal consistency for combined scales was good, and test–retest reliability was acceptable. Convergent and discriminant validity results were as predicted.

The QUOLIS was validated in one study that included ten LD patients, more than 50% of whom were cognitively impaired.[46] The study reported a moderate level of validation, showing that the internal consistency of the measure was good and inter-rater agreement acceptable for 73% of the scores.

The Oregon QLQ was validated in one study that included 50 LD patients,[50] 20–50% of whom were cognitively impaired, and the level of validation was limited. The internal consistency was below 0.5 for seven out of 17 of the subscales, and overall the inter-rater reliability was described by the authors as 'satisfactory'.

The QoL-Interview was validated in two studies, which included a total of 232 LD patients.[56,57] However, in both studies less than 20% of the respondents were cognitively impaired.

Summary

In LD patients, six measures were found that have been validated in populations where more than 50% of the patients were cognitively impaired: ComQol-I5, LSS, LSI, SF-36, CQOL and QUOLIS. The ComQol-I5 was the most extensively validated of the six measures, with a number of the validity assessments reported as positive, suggesting the measure was valid in this population. Only a moderate level of validation was reported for the LSS, LSI and SF-36 measures and the results were less favourable but fair. The QUOLIS was validated in only ten LD patients, and the CQOL was validated in 26 mothers of LD patients. Again both of the measures had a moderate level of validation and the results were less favourable but fair. The Oregon QLQ was validated in a population where 20–50% of the respondents were cognitively impaired, and the level of validation was limited.

Brain injury

Five measures were validated in patients with brain injury:

- SIP
- BICRO-39
- VAS-QOL
- Life Satisfaction (Viitanen)
- PCRS.

The SIP was validated in five studies, which included a total of 576 brain injury patients. One study included 20–50% of respondents who were

cognitively impaired.[24] A moderate level of validation was reported. The correlations between patient scores and those of 'close others' were high for all categories, with the exceptions of the 'emotion', 'sleep and rest' and 'household management' subscales. The four remaining studies included less than 20% of respondents with cognitive impairment.[30,32,33,35]

The VAS-QoL was validated in one study, which included 92 brain injury patients.[43] More than 50% of the respondents were cognitively impaired, and the level of validation was limited. Strong correlations existed between the VAS-QoL and 'mood', 'physical and social functioning' and 'severity of disability', with a moderate correlation between VAS-QoL and the 'time since injury'.

The BICRO-39,[25] the PCRS[58] and Life Satisfaction (Viitanen)[20,54] were each validated in one study, which included 235, 53 and 19 brain injury patients, respectively. All of the populations included less than 20% of respondents with cognitive impairment.

Summary

In brain injury patients only the VAS-QoL was validated in a population of MS patients that included more than 50% of respondents with cognitive impairment.[43] However, the level of validation of the measure was limited, although it did show positive results in terms of construct validity. The SIP was validated in one study where 20–50% of the respondents were cognitively impaired.[24] The level of validation was moderate and results were fair.

Chapter 4
Discussion

The aim of this review was to identify which general health status measures have been validated in patients with cognitive impairment, to assess the extent of this validation, and to assess the implications of these findings for the use of existing measures and for future primary research in this area.

Scope of the review

First, the scope of the review was identified. In particular, terms such as 'general health status measures', 'cognitive impairment' and the extent of the validation were defined. The expert panel was consulted to define these concepts and to assess the scope of the review.

Although, only studies that either definitely or probably included people with cognitive impairment were considered, it was still uncertain what percentage of respondents were actually cognitively impaired. The included studies were divided into three groups: studies including at least 50% of respondents with cognitive impairment, studies including 20–50% of respondents with cognitive impairment, and studies including less than 20% of respondents with cognitive impairment. For instance, in one study only two out of 98 stroke patients refused to cooperate; it was assumed, therefore, that people with cognitive impairment were probably included. However, there was no indication of how many people with cognitive impairment were included, so it was assumed that less than 20% of respondents had cognitive impairment.[55]

Most studies did not report any information on the level of cognitive impairment experienced by the respondents. These studies were therefore considered as having less than 20% of respondents with cognitive impairment. When the level of cognitive impairment was reported, this was done in various ways. Four studies reported an MMSE score, with a score below 24 generally considered to be an indication of cognitive impairment.[19,35,38,39,47] Two studies reported an age equivalent to describe the level of cognitive impairment of the participants,[21,22,44] and two other studies reported the respondents' IQ.[31,45] Other studies used a self-report measure, such as the London Handicap Scale[48] or the Incapacity Status Scale,[52,53] or the respondents were described in terms of severely affected, having severe neurological impairment or having cognitive difficulties. In one study the level of cognitive impairment of the respondents was assessed by a neurologist and by self-report; however the results of these assessments were not reported.[43]

As non-validated measures are of little interest to this review, only those measures that have been validated on at least one of the five steps identified by Feinstein in evaluating clinimetric tests (i.e. choice of component variables, sensibility, consistency, accuracy and suitability),[18] or the extent to which an instrument provides the same results between patients themselves and proxies (patient–proxy agreement) were included.

The last item was added because the assessment of general health status in cognitively impaired people using third parties or proxies is an important issue. However, few data relating to the use of proxies in general health status measurement were identified for cognitively impaired persons.

Summary of research evidence

Eighty-three publications met the inclusion criteria, and data were extracted from these reports. In total 34 different general health status measures were described in the 83 publications, with the SIP and the SF-36 the most frequently used measures (20 and 19 studies, respectively).

Although 98 validity assessments of a general health status measure were carried out in people with acquired brain injuries and LD, sufficient information was only provided for 52 of the validity assessments to determine whether they definitely or probably included people with cognitive impairment.

The content of the general health status measures ranged from 136 items and 12 dimensions in the SIP to a one-dimensional VAS. It would seem likely that the more comprehensive measures would be most difficult for cognitively impaired respondents to complete. However, different studies need

different ways of assessing general health status. It is important, therefore, that general health status measures aimed at specific populations with various contents are validated for people with cognitive impairment. The content of each general health status measure included in this study is described in appendix 6.

Most studies had low quality scores. In most cases this was due to the fact that the study was not designed to validate the general health status measure(s) used. A large number of measures (nine out of 23) scored 0.25 or less, and six measures scored 0.50 or more. These were:

- the ComQol-I5,[21,22] which was validated in a population where more than 50% of the respondents were cognitively impaired
- the MS-QLI[23] and the SIP,[14,23,24] where 20–50% of the respondents had cognitive impairment
- the BICRO-39,[25] SF-36,[15,26–29] SIP,[30–36] EQ-5D,[13,26,27,37] MSQOL-54[38,39] and LSS,[40] where less than 20% of the respondents had cognitive impairment. Four of these measures were new (ComQol-I5, MS-QLI, BICRO-39 and MSQOL-54).

This review mainly focused on studies in which at least 20% of the respondents had cognitive impairment, and the instrument had a quality score of 0.25 or more, based on all relevant studies. This yielded validity assessments of eight different measures in 13 publications:

- ComQol-I5[21,22]
- SF-36[23,41,47]
- LSI[44]
- LSS[42]
- CQOL[45]
- QUOLIS[46]
- SIP[14,23,24]
- MS-QLI.[23]

Six measures (SF-36, ComQol-I5, LSS, LSI, CQOL and QUOLIS) were validated in studies that included populations in which more than 50% of the people were cognitively impaired, and where the level of validation was more extensive (quality score > 0.2).

- For the ComQol-I5 the internal consistency was under 0.60 for the total objective scale and for all subscales except 'health'. Test–retest reliability was generally high within 2 weeks, and reported by the authors as 'somewhat erratic' beyond 2 weeks. Patient–proxy assessments showed generally little correspondence; carer estimations correlated more strongly with objective cues than with patient estimates.
- The results from the SF-36 study, which included more than 50% of respondents with cognitive impairment, showed that the internal consistency and item total biserial correlations for each subscale were largely acceptable. The only exception was the social functioning scale where both the mean item-total correlation and the Cronbach's alpha score (0.33) were low.[41] Inter-rater reliability between patients with LD and their carers was moderate. However, the 'role-physical' and the 'role-emotional' subscales were considerably less reliable.
- The internal consistency for the LSS was 0.74 for the high-IQ group and 0.61 for the low-IQ group, which showed that the internal consistency was lower but sufficient for people with cognitive impairment, according to the authors. Convergent validity showed that the LSS was sensitive to living conditions (PASSING), and less sensitive to mental deficiencies (ABS) and independence (RCAS).
- For the LSI the validity was assessed only in 48 LD patients. Internal consistency was acceptable to good for most scales, but poor for 'relation-ships and social activities'. Convergent validity results were somewhat unclear.
- For the CQOL validity was assessed only in 26 mothers of LD patients. The results showed that the internal consistency for combined scales was good, and the test–retest reliability was acceptable. Convergent and discriminant validity results were as predicted.
- For the QUOLIS validity was assessed in only ten LD patients. The results showed that internal consistency was good and that inter-rater agreement was acceptable for 73% of the scores.

Two measures (the SIP and MS-QLI) were validated in studies that included 20–50% of cognitively impaired respondents. The level of validation was extensive for these measures; however, it was uncertain whether the results applied to people with cognitive impairment.

- The SIP results showed that the internal consistency for the total SIP score was generally good to excellent. For SIP subscales the Cronbach's alpha scores were lower, and the internal consistencies of the 'sleep and rest' and 'eating' subscales were relatively poor. Correlations between SIP subscales and quantitative measures of function were not entirely predictable. Agreement between patient and proxy scores ranged from moderate to excellent. However, proxy scores were systematically higher than patient scores. The 'emotion' and 'sleep and

rest' subscales were the most difficult categories for relatives to assess.
- The MS-QLI results showed that the internal consistency of the total MS-QLI and the subscales was good, with the exception of the SF-36 'social functioning' subscale. The measure also had good construct validity.

Three measures (VAS-QoL, Life Satisfaction (Cantril's ladder) and the SWB) were validated in studies that included more than 50% of respondents with cognitive impairment. However, the level of validation was limited (quality score ≤ 0.2).

- For the VAS-QoL strong correlations were found between 'mood', 'physical and social functioning' and 'severity of disability', and a moderate correlation with the 'time since injury'.
- For the Life Satisfaction (Cantril's ladder) the results showed that the internal consistency of the instrument was good in people with cognitive impairment.
- For the SWB the results showed that the internal consistency of the instrument was good in people with cognitive impairment.

Nine measures (EQ-5D, MSQOL-54, QoL-Index, Life Satisfaction (Viitanen), QoL-Interview, PCRS, BICRO-39, QoL for MS, and Ho's QoL) were only validated in studies where less than 20% of the respondents were cognitively impaired. For these measures it was unclear whether the findings applied to people with cognitive impairment. Three measures (NHP, COOP/WONCA charts and the Oregon QLQ) were validated in studies, where 20–50% of the respondents were cognitively impaired, but the level of validation was limited (quality score ≤ 0.2).

The use of proxies in general health status measurement

The use of proxies in general health status measurement is an important and controversial issue. Only seven studies looked at the use of proxy measurements and of these only two were considered likely to reflect the situation in cognitively impaired people (see chapter 3). One study used SIP in stroke patients and the other used ComQol-I5 in people with LD. The study conducted in stroke patients suggested patient–proxy agreement was poor, and the other study found moderate-to-excellent patient–proxy agreement for people with LD. Less than 20% of the participants in the other five studies were cognitively impaired and so these data may not relate to cognitively impaired persons. Therefore little evidence was found on which to base recommendations about the use of proxies in general health status measurement.

General health status measures in economic evaluations

Extensive efforts were made to ensure that all relevant economic evaluations relating to the validity of general health status measurement in people with cognitive impairments were identified. In particular the following economic databases were searched: Economic Literature Index (EconLit), BibEc, the Economics Working Paper Archive, and IDEAS. In addition the following search terms were included to identify economic evaluations: 'utility index', 'quality-adjusted life-years', 'contingent valuation', 'human capital', and 'health life equivalents' (see appendix 2). Despite these efforts, no economic evaluations were found. However, Brazier and co-workers have recently published an extensive review of the use of health status measures in economic evaluations.[7] Although they do not address the issue of economic evaluations in people with cognitive impairments, their review can be used in combination with our findings to draw conclusions about the applicability of general health status measures in economic evaluations in people with cognitive impairments.

In the review a distinction is made between instruments that generate scores based on people's preferences (e.g. EQ-5D) and instruments that use arbitrary scoring procedures (e.g. SF-36, which assumes equal weighting for most items).[7] Preference-based health status measures are known as multi-attribute utility scales. These produce a single index score for each state of health which can have a value of one or less, where one is equivalent to full health and zero is death. These scores are known as health state utilities and are used to calculate QALYs. These scores are used in cost–utility analysis.[7]

Brazier and co-workers state that non-preference-based health status measures are not designed for use in economic evaluation and have a number of problems which make them unsuitable for use in economic evaluations. The main objection is that they do not reflect patient preferences and therefore they recommend that a preference-based measure be used alongside a health status measure in trials where the intention is to undertake an economic evaluation.[7] The following techniques for valuing health states were reviewed: standard

gamble, time trade-off, VAS, magnitude estimation and person trade-off. When considering theoretical validity it was concluded that only choice-based techniques should be used, that is, standard gamble, time trade-off and person trade-off.[7]

Five preference-based measures of health used in economic evaluation were reviewed: the QWB scale, Rosser's disability/distress scale, the HUI (mark I to III), the EQ-5D (EuroQoL) and the 15D. The QWB, Rosser's scale and 15D were regarded as inferior to the other two measures because their values were not obtained using one of the choice-based techniques. The HUI and EQ-5D use different methods of eliciting weights (standard gamble and time trade-off, respectively) and there is currently no consensus among health economists as to which is better.[7]

Given the evidence compiled by Brazier and co-workers and the findings presented in this review, two conclusions can be drawn on the use of general health status measures for economic evaluations in people with cognitive impairments. First, Brazier and co-workers concluded that only choice-based techniques should be used for valuing heath states.[7] However, from the evidence identified in the current review, the feasibility of using choice-based techniques in people with cognitive impairments has never been assessed. The techniques are complicated and in most cases the choices very difficult to make. It is uncertain whether the use of these techniques could be successful in people with cognitive impairments. Second, the EQ-5D and HUI were found to be superior compared with other preference-based measures of health.[7] Neither the HUI,[12] nor the EQ-5D[13,26,27,37] has been assessed in populations with more than 20% of people with cognitive impairments, therefore the validity of these instruments for people with cognitive impairments could not be assessed.

Chapter 5
Conclusions

Very few of the measures have been specifically validated in cognitively impaired respondents. Studies that involved at least 50% of respondents with cognitive impairment generally showed poorer validity results compared with studies with fewer cognitively impaired persons. This suggests that general health status measures designed for the general population are not automatically suitable for people with cognitive impairment. The few measures that were specifically developed for people with cognitive impairment also showed poor results. **Therefore the main conclusion from this report is that there are no validated instruments available for use in cognitively impaired respondents; existing measures, specifically designed for use in these populations, should be used with caution.**

The measures that looked most promising were the SF-36 for LD, MS and stroke patients, the MS-QLI for MS patients, and the LSS for LD patients. The SF-36 was evaluated in 71 LD patients (> 50% cognitively impaired), 96 stroke patients (20–50% cognitively impaired) and 300 MS patients (20–50% cognitively impaired). The results showed acceptable levels of internal consistency (> 0.70) and good construct validity for most subscales. The MS-QLI was more extensively validated in 300 MS patients and the results were relatively good, except for the 'social function' subscale. However, only 20–50% of the respondents in this study were cognitively impaired. The level of validation for the LSS was moderate in 120 patients with LD and other conditions, where at least 50% of the respondents had cognitive impairment. Overall, the results were acceptable.

Only two studies described the differences in an instrument's validity for groups of patients with different levels of cognitive impairment. One study among 120 LD patients compared the internal consistency of the LSS between respondents with high-IQ scores (Cronbach's alpha = 0.74) and low-IQ scores (0.61).[42] The other study compared mean completion time of the MSQOL-54, help needed, number of missing values and number of inconsistencies between 18 MS patients with cognitive impairment (MMSE score ≤ 23.8) and the remaining 186 MS patients.[38] In this study MS patients with cognitive impairment needed more time to complete the MSQOL-54 (23 minutes versus 19 minutes for non-impaired), more patients needed help (67% versus 38%), there were more missing items (5.5–33.3% versus 0.5–2.9%), and more logical inconsistencies (50% none and 67% one or less versus 66% and 82%). These results clearly indicate that there are differences in the validity of general health status measures for people with and without cognitive impairment, which have to be taken into account when assessing health status in cognitively impaired patients.

It is difficult to speculate why some measures are better than others. Issues such as instrument design, complexity and the baseline validity with healthy adults have not yet been explored. How the severity and nature of the impairment affects the effectiveness of the measure, and what the most appropriate uses of a tool are given its limitations, also need to be addressed in future studies. Likewise, the structure or length of an instrument may affect its validity for use in patients with cognitive impairment. However, this review did not find any research into the comprehensibility of different formats, so again, this needs to be assessed in future studies. For example, it would be useful to look at the length of an instrument and whether it predicts test–retest reliability. Similarly, the format of an instrument may affect the accuracy of its use with proxies and this may be relevant in cases of severe cognitive impairment, where the use of proxies is required. However, little evidence relating to the use of proxies was identified in this review.

It may seem logical that measures using short and simple formats (such as EQ-5D) may provide more reliable and consistent data compared with measures with longer formats, such as the SF-36 and NHP. However, there was no evidence to support this assumption, mainly due to the fact that few short and simple formats were assessed in patients with cognitive impairment.

There was some evidence to suggest that social dimensions are difficult to assess in patients with cognitive impairment. The domain of 'social function' had a low internal consistency in studies evaluating the MS-QLI[23] and the SF-36;[41] the domain of 'relationships and social activities' had poor internal consistency in a study evaluating the

LSI;[44] and the domain of 'place in the community' had poor internal consistency in a study evaluating the ComQol-I5.[21,22] These findings suggest that there are considerable differences in the reliability and also probably the validity of these subscales. The extent of these differences needs to be examined in future research.

In the current review the assessment of validity focused mainly on the extent to which the instrument correlated with measures from other instruments. There were no references as to how the measures correlated with clinical indicators of disease such as degree of atherosclerosis, respiratory function, or use of analgesics. These measures could give an indication of disease status, which may in turn affect general health status. For example, it would be useful to demonstrate that hemiplegic patients with cognitive impairment accurately report difficulties with mobility, or that patients with respiratory disease report poor physical function. Alternatively, if there were evidence that stroke patients with cognitive impairment reported better general health status than stroke patients without cognitive impairment, it would make one suspicious regarding the validity of the measure. Unfortunately, this information was not available in the studies included in this review and needs to be addressed in future studies.

In order to determine the suitability and validity of general health status measures it is important that studies carry out an objective assessment of the cognitive function of respondents (even something as simple as the MMSE). Unfortunately, few of the studies included in this review carried out this assessment. This needs to be addressed in future research. It would also be useful to know if the cognitive test scores affect factors such as test–retest reliabilities and internal consistency. A person who does not understand the questions, or has difficulty interpreting the layout of the instrument will probably produce inconsistent responses. Obviously a cognitive impairment may also affect the validity of an instrument and it would be interesting to know more about instrument validity for patients with differing levels of cognitive impairment. As no such studies exist, this is a major recommendation for future research.

Cognitive impairment can take many different forms. The type and severity of a cognitive impairment will determine the degree to which a person's ability to accurately self-report their health status is affected. Different aetiologies are also associated with different patterns of cognitive impairment. Stroke often produces relatively localised damage whereas dementia is commonly associated with global cerebral atrophy. This has implications for the nature of cognitive impairment seen in different diseases and so will, in turn, determine which instruments are appropriate for different patients. The degree to which these differences affect the ability of individuals to complete general health status measures has yet to be established.

With respect to the use of general health status measures for economic evaluations in people with cognitive impairments we can draw two conclusions. First, Brazier and co-workers concluded that only choice-based techniques should be used for valuing health states.[7] However, as far as we know, the feasibility of using choice-based techniques in people with cognitive impairments has never been assessed. The techniques are complicated and in most cases the choices are very difficult to make. It is uncertain whether the use of these techniques could be successful in people with cognitive impairments. Second, the EQ-5D and HUI were found to be superior to other preference-based measures of health.[7] Neither the HUI,[12] nor the EQ-5D[13,26,27,37] has been assessed in populations with more than 20% of people with cognitive impairments; therefore the validity of these instruments for people with cognitive impairments could not be assessed.

Limitations of the review

There are two main limitations to this review. First, in most cases there was very limited information on the number of respondents who actually were cognitively impaired. Objective validated psychometric tests and a neurologists' diagnoses are the best markers of cognitive impairment; however these were only provided in a small number of studies. In most cases presumed impairment based upon epidemiology was relied on, which is a weaker basis for assessing the number of cognitively impaired respondents included. Therefore decisions concerning inclusion or exclusion, based on the number of cognitively impaired respondents included in the study, were arbitrary and may be subject to debate.

Second, the extent of the validation was limited in most cases. The fact that the level of validation was limited for general health status measures in people with cognitive impairment is reflected in the main conclusion of the review. Most information

on the validity of instruments seems to concentrate on measures of consistency, which is a very limited way of assessing the validity of general health status measures. It should be noted that reports of cognitively impaired people completing an instrument in a coherent way does not mean that the instrument can be regarded as a valid tool for use with this population. A proper validation should also include measures of accuracy and suitability.

There are other limitations that also need to be addressed. The instrument that was used to assess the quality of the included validity assessments was a self-developed instrument. The instrument itself has not been validated, so the quality scores should be interpreted with caution. Five of the six items, however, were derived from Feinstein's guidelines on evaluating clinimetric tests.[18] Patient–proxy agreement and inter-rater agreement were added as a dimension to assess the validity of a measure because it was assumed that agreement between patients and proxies and between different raters shows the strength of the measure. However, as mentioned before, it should be noted that agreement between two people does not automatically make their judgements more accurate. Therefore the outcomes of patient–proxy and inter-rater agreement should be interpreted with caution.

There may also be some debate as to how various disease states have been defined and which have been included or excluded. Although the expert panel was consulted on this matter, there remained some discussion between the panel members on certain issues. Included in the review were studies that validated general health status measures in people with cognitive impairment due to acquired brain injury (caused by stroke, trauma, MS or other causes) and LD. Although LD can be secondary to road traffic accident or due to a prenatal brain injury, it was decided to treat LD separately from acquired brain injury. Also, although MS is a degenerative disease of the whole nervous system and is not specifically an acquired brain injury, it was decided to treat MS as an acquired brain injury.

Implications for practice

The main conclusion from this review is that few measures have been validated specifically for cognitively impaired respondents, and there is little evidence available on which to base recommendations for practice. The main implications for practice are therefore as follows.

- Existing general health status measures should be used with caution in individuals with cognitive impairment.
- There is no evidence to indicate the most suitable general health status measure for use in economic evaluations of cognitive impairment.
- There is little evidence to support the validity of proxy assessments in cognitively impaired populations.

Based on the limited evidence identified in this review the following brief statements summarise the current position of general health status measurement in cognitively impaired persons. However, these should be viewed with caution and are not intended to be definitive implications for practice.

- Overall, the SIP and MS-QLI appear to be the most extensively validated measures in populations containing relatively large numbers of cognitively impaired persons, with both measures appearing to be reasonably valid, based on the limited evidence available.
- The ComQol-I5 and the SF-36 also appear to be extensively validated compared with other measures, but their study populations did not always contain large numbers of cognitively impaired persons. The ComQol-I5 showed some weaknesses in terms of its validity in cognitively impaired populations compared with the SF-36.
- Compared with other measures, the VAS-QoL and the SIP appear to be the most validated measures for use in patients suffering from brain injury, but the studies included limited numbers of cognitively impaired patients (i.e. 20–50%).
- The SIP appears to be the most extensively validated measure for use in stroke patients, but the studies included limited numbers of cognitively impaired patients (i.e. 20–50%).
- The MS-QLI appears to show the greatest validity in MS patients, but again the studies included limited numbers of cognitively impaired patients (i.e. 20–50%).
- Compared with the other disease states, most information relates to people with LD, in particular with regard to the ComQol-I5, LSS, LSI, SF-36, CQOL and QUOLIS. The ComQol-I5, LSS, LSI, and SF-36 all appear acceptable for use in patients with cognitive impairment, while the ComQol-I5 is the most extensively validated measure out of the six. However, some of the measures contain subscales that seem to show some weaknesses in people with cognitive impairment.

Recommendations for further research

- There is a need for the development of new health status measures for use in people with cognitive impairment, particularly for people with acquired brain injury due to stroke, MS or trauma.
- Existing general health status measures, such as the SIP, SF-36, EQ-5D and NHP need to be validated for people with cognitive impairment. More research is needed into how these existing measures could be modified so that they are more suitable for people with cognitive impairment.
- The validity of general health status measures for use in economic evaluations needs to be assessed in cognitively impaired populations, as well as the feasibility of using choice-based techniques in people with cognitive impairments.
- Several large trials have been performed that included respondents with cognitive impairments; however data on the validity of general health status measures were not reported separately for these respondents.[19,26,27] If data are still available, subgroup analyses should be done to validate existing general health status measures for people with cognitive impairment.
- For some populations general health status can only be measured through proxy assessments. Therefore, health status measures need to be validated for use by proxies in these populations.
- The structure and length of an instrument, as well as its design and complexity, may affect its validity for use in patients with cognitive impairment. This needs to be tested in future studies.
- In many cases validations of general health status measurements for people with cognitive impairment have previously been the secondary focus of studies. In future, this issue should be addressed in studies specifically designed for this purpose.
- In studies aimed at the validation of general health status measures for people with cognitive impairment, objective validated psychometric tests or a neurologist's diagnosis should assess the level of cognitive impairment. Separate analyses should be performed to assess the validity of the instrument for different levels of cognitive impairment.
- The validation of general health status measures should include information on the choice of component items, sensibility, consistency, accuracy and suitability. When there is need for proxy assessments the instrument should be assessed for patient–proxy agreement and inter-rater agreement. Studies should include a large number of respondents with different levels of cognitive impairment, so that differences in the instrument's validity for different groups of people with cognitive impairment can be assessed.
- In most studies the assessment of validity focused on the extent to which the instrument correlated with measures from other instruments. No references to clinical indicators of disease, such as degree of atherosclerosis, respiratory function, or use of analgesics, were found. These measures could give an indication of disease status and so, in turn, may affect general health status. This needs to be assessed in future research.

Based on these recommendations the following characteristics of an ideal study are suggested:

- an adequate number of participants in order to provide sufficient power to validate the health status measure
- a definition of the level of cognitive impairment of participants using either a validated psychometric test (e.g. MMSE), a neurologist's examination or both
- separate subgroup analyses according to the level of cognitive impairment experienced by participants
- a comparison group consisting of individuals with the same primary disease (i.e. stroke, brain injury, MS) but no cognitive impairment. This may not always be feasible depending on the aetiology of the underlying disease; however the value of any health status measure that fails to identify the two separate groups of individuals must be questionable
- studies should consider all of the six validity criteria outlined in this review (i.e. choice of component items, sensibility, consistency, accuracy, suitability, and proxy assessment where appropriate).

Acknowledgements

We would like to thank the following expert panel members for commenting on drafts of the protocol and review:

Dr Sube Banerjee, Senior Lecturer in Psychiatry (Co-investigator on the HTA proposal to develop HRQL measures in dementia)

Ms Carol Boys, Director Down's Syndrome Association

Miss Tig Calvert, Lecturer in Psychology, (currently developing outcome measures in brain injury)

Miss Margaret Goose, Chief Executive of the Stroke Association

Dr Jeremy Hobart, Clinical Lecturer in Neurology, Neurological Outcomes Measurement Unit

Professor Paul Kind, Centre for Health Economics

Dr Andrew Lloyd, OCHRAD, Oxford Brookes University

Dr Sue Macran, Centre for Health Economics

Dr Jane O'Dwyer, Senior Lecturer, Learning Disability Psychiatry, School of Health and Related Research, Psychiatry

Dr Mark Petticrew, Associate Director MRC Social and Public Health Sciences Unit.

We would like to acknowledge in particular the contribution of Dr Mark Petticrew who was responsible for preparing the original research proposal and was actively involved in writing the review protocol.

This review was funded by the HTA Programme and we are grateful for the comments received from the referees. The final contents are the sole responsibility of the authors.

References

1. Kessler RC, Mrozczk DK. Some methodological issues in the development of quality of life measures for the evaluation of medical interventions. *J Eval Clin Pract* 1996;**2**(3):181–91.
2. Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. *Ann Intern Med* 1993;**118**(8):622–9.
3. Sbordone RJ, Seyranian GD, Ruff RM. Are the subjective complaints of traumatically brain injured patients reliable? *Brain Injury* 1998;**12**(6):505–15.
4. Hillier SL, Metzer J. Awareness and perceptions of outcomes after traumatic brain injury. *Brain Injury* 1997;**11**(7):525–36.
5. Livingston MG, Livingston HM. The Glasgow Assessment Schedule: clinical and research assessment of head injury outcome. *Int Rehabil Med* 1985;**7**(4):145–9.
6. Knight RG, Devereux R, Godfrey HP. Caring for a family member with a traumatic brain injury. *Brain Injury* 1996;**12**(6):467–81.
7. Brazier J, Deverill M, Green C, Harper R, Booth A. A review of the use of health status measures in economic evaluation. *Health Technol Assess* 1999;**3**(9).
8. Howard S, Spencer A. Effects of resettlement on people with learning disabilities. *Br J Nurs* 1997;**6**(8):436–41.
9. Donnelly M, McGilloway S, Mays N, Knapp M, Kavanagh S, Beecham J, *et al.* One and two year outcomes for adults with learning disabilities discharged to the community. *Br J Psychiatry* 1996;**168**(5):598–606.
10. Cullen D, Whoriskey M, Mackenzie K, Mitchell W, Ralston K, Shreeve S, *et al.* The effects of deinstitutionalization on adults with learning disabilities. *J Intellect Disabil Res* 1995;**39**(6):484–94.
11. Dockrell JE, Gaskell GD, Normand C, Rehman H. An economic analysis of the resettlement of people with mild learning disabilities and challenging behaviour. *Soc Sci Med* 1995;**40**(7):895–901.
12. Mathias SD, Bates MM, Pasta DJ, Cisternas MG, Feeny D, Patrick DL. Use of the Health Utilities Index with stroke patients and their caregivers. *Stroke* 1997;**28**(10):1888–94.
13. Dorman PJ, Waddell F, Slattery J, Dennis M, Sandercock P. Is the EuroQol a valid measure of health-related quality of life after stroke? *Stroke* 1997;**28**(10):1876–82.
14. Sneeuw KC, Aaronson NK, de Haan RJ, Limburg M. Assessing quality of life after stroke. The value and limitations of proxy ratings. *Stroke* 1997;**28**(8):1541–9.
15. Anderson C, Laubscher S, Burns R. Validation of the Short Form 36 (SF-36) health survey questionnaire among stroke patients. *Stroke* 1996;**27**(10):1812–16.
16. Melzer D. Profile of disability in elderly people: estimates from a longitudinal population study. *BMJ* 1999;**318**:1108–11.
17. Whitehouse PJ, Orgogozo JM, Becker RE, Gauthier S, Pontecorvo M, Erzigkeit H, *et al.* Quality-of-life assessment in dementia drug development. Position paper from the International Working Group on Harmonization of Dementia Drug Guidelines. *Alzheimer Dis Assoc Disord* 1997;**3**:56–60.
18. Feinstein AR. Clinimetrics. London: Yale University Press; 1987.
19. Ormel J, Kempen GIJM, Deeg DJH, Brilman EI, Van Sonderen E, Relyveld J. Functioning, well-being, and health perception in late middle-aged and older people: comparing the effects of depressive symptoms and chronic medical conditions. *J Am Geriatr Soc* 1998;**46**(1):39–48.
20. Koskinen S. Quality of life 10 years after a very severe traumatic brain injury (TBI): the perspective of the injured and the closest relative. *Brain Injury* 1998;**12**:631–48.
21. Cummins RA, McCabe MP, Romeo Y, Reid S, Waters L. An initial evaluation of the Comprehensive Quality of Life Scale – Intellectual Disability. *Int J Disabil Dev Educ* 1997;**44**(1):7–19.
22. Cummins R. The comprehensivec Quality of life scale – Intellectual Disability (5th. ed.) (ComQol-I5): Manual. Toorak: Deakin University School of Psychology; 1997.
23. Fischer JS, LaRocca NG, Miller DM, Ritvo PG, Andrews H, Paty D. Recent developments in the assessment of quality of life in multiple sclerosis (MS). *Multiple Sclerosis* 1999;**5**(4):251–9.
24. Smith HE. Head injury follow-up: is the Sickness Impact Profile a useful clinical tool. *Clin Rehabil* 1992;**6**(1):31–9.
25. Powell JH, Beckers K, Greenwood RJ. Measuring progress and outcome in community rehabilitation after brain injury with a new assessment instrument – the BICRO-39 scales. Brain Injury Community Rehabilitation Outcome. *Arch Phys Med Rehabil* 1998;**79**(10):1213–25.

26. Dorman P, Slattery J, Farrell B, Dennis M, and Sandercock P. Qualitative comparison of the reliability of health status assessments with the EuroQol and SF-36 questionnaires after stroke. United Kingdom Collaborators in the International Stroke Trial. *Stroke* 1998;**29**(1):63–8.

27. Dorman PJ, Slattery J, Farrell B, Dennis MS, Sandercock PA. A randomised comparison of the EuroQol and Short Form-36 after stroke. United Kingdom collaborators in the International Stroke Trial. *BMJ* 1997;**315**(7106):461.

28. Pickard EA. Replicability of SF-36 summary scores by the SF-12 in stroke patients. *Stroke* 1999;**30**(6):1213–17.

29. Vickrey BG, Hays RD, Genovese BJ, Myers LW, Ellison GW. Comparison of a generic to disease-targeted health-related quality-of-life measures for multiple sclerosis. *J Clin Epidemiol* 1997;**50**(5):557–69.

30. Dikmen SS, Ross BL, Machamer JE, Temkin NR. One year psychosocial outcome in head injury. *J Int Neuropsychol Soc* 1995;**1**:67–77.

31. Baird AD, Brown GG, Adams KM, Shatz MW, McSweeny AJ, Ausman JI, *et al.* Neuropsychological deficits and real-world dysfunction in cerebral revascularization candidates. *J Clin Exp Neuropsychol* 1987;**9**(4):407–22.

32. Burton LA, Volpe BT. Depression after head injury: do physical and cognitive sequelae have similar impact? *J Neurol Rehabil* 1994;**8**(2):63–7.

33. Temkin N, McLean A, Jr., Dikmen S, Gale J, Bergner M, Almes MJ. Development and evaluation of modifications to the Sickness Impact Profile for head injury. *J Clin Epidemiol* 1988;**41**(1):47–57.

34. Visser MC, Koudstaal PJ, Erdman RA, Deckers JW, Passchier J, van Gijn J, *et al.* Measuring quality of life in patients with myocardial infarction or stroke: a feasibility study of four questionnaires in The Netherlands. *J Epidemiol Community Health* 1995;**49**(5):513–17.

35. Smith JL, Magill Evans J, Brintnell S. Life satisfaction following traumatic brain injury. *Can J Rehabil* 1998;**11**(3):131–40.

36. Zeldow PB, Pavlou M. Physical and psychosocial functioning in multiple sclerosis: descriptions, correlations, and a tentative typology. *Br J Med Psychol* 1988;**61**(2):185–95.

37. Dorman PJ, Waddell F, Slattery J, Dennis M, Sandercock P. Are proxy assessments of health status after stroke with the EuroQol questionnaire feasible, accurate, and unbiased? *Stroke* 1997;**28**(10):1883–7.

38. Solari A, Ghezzi A, Mendozzi L, Filippini G, Cifani S, Barbieri E, *et al.* Relation of cognitive impairment and depression to quality of life in multiple sclerosis patients. *Ital J Neurol Sci* 1998;**19**:S392–S398.

39. Solari A, Filippini G, Mendozzi L, Ghezzi A, Cifani S, Barbieri E, *et al.* Validation of Italian Multiple-Sclerosis-Quality-of-Life-54 questionnaire. *J Neurol Neurosurg Psychiatry* 1999;**67**(2):158–62.

40. Gulick EE. Correlates of quality of life among persons with multiple sclerosis. *Nurs Res* 1997;**46**(6):305–11.

41. Jones J, Dagnan P, Ruddick L. A pilot study of the use of the SF-36 to assess health status of adults with learning disabilities living in small community based homes. *Br J Dev Disabil* 1997;**43**(1):27–35.

42. Vandergriff DV, Chubon RA. Quality of life experienced by persons with mental retardation in various residential settings. *J Rehabil* 1994;**60**(4):30–7.

43. Kreuter M, Sullivan M, Dahllof AG, Siosteen A. Partner relationships, functioning, mood and global quality of life in persons with spinal cord injury and traumatic brain injury. *Spinal Cord* 1998;**36**(4):252–61.

44. Hawkins BA, Kim K, Eklund SJ. Validity and reliability of a Five Dimensional Life Satisfaction Index. *Ment Retard* 1995;**33**(5):295–303.

45. Graham P, Stevenson J, Flynn D. A new measure of health-related quality of life for children: preliminary findings. *Psychol Health* 1997;**12**(5):655–65.

46. Ouellette Kuntz H. A pilot study in the use of the Quality of Life Interview Schedule. *Soc Ind Res* 1990;**23**(3):283–98.

47. Wilkinson PR, Wolfe CDA, Warburton FG, Rudd AG, Howard RS, Ross Russell RW, *et al.* Longer term quality of life and outcome in stroke patients: is the Barthel index alone an adequate measure of outcome? *Qual Health Care* 1997;**6**(3):125–30.

48. Harwood RH, Gompertz P, Ebrahim S. Handicap one year after a stroke: validity of a new scale. *J Neurol Neurosurg Psychiatry* 1994;**57**(7):825–9.

49. Lam CLK, Van Weel C, Lauder IJ. Can the Dartmouth COOP/WONCA charts be used to assess the functional status of Chinese patients? *Fam Pract* 1994;**11**(1):85–94.

50. Turner JC. A Comparison of the quality of life of developmentally handicapped adults living in community placements and receiving case management or traditional community services (Thesis). Cleveland OH 44106: Case Western Reserve University; 1997.

51. Gompertz P, Pound P, Ebrahim S. Validity of the extended activities of daily living scale. *Clin Rehabil* 1994;**8**(4):275–80.

52. Stuifbergen AK. Cognitive impairment and perceptions of quality of life among individuals with multiple sclerosis. *Rehabil Nurs Res* 1995;**4**(1):11–18.

53. Stuifbergen AK. Health-promoting behaviors and quality of life among individuals with multiple sclerosis. *Scholarly Inquiry for Nursing Practice* 1995;**9**(1):31–50.

54. Kaitaro T, Koskinen S, Kaipo ML. Neuropsychological problems in everyday life: a 5-year follow-up study of young severely closed-head-injured patients. *Brain Injury* 1995;**9**(7):713–27.

55. Ahlsio B, Britton M, Murray V, Theorell T. Disablement and quality of life after stroke. *Stroke* 1984;**15**(5):886–90.

56. Lehman AF, Possidente S, Hawker F. The quality of life of chronic patients in a state hospital and in community residences. *Hosp Community Psychiatry* 1986;**37**(9):901–7.

57. Lehman A. A quality of life interview for chronically mentally ill. *Evaluation and Program Planning* 1988;**11**:51–62.

58. Fleming J, Strong J. A longitudinal study of self-awareness: functional deficits underestimated by persons with brain injury. *Occup Ther J Res* 1999;**30** 19(1):3–17.

59. Chow EO. Quality of life and associated factors in Chinese stroke survivors in Hong Kong. *Health Care in Later Life* 1997;**2**(4):227–38.

Excluded studies

Cella DF, Dineen K, Arnason B, Reder A, Webster KA, karabatsos G, *et al.* Validation of the functional assessment of multiple sclerosis quality of life instrument. *Neurology* 1996;**47**(1):129–39.

Corrigan JD, Smith Knapp K, Granger CV. Validity of the functional independence measure for persons with traumatic brain injury. *Arch Phys Med Rehabil* 1997;**78**(8):828–34.

De Haan R, Horn J, Limburg M, Van Der Meulen J, Bossuyt P. A comparison of five stroke scales with measures of disability, handicap, and quality of life. *Stroke* 1993;**24**(8):1178–81.

Di Fabio RP, Choi T, Soderberg J, Hansen CR. Health-related quality of life for patients with progressive multiple sclerosis: influence of rehabilitation. *Phys Ther* 1997;**77**(12):1704–16.

Duncan PW, Samsa GP, Weinberger M, Goldstein LB, Bonito A, Witter DM, *et al.* Health status of individuals with mild stroke. *Stroke* 1997;**28**(4):740–5.

Ebrahim S, Barer D, Nouri F. Use of the Nottingham Health Profile with patients after a stroke. *J Epidemiol Community Health* 1986;**40**(2):166–9.

Freeman JA, Langdon DW, Hobart JC, Thompson AJ. Health-related quality of life in people with multiple sclerosis undergoing inpatient rehabilitation. *J Neurol Rehabil* 1996;**10**(3):185–94.

Gage BF, Cardinalli AB, Owens DK. The effect of stroke and stroke prophylaxis with aspirin or warfarin on quality of life. *Arch Intern Med* 1996;**156**:1829–36.

Gompertz P, Pound P, Ebrahim S. The reliability of stroke outcome measures. *Clin Rehabil* 1993;**7**(4):290–6.

Granger CV, Cotter AC, Hamilton BB, Fiedler RC. Functional assessment scales: a study of persons after stroke. *Arch Phys Med Rehabil* 1993;**74**(2):133–8.

Hale LA, Eales CJ, Fritz VU. The Soweto Stroke Questionnaire. *S Afr J Physiother* 1998;**54**(4):16–20.

Harper AC, Harper DA, Chambers LW, Cino PM, Singer J. An epidemiological description of physical, social and psychological problems in multiple sclerosis. *J Chron Dis* 1986;**39**(4):305–10.

Hutchinson J, Hutchinson M. The Functional Limitations Profile may be a valid, reliable and sensitive measure of disability in multiple sclerosis. *J Neurology* 1995;**242**(10):650–7.

Indredavik B, Bakke F, Slordahl SA, Rokseth R, Haheim LL. Stroke unit treatment improves long-term quality of life: a randomized controlled trial. *Stroke* 1998;**29**(5):895–9.

Jonsson A, Dock J, Ravnborg MH. Quality of life as a measure of rehabilitation outcome in patients with multiple sclerosis. *Acta Neurol Scand* 1996;**93**:229–35.

King RB. Quality of life after stroke. *Stroke* 1996;**27**(9):1467–72.

Lankhorst GJ, Jelles F, Smits RCF, Polman CH, Kuik DJ, Pfennings L, *et al.* Quality of life in multiple sclerosis: the disability and impact profile (DIP). *J Neurol* 1996;**243**(6):469–74.

Mathias SD, Bates MM, Pasta DJ, Cisternas MG, Feeny D, Patrick DL. Use of the Health Utilities Index with stroke patients and their caregivers. *Stroke* 1997;**28**(10):1888–94.

O'Mahony PG, Rodgers H, Thomson RG, Dobson R, James OF. Is the SF-36 suitable for assessing health status of older stroke patients? *Age Ageing* 1998;**27**(1):19–22.

Parkin D, McNamee P, Jacoby A, Miller P, Thomas S, Bates D. A cost–utility analysis of interferon beta for multiple sclerosis. *Health Technol Assess* 1998;**2**(4):1–45.

Pfennings L, VanderPloeg HM, Cohen L, Bramsen I, Polman CH, Lankhorst GJ, *et al.* A health-related quality of life questionnaire for multiple sclerosis patients. *Acta Neurol Scand* 1999;**100**(3):148–55.

Pfennings L, Cohen L, Miller D, Gerbaud L, Vleugels L, Freeman J, *et al.* Using the Short Form-36 with multiple sclerosis patients in five countries: a cross-cultural comparison. *Psychol Rep* 1999;**85**(1):19–31.

Pfennings L, Cohen L, Ader H, Polman C, Lankhorst G, Smits R, *et al.* Exploring differences between subgroups of multiple sclerosis patients in health-related quality of life. *J Neurol* 1999;**246**(7):587–91.

Pfennings L, Cohen L, van der Ploeg H, Polman C, Lankhorst G. Reliability of two measures of health-related quality of life in patients with multiple sclerosis. *Percept Mot Skills* 1998;**87**(1):111–14.

Provinciali L, Ceravolo MG, Bartolini M, Logullo F, Danni M. A multidimensional assessment of multiple sclerosis: relationships between disability domains. *Acta Neurol Scand* 1999;**100**(3):156–62.

Robnett RH, Gliner JA. Qual-OT: a quality of life assessment tool. *Occup Ther J Res* 1995;**15**(3):198–214.

Rothwell PM, McDowell Z, Wong CK, Dorman PJ. Doctors and patients don't agree: cross sectional study of patients' and doctors' perceptions and assessments of disability in multiple sclerosis. *BMJ* 1997;**314**(7094):1580–3.

Rudick RA, Miller D, Clough JD, Gragg LA, Farmer RG. Quality of life in multiple sclerosis. Comparison with inflammatory bowel disease and rheumatoid arthritis. *Arch Neurol* 1992;**49**(12):1237–42.

Schuling J, De Haan R, Limburg M, Groenier KH. The Frenchay Activities Index. Assessment of functional status in stroke patients. *Stroke* 1993;**24**(8):1173–7.

Schuling J, Greidanus J, Meyboom-De Jong B. Measuring functional status of stroke patients with the Sickness Impact Profile. *Disabil Rehabil* 1993;**15**(1):19–23.

Schwartz CE, Coulthard Morris L, Cole B, Vollmer T. The quality-of-life effects of interferon beta-1b in multiple sclerosis. An extended Q-TWiST analysis. *Arch Neurol* 1997;**54**(12):1475–80.

Schwartz CE, Vollmer T, Lee H. Reliability and validity of two self-report measures of impairment and disability for MS. North American Research Consortium on Multiple Sclerosis Outcomes Study Group. *Neurology* 1999;**52**(1):63–70.

Segal ME, Schall RR. Determining functional/ health status and its relation to disability in stroke survivors. *Stroke* 1994;**25**(12):2391–7.

Temkin NR, Dikmen S, Machamer J, McLean A. General versus disease-specific measures. Further work on the Sickness Impact Profile for head injury. *Med Care* 1989;**27**(3 Suppl):S44–53.

Van Straten A, de Haan RJ, Limburg M, Schuling J, Bossuyt PM, van den Bos GAM. A stroke-adapted 30-item version of the Sickness Impact Profile to assess quality of life (SA-SIP30). *Stroke* 1997;**28**(11):2155–61.

Vazquez Barquero JL, Arias Bal MA, Pena C, Diez Manrique JF, *et al.* El cuestionario "perfil de impacto de la enfermedad" (SIP): version espanola de una medida del estado de salud. The Sickness Impact Profile (SIP): a Spanish version of a health status assessment. *Actas Luso Espanolas de Neurologia y Psiquiatria y Ciencias Afines* 1991;**19**(2):127–34.

Vickrey BG, Hays RD, Harooni R, Myers LW, Ellison GW. A health-related quality of life measure for multiple sclerosis. *Qual Life Res* 1995;**4**(3):187–206.

Viitanen M, Fugl Meyer KS, Bernspang B, Fugl Meyer AR. Life satisfaction in long-term survivors after stroke. *Scand J Rehabil Med* 1988;**20**(1):17–24.

Yoon H. Factors affecting quality of life of the Korean aged stroke patients. *Int J Aging Hum Dev* 1997;**44**(3):167–81.

Appendix 1

Handsearched bibliographies

Sintonen H, editor. Euroqol Conference Proceedings. Helsinki October 1992. Kuopio, Finland: Kuopio University, Department of Social Sciences; 1993.

Quality of Life assessment: an annotated bibliography. Geneva: WHO Division of Mental Health; 1994.

Busschbach JJV, Bonsel GJ, de Charro F, editors. Euroqol plenary meeting Rotterdam 1992 6–8 October. Discussion papers. Rotterdam: Erasmus University, Department of Public Health; 1994.

O'Hanlon M, Buxton M, editors. Euroqol plenary meeting, London October 1994. Uxbridge: Brunel University, Health Economics Research Group; 1995.

Badia X, Herdman M, Segura A, editors. Euroqol plenary meeting, Barcelona 3–6 October 1995. Barcelona: Institut de Salut Publica de Catalunya; 1996.

Nord E, editor. Euroqol plenary meeting, Oslo 17–18 October 1996. Oslo: National Institute of Public Health; 1997.

Psychosocial factors in multiple sclerosis. Proceedings of the Multiple Sclerosis Forum Modern Management Workshop, Rome April 1995. Worthing: Profes-sional Postgraduate Services Europe; 1995.

Spilker B, Molinek FR, Johnston KA, Simpson RL, Tilson HH. Quality of life: bibliography and indexes. *Med Care* 1990;**28** (12 Suppl):DS1–DS77.

Spilker B, White WSA, Simpson RL, Tilson HH. Quality of life bibliography and indexes. 1990 update. *J Clin Res Pharmacoepidemiol* 1992;**6**:87–156.

Spilker B, Simpson RL, Tilson HH. Quality of life bibliography and indexes: 1991 update. *J Clin Res Pharmacoepidemiol* 1992;**6**:205–66.

Berzon RA, Simeon GP, Simpson RL, Donnelly MA, Tilson HH. Quality of life bibliography and indexes: 1992 update. *J Clin Res Drug Dev* 1993;**7**:203–42.

Berzon RA, Simeon GP, Simpson RL, Donnelly MA, Tilson HH. Quality of life bibliography and indexes: 1993 update. *Qual Life Res* 1995;**4**:53–73.

Berzon RA, Donnelly MA, Simpson RL, Simeon GP, Tilson HH. Quality of life bibliography and indexes: 1994 update. *Qual Life Res* 1995;**4**:547–69.

Quality of life research. Oxford: Rapid Science; 1992–98;**1**(1)–**7**(8).

Appendix 2

Details of search strategies

Searches on a range of databases were carried out. The search strategy development followed an iterative process. Initial searches were used to develop the structure of the strategy and to identify key papers which could provide more key terms. Further iterations refined the structure of the search and added key terms. The final collection of papers examined in the review include the results from all iterations of the search strategies. The strategy development took place from March 1999 to June 1999.

A further sequence of updating searches based on the final iteration was undertaken at the end of September 1999. At that stage a few further terms were added to the searches:

Learning adj6 difficult$
Imbecility
Mental adj6 handicap
Disease repercussion profile
Flic
Health assessment questionnaire NOT arthritis

The search process followed a typical pattern of iterative testing. Each iteration produced studies of interest, and further terms to enrich the strategy. Because the searches take many pages the versions for each iteration are only shown for MEDLINE. The searches for the other databases are only presented in their final iteration version – all versions are available from the Centre for Reviews and Dissemination on request.

MEDLINE (searched using the SilverPlatter interface)

First iteration (March 1999):

No.	Records	Request
1	897	explode "Learning-Disorders"/ all subheadings
2	6053	explode "Mental-Retardation"/ all subheadings
3	479	"Mental-Competency"/ all subheadings
4	2339	explode "Brain-Injuries"/ all subheadings
5	5528	explode "Head-Injuries"/ all subheadings
6	12828	#1 or #2 or #3 or #4 or #5
7	1601	explode "Brain-Damage,-Chronic"/ all subheadings
8	17525	explode "Cerebrovascular Disorders"/ all subheadings
9	2304	"Multiple-Sclerosis"/ all subheadings
10	1018	(learning with (disab* or disorder*)) in ti,ab
11	2269	(mental* near (disab* or retard*)) in ti,ab
12	2568	(cognitiv* near (disorder* or disab* or impair*)) in ti,ab
13	358	(intellect* near (disab* or retard* or disorder* or impair*)) in ti,ab
14	3186	(head injur* or brain inju*) in ti,ab
15	1136	(brain damag*) in ti,ab
16	2940	stroke in ti
17	336	(cerebrovascular accident*) in ti,ab
18	37196	#6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #14 or #16 or #17
19	3806	"Activities-of-Daily-Living"/ all subheadings
20	1194	explode "Health-Status-Indicators"/ all subheadings
21	6510	explode "Quality-of-Life"/ all subheadings
22	3387	"Health-Status"
23	134	"Karnofsky-Performance-Status"/ all subheadings
24	56473	(health status or (quality near life) or adl or qol or health state) in ti,ab
25	2157	(short form or sf or mos or disability rating scale or euroqol) in ti,ab
26	235	((qualite near vie) or health profile or qwb or (quality near wellbeing)) in ab,ti
27	62835	#19 or #20 or #21 or #22 or #23 or #24 or #25 or #26
28	96402	(scale or scales or index or measure or measures or measurement) in ti,ab
29	118553	(schedule or assessment or test or tests or questionnaire*) in ti,ab
30	10856	(schedules or instrument or instruments or checklist* or inventory) in ti,ab
31	192762	#28 or #29 or #30

32	8824	#31 near (rating or evaluating or evaluation)
33	8814	#31 near (valid* or precision or rated or rating)
34	10208	#31 near (evaluated or reliability or reliable or predictability or consistency)
35	19604	"Reproducibility-of-Results"
36	9251	"Predictive-Value-of-Tests"
37	44923	#32 or #33 or #34 or #35 or #36
38	428	#18 and #27 and #37

Second iteration (May 1999):

No.	Records	Request
1	854	explode "Learning-Disorders"/ all subheadings
2	5709	explode "Mental-Retardation"/ all subheadings
3	447	"Mental-Competency"/ all subheadings
4	2173	explode "Brain-Injuries"/ all subheadings
5	5161	explode "Head-Injuries"/ all subheadings
6	12046	#1 or #2 or #3 or #4 or #5
7	1508	explode "Brain-Damage,-Chronic"/ all subheadings
8	16404	explode "Cerebrovascular Disorders"/ all subheadings
9	2130	"Multiple-Sclerosis"/ all subheadings
10	962	(learning with (disab* or disorder*)) in ti,ab
11	2140	(mental* near (disab* or retard*)) in ti,ab
12	2388	(cognitiv* near (disorder* or disab* or impair*)) in ti,ab
13	342	(intellect* near (disab* or retard* or disorder* or impair*)) in ti,ab
14	2962	(head injur* or brain inju*) in ti,ab
15	1067	(brain damag*) in ti,ab
16	2745	stroke in ti
17	310	(cerebrovascular accident*) in ti,ab
18	34833	#6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #14 or #16 or #17
19	3595	"Activities-of-Daily-Living"/ all subheadings
20	1080	explode "Health-Status-Indicators"/ all subheadings
21	6077	explode "Quality-of-Life"/ all subheadings
22	3153	"Health-Status"
23	128	"Karnofsky-Performance-Status"/ all subheadings
24	52707	(health status or (quality near life) or adl or qol or health state) in ti,ab
25	1978	(short form or sf or mos or disability rating scale or euroqol) in ti,ab
26	215	((qualite near vie) or health profile or qwb or (quality near wellbeing)) in ab,ti
27	58693	#19 or #20 or #21 or #22 or #23 or #24 or #25 or #26
28	90191	(scale or scales or index or measure or measures or measurement) in ti,ab
29	110979	(schedule or assessment or test or tests or questionnaire*) in ti,ab
30	10176	(schedules or instrument or instruments or checklist* or inventory) in ti,ab
31	180437	#28 or #29 or #30
32	8248	#31 near (rating or evaluating or evaluation)
33	8232	#31 near (valid* or precision or rated or rating)
34	9575	#31 near (evaluated or reliability or reliable or predictability or consistency)
35	18329	"Reproducibility-of-Results"
36	8676	"Predictive-Value-of-Tests"
37	42042	#32 or #33 or #34 or #35 or #36
38	397	#18 and #27 and #37
39	2608	"Feasibility-Studies"
40	46294	explode "Evaluation-Studies"/ all subheadings
41	1182	"Factor-Analysis,-Statistical"
42	114050	exact{COMPARATIVE-STUDY} in TG
43	2845	#31 near (suitable or appropriate)
44	152896	#39 or #40 or #41 or #42
45	160	ql
46	32	coop
47	4059	chart*
48	1563	mcmaster
49	132142	health
50	30671	index
51	1972	nottingham
52	132142	health
53	13627	profile
54	360	ql or coop chart* or mcmaster health index or nottingham health profile or nhp or sf36
55	226	(sickness impact profile or sip) in ti,ab
56	208	"Sickness-Impact-Profile"
57	7143	rating
58	7460	evaluating
59	54745	evaluation
60	51926	(rating or evaluating or evaluation) in ti,ab
61	15999	valid*
62	3415	precision

63	3282	rated
64	7143	rating
65	24414	(valid* or precision or rated or rating) in ti,ab
66	51501	evaluated
67	5350	reliability
68	10583	reliable
69	450	predictability
70	2200	consistency
71	66409	(evaluated or reliability or reliable or predictability or consistency) in ti,ab
72	318208	exact{ANIMAL} in TG
73	778073	exact{human} in tg
74	35224	#6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 or #17
75	58797	#19 or #20 or #21 or #22 or #24 or #25 or #26 or #54 or #55 or #56
76	244050	#35 or #36 or #39 or #40 or #41 or #42 or #43 or #60 or #65 or #71
77	Failed	#74 and #75 and #76
78	2418	#74 and #75
79	952	#78 and #76
80	222615	#72 not (#72 and #73)
81	Failed	#79 not #80
82	555	#79 not #38

Third iteration (June 1999). Searches were transferred to the OVID CD-ROM interface because the size of the strategy was making the SilverPlatter searches unstable:

1 exp learning disorders/
2 exp mental retardation/
3 mental competency/
4 exp brain injuries/
5 exp head injuries/
6 exp brain damage chronic/
7 exp cerebrovascular disorders/
8 (learning adj6 (disab$ or disorder$)).ti,ab.
9 (mental$ adj6 (disab$ or retard$)).ti,ab.
10 (cognitiv$ adj6 (disorder$ or disab$ or impair$)).ti,ab.
11 (intellect$ adj2 (disab$ or retard$ or disorder$ or impair$)).ti,ab.
12 ((head adj injur$) or (brain adj injur$)).ti,ab.
13 (brain adj damag$).ti,ab.
14 stroke.ti.
15 (cerebrovascular adj accident$).ti,ab.
16 or/1-15
17 activities of daily living/
18 exp health status indicators/
19 exp quality of life/
20 health status/
21 ((health adj status) or (quality adj4 life) or adl or qol or (health adj state)).ti,ab.
22 ((short adj form) or sf or mos or (disability adj rating adj scale) or euroqol).ti,ab.
23 ((qualite adj3 vie) or (health adj profile) or qwb or (quality adj3 wellbeing)).ti,ab.
24 (ql or (coop adj chart$) or (mcmaster adj health adj index) or (nottingham adj health adj profile) or nhp or sf36).ti,ab.
25 ((sickness adj impact adj profile) or sip).ti,ab.
26 sickness impact profile/
27 or/17-26
28 reproducibility of results/
29 predictive value of tests/
30 feasibility studies/
31 exp evaluation studies/
32 factor analysis statistical/
33 comparative study/
34 ((scale or scales or index or measure or measures or measurement) adj3 (suitable or appropriate)).ti,ab.
35 ((schedule or assessment or test or tests or questionnaire$) adj3 (suitable or appropriate)).ti,ab.
36 ((schedules or instrument or instruments or checklist$ or inventory) adj3 (suitable or appropriate)).ti,ab.
37 (rating or evaluating or evaluation or valid$ or precision or rated).ti,ab.
38 (evaluated or reliability or reliable or predictability or consistency).ti,ab.
39 or/28-38
40 16 and 27 and 39
41 animal/
42 human/
43 41 not (41 and 42)
44 40 not 43
45 (utility adj scale$).mp.
46 (health adj utility).mp.
47 (utility adj index$).mp.
48 (quality adj adjusted adj life).mp.
49 (qaly$ or qalm$ or qald$ or qale).mp.
50 (quality adj adjusted adj life adj expectancy).mp.
51 (contingent adj valuation).mp.
52 (human adj capital).mp.
53 (global adj health adj index).mp.
54 (global adj health adj indices).mp.
55 ((hrqol$ or hr) adj qol).mp.
56 (health adj life adj equivalent$).mp.
57 hye$.tw.
58 (mutli adj attribute adj scale$).mp
59 or/45-58
60 ms.ti,ab.
61 (multiple adj sclerosis).ti,ab.
62 nhp.ti,ab.
63 or/59,62
64 (60 or 61) and 27 and 39
65 64 not 43
66 65 not 44

67 (16 or 60 or 61) and 62
68 67 not 44
69 68 not 66
70 (multi adj attribute adj scale$).mp.
71 or/59,70
72 (16 or 60 or 61) and 71 and 39
73 72 not (43 or 44)
74 (16 or 60 or 61) and 71
75 74 not (43 or 44)
76 75 not 73

Fourth (and final) iteration (September 1999):

1 exp learning disorders/
2 exp mental retardation/
3 mental competency/
4 exp brain injuries/
5 exp head injuries/
6 exp brain damage chronic/
7 exp cerebrovascular disorders/
8 (learning adj6 (disab$ or disorder$)).ti,ab.
9 (mental$ adj6 (disab$ or retard$)).ti,ab.
10 (cognitiv$ adj6 (disorder$ or disab$ or impair$)).ti,ab.
11 (intellect$ adj2 (disab$ or retard$ or disorder$ or impair$)).ti,ab.
12 ((head adj injur$) or (brain adj injur$)).ti,ab.
13 (brain adj damag$).ti,ab.
14 stroke.ti.
15 (cerebrovascular adj accident$).ti,ab.
16 ms.ti,ab.
17 (multiple adj sclerosis).ti,ab.
18 (learning adj6 difficult$).ti,ab.
19 imbecility.ti,ab.
20 (mental adj6 handicap$).ti,ab.
21 or/1-20
22 activities of daily living/
23 exp health status indicators/
24 exp quality of life/
25 health status/
26 ((health adj status) or (quality adj4 life) or adl or qol or (health adj state)).ti,ab.
27 ((short adj form) or sf or mos or (disability adj rating adj scale) or euroqol).ti,ab.
28 ((qualite adj3 vie) or (health adj profile) or qwb or (quality adj3 wellbeing)).ti,ab.
29 (ql or (coop adj chart$) or (mcmaster adj health adj index) or (nottingham adj health adj profile) or nhp or sf36).ti,ab.
30 ((sickness adj impact adj profile) or sip).ti,ab.
31 sickness impact profile/
32 (utility adj scale$).mp.
33 (health adj utility).mp.
34 (utility adj index$).mp.
35 (quality adj adjusted adj life).mp.
36 (qaly$ or qalm$ or qald$ or qale).mp.
37 (quality adj adjusted adj life adj expectancy).mp.
38 (contingent adj valuation).mp.
39 (human adj capital).mp.
40 (global adj health adj index).mp.
41 (global adj health adj indices).mp.
42 ((hrqol$ or hr) adj qol).mp.
43 (health adj life adj equivalent$).mp.
44 hye$.tw.
45 (multi adj attribute adj scale$).mp.
46 (disease adj repercussion adj profile).ti,ab.
47 (flic or (health adj assessment adj questionnaire)).ti,ab. not arthritis.ti.
48 or/22-47
49 21 and 48
50 animal/
51 human/
52 50 not (50 and 51)
53 49 not 52

CINAHL (searched using the SilverPlatter interface)

Third iteration search (June 1999):

No.	Records	Request
1	262	explode "Learning-Disorders"/ all topical subheadings / all age subheadings
2	463	explode "Mental-Retardation"/ all topical subheadings / all age subheadings
3	693	explode "Brain-Injuries"/ all topical subheadings / all age subheadings
4	1407	#1 or #2 or #3
5	224	explode "Brain-Damage,-Chronic"/ all topical subheadings / all age subheadings
6	1411	"Cerebral-Vascular-Accident"/ all topical subheadings / all age subheadings
7	261	"Multiple-Sclerosis"/ all topical subheadings / all age subheadings
8	1884	#5 or #6 or #7
9	6241	learning
10	7858	disab*
11	12066	disorder*
12	434	(learning with (disab* or disorder*)) in ti,ab
13	244	(mental* near (disab* or retard*)) in ti,ab
14	466	(cognitiv* near (disorder* or disab* or impair*)) in ti,ab
15	68	(intellect* near (disab* or retard* or disorder* or impair*)) in ti,ab
16	3922	#4 or #8 or #12 or #13 or #14 or #15
17	1077	explode "Activities-of-Daily-Living"/ all topical subheadings / all age subheadings

18	1704	explode "Health-Status"/ all topical subheadings / all age subheadings
19	123	"Health-Status-Indicators"/ all topical subheadings / all age subheadings
20	1349	explode "Quality-of-Life"/ all topical subheadings / all age subheadings
21	3828	#17 or #18 or #19 or #20
22	2078	(health status or (quality near life) or adl or qol or health state) in ti,ab
23	173	(short form or sf or mos or disability rating scale or euroqol) in ti,ab
24	46	((qualite near vie) or health profile or qwb or (quality near wellbeing)) in ti,ab
25	4869	#21 or #22 or #23 or #24
26	9692	(scale or scales or index or measure or measures or measurement) in ti,ab
27	12771	(schedule* or assessment or test or tests or questionnaire*) in ti,ab
28	2159	(instrument or instruments or checklist* or inventory) in ti,ab
29	19354	#26 or #27 or #28
30	1179	#29 near (rating or evaluation or evaluating)
31	1463	#29 near (valid* or precision or rated or rating)
32	1240	#29 near (evaluated or reliability or reliable or predictability or consistency)
33	230	"Reproducibility-of-Results"/ all topical subheadings / all age subheadings
34	5046	explode "Reliability-and-Validity"/ all topical subheadings / all age subheadings
35	238	"Predictive-Value-of-Tests"/ all topical subheadings / all age subheadings
36	5170	#33 or #34 or #35
37	6838	#30 or #31 or #32 or #36
38	767	"Instrument-Validation"/ all topical subheadings / all age subheadings
39	138	#16 and #25 and (#37 or #38)
40	47	(ql or coop chart* or mcmaster health index or nottingham health profile or nhp or sf36) in ti,ab
41	27	#31 near (suitable or appropriate) in ti,ab
42	7747	(rating or evaluation or evaluating or valid* or precision or rated) in ti,ab
43	3851	(evaluated or reliability or reliable or predictability or consistency) in ti,ab
44	4873	#25 or #40
45	13238	#36 or #38 or #41 or #42 or #43
46	177	#16 and #44 and #45
47	39	#46 not #39
48	4961	ms
49	279	ms in ti,ab
50	6327	multiple
51	706	sclerosis
52	215	multiple sclerosis in ti,ab
53	405	#49 or #52
54	20	#53 and #44 and #45
55	6	#54 not (#39 or #47)
56	973	utility
57	7536	scale*
58	3	utility scale*
59	2	health utility
60	0	utility index*
61	0	utility indices
62	40	quality adjusted life
63	19	qaly* or qalm* or qald* or qale
64	3	quality adjusted life expectanc*
65	5	contingent valuation
66	26	human capital
67	1	global health ind*
68	15	hrqol* or hr qol
69	0	healthy year equivalen*
70	1	healthy life year*
71	35	hye*
72	0	multi attribute scale*
73	136	#58 or #59 or #60 or #61 or #62 or #63 or #64 or #65 or #66 or #67 or #68 or #69 or #70 or #71 or #72
74	11	(#16 or #53) and #73
75	9	#74 not (#39 or #47 or #55)

Sociological Abstracts (searched using the SilverPlatter interface)

Third iteration (June 1999):

No.	Records	Request
1	681	"Mentally-Retarded" in DE
2	27	"Downs-Syndrome" in DE
3	269	"Learning-Disabilities" in DE
4	0	"brain-injuries" in de
5	291	(learning with (disab* or disorder*)) in ti,ab
6	700	(mental* near (disab* or retard*)) in ti,ab
7	183	(cognitiv* near (disorder* or disab* or impair*)) in ti,ab
8	74	(intellect* near (disab* or retard* or disorder* or impair*)) in ti,ab
9	32	(head injur* or brain injur*) in ti,ab

10	15	(brain damag*) in ti,ab
11	81	stroke in ti,ab
12	49	multiple sclerosis
13	1146	explode "Quality-of-Life"
14	3009	(health status or (quality near life) or adl or qol or health state) in ti,ab
15	68	(short form or sf or sf36 or mos or disability rating scale or euroqol) in ti,ab
16	37	((qualite near vie) or health profile or qwb or (quality near wellbeing)) in ti,ab
17	6	(ql or coop chart* or mcmaster adj health adj index or nottingham health profile or nhp) in ti,ab
18	11	(sickness impact profile or sip) in ti,ab
19	1704	#1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12
20	3447	#13 or #14 or #15 or #16 or #17 or #18
21	88	#19 and #20
22	3392	ms in ti,ab
23	66	#22 and #20
24	63	#23 not #21
25	53	utility scale* or health utility or utility ind* or quality adjusted life
26	6	qaly* or qalm* or qald* or qale
27	2	quality adjusted life adj expectan*
28	13	contingent valuation
29	898	human capital
30	5	global health ind*
31	3	hrqol* or hr qol
32	0	healthy life equivalent*
33	0	healthy year equivalen*
34	34	hye*
35	0	multi attribute scale*
36	1005	#25 or #26 or #27 or #28 or #29 or #30 or #31 or #32 or #33 or #34 or #35
37	36	(#19 or #22) and #36
38	35	#37 not #21
39	98	#24 or #38

EMBASE, CAB HEALTH Dissertation Abstracts Online, JICST-EPlus, Pascal, Conference Papers Index

Searches of these databases were carried out on the Dialog Corporation Dialog Service and were searched in one sweep. The results were de-duplicated against MEDLINE in the same exercise.

Third search iteration (September 1999):

b155,86,163,6,162,35,94,144,77
s Learning Disorders!/de from 155
s Mental Retardation!/de from 155
s Mental Competency/de from 155
s Brain Injuries!/de from 155
s Head Injuries!/de from 155
s Brain-Damage,-Chronic!/de from 155
s Cerebrovascular Disorders!/de from 155
s Multiple Sclerosis!/de
s (learning?(s)(disab? or disorder?))/ti,ab from 155
s (mental?(s)(disab? or retard?))/ti,ab from 155
s (cognitiv?(s)(disorder? or disab? or impair?))/ti,ab from 155
s (intellect?(s)(disab? or retard? or disorder? or impair?))/ti,ab from 155
s (head()injur? or brain()inju?)/ti,ab from 155
s (brain()damag?)/ti,ab from 155
s stroke/ti from 155
s (cerebrovascular()accident?)/ti,ab from 155
s (learning?(s)(disab? or disorder?)) from 86,163,6,162,35,94,144,77
s (mental?(s)(disab? or retard?)) from 86,163,6,162,35,94,144,77
s (cognitiv?(s)(disorder? or disab? or impair?)) from 86,163,6,162,35,94,144,77
s (intellect?(s)(disab? or retard? or disorder? or impair?)) from 86,163,6,162,35,94,144,77
s (head()injur? or brain()inju?) from 86,163,6,162,35,94,144,77
s (brain()damag?) from 86,163,6,162,35,94,144,77
s stroke from 86,163,6,162,35,94,144,77
s (cerebrovascular()accident?) from 86,163,6,162,35,94,144,77
s ms/ti,ab
s multiple()sclerosis/ti
s (learning(6w)difficult?)/ti,ab
s imbecility/ti,ab
s (mental(6w)handicap?)/ti,ab
ss s1:s29
s Activities of Daily Living/de from 155
s Health Status Indicators!/de from 155
s Quality of Life!/de from 155
s Health Status/de from 155
s Karnofsky Performance Status/de from 155
s (health()status or (quality(3w)life) or adl or qol or health()state? ?)/ti,ab from 155
s (short()form or sf or mos or disability()rating()scale or euroqol)/ti,ab from 155
s ((qualite (3w) vie) or health()profile or qwb or (quality(3w)wellbeing))/ab,ti from 155
s (health()status or (quality(3w)life) or adl or qol or health()state? ?) from 86,163,6,162,35,94,144,77
s (short()form or sf or mos or disability()rating()scale or euroqol)from 86,163,6,162,35,94,144,77
s ((qualite (3w) vie) or health()profile or qwb or (quality(3w)wellbeing))from 86,163,6,162,35,94,144,77
s (ql or coop()chart? Or mcmaster()health()index

or nottingham()health()profile or nhp or sf36)/ti,ab
s (sickness()impact()profile or sip)/ti,ab
s (nhp or utility()scale? ? or health()utility)/ti,ab
s (utility()index? or quality()adjusted()life)/ti,ab
s (qaly? or qalm? or qald? or qale)/ti,ab
s (quality()adjusted()life()expectancy)/ti,ab
s (contingent()valuation or human()capital)/ti,ab
s (global()health()index or global()health()indices)/ti,ab
s (hrqol? or hr()qol)/ti,ab
s (healthy()life()equivalent?)/ti,ab
s (healthy()year()equivalen?)/ti,ab
s hye?/ti,ab
s (multi()attribute()scale?)/ti,ab
s (disease()management()profile)/ti,ab
s (FLIC or (health()assessment()questionnaire))/ti,ab not arthritis/ti
ss s31:s56
s s30 and s57

PsycLIT (searched using the SilverPlatter interface)

Third search iteration (June 1999):

No.	Records	Request
1	4514	explode "Mental-Retardation"
2	4983	explode "Learning-Disabilities"
3	8306	explode "Mentally-Retarded"
4	723	"Learning-Disorders" in DE
5	508	"Cerebrovascular-Accidents" in DE
6	49	"Cerebral-Hemorrhage" in DE
7	60	"Cerebral-Ischemia" in DE
8	235	"Multiple-Sclerosis" in DE
9	2417	explode "Brain-Damage"
10	1421	explode "Brain-Damaged"
11	600	explode "Head-Injuries"
12	22580	#1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11
13	87658	learning
14	19796	disab*
15	79579	disorder*
16	5505	(learning near (disab* or disorder*)) in ti,ab
17	11613	(mental* near (disab* or retard*)) in ti,ab
18	3110	(cognitiv* near (disorder* or disab* or impair*)) in ti,ab
19	2377	(intellect* near (disab* or retard* or disorder* or impair*)) in ti,ab
20	1888	(head injur* or brain injur*) in ti,ab
21	3300	brain damage* in ti,ab
22	228	stroke in ti
23	111	cerebrovascular accident* in ti,ab
24	505	multiple sclerosis in ti,ab
25	34210	#12 or #16 or #17 or #18 or #19 or #20 or #21 or #22 or #23 or #24
26	335	explode "Quality-of-Life"
27	1	"Activities-of-Daily-Living" in DE
28	335	#26 or #27
29	2107	(health status or (quality near life) or adl or qol or health state) in ti,ab
30	1161	(short form or sf or mos or disability rating scale or euroqol) in ti,ab
31	15	((qualite near vie) or health profile or qwb or (quality near wellbeing)) in ti,ab
32	3366	#28 or #29 or #30 or #31
33	114957	(scale or scales or index or measure or measures or measurement) in ti,ab
34	157256	(schedule or assessment or test or tests or questionnaire*) in ti,ab
35	33521	(schedules or instrument or instruments or checklist* or inventory) in ti,ab
36	239451	#33 or #34 or #35
37	19921	#36 near (rating or evaluation or evaluation)
38	35216	#36 near (valid* or precision or rated or rating)
39	21798	#36 near (evaluated or reliability or reliable or predictability or consistency)
40	3670	explode "Statistical-Validity"
41	704	"Construct-Validity" in DE
42	6567	"Test-Validity" in DE
43	395	"Test-Standardization" in DE
44	3913	"Test-Reliability" in DE
45	470	"Interrater-Reliability" in DE
46	1554	"Statistical-Reliability" in DE
47	0	"Consistency-(Measurement)" in DE
48	163	explode "Prediction-Errors"
49	13565	#40 or #41 or #42 or #43 or #44 or #45 or #46 or #47 or #48
50	59534	#37 or #38 or #39 or #49
51	71	#25 and #32 and #50
52	15974	rating
53	18973	scales
54	2461	rating scales in de
55	3	ql
56	0	sf36
57	26	coop
58	2453	chart*
59	38	mcmaster
60	40719	health
61	12466	index
62	41	nottingham
63	40719	health
64	5199	profile

65	2	nhp
66	7	(ql or sf36 or coop chart* or mcmaster health index or nottingham health profile or nhp) in ti,ab
67	1029	sickness
68	11505	impact
69	5199	profile
70	71	sip
71	84	(sickness impact profile or sip) in ti,ab
72	3548	explode "Factor-Analysis"
73	3483	suitable
74	14217	appropriate
75	5856	#36 near (suitable or appropriate)
76	15974	rating
77	5938	evaluating
78	40720	evaluation
79	31573	valid*
80	1632	precision
81	13533	rated
82	79155	(rating or evaluating or evaluation or valid* or precision or rated) in ti,ab
83	14376	evaluated
84	14642	reliability
85	7452	reliable
86	1021	predictability
87	5836	consistency
88	37733	(evaluated or reliability or reliable or predictability or consistency) in ti,ab
89	5863	#32 or #54 or #66 or #71
90	93393	#49 or #72 or #75 or #82 or #89
91	394	#25 and #89 and #90
92	323	#91 not #51
93	45	(utility scale* or health utility or utility ind* or quality adjusted life) in ti,ab
94	0	(qaly* or qalm* or qald* or qale or quality adjusted life expectan*) in ti,ab
95	30	(contingent valuation or human capital or global health ind*) in ti,ab
96	0	(hrqol* or hr qol or healthy life equivalen* or healthy year equivalen*) in ti,ab
97	40	(hye* or multi attribute scale*) in ti,ab
98	114	#93 or #94 or #95 or #96 or #97
99	1	#25 and #98
100	1	#99 not (#51 or #92)

Linguistics and Language Behavior Abstracts (searched using the SilverPlatter interface)

Second iteration (June 1999):

No.	Records	Request
1	260	"Nervous-System-Disorders" in DE
2	575	"Brain-Damage" in DE
3	1171	explode "Mental-Retardation"
4	1973	#1 or #2 or #3
5	338	stroke
8	371	stroke or cerebrovascular accident*
12	35	multiple sclerosis
17	1077	brain injur* or brain damage*
23	3428	learning near (disabl* or disabili* or disorder* or impair*)
29	1441	mental* near (disab* or disabil* or retard* or impair*)
34	378	cognitiv* near (disorder* or disab* or impair*)
40	146	intellect* near (disab* or retard* or disorder* or impair*)
41	6443	#4 or #8 or #12 or #17 or #23 or #29 or #34 or #40
50	69	health status or (quality near life) or adl or qol or health state
59	1355	short form or sf or mos or disability rating scale or euroqol
67	1	(qualite near vie) or health profile or qwb or (quality near wellbeing)
70	20	activities near living
71	1444	#50 or #59 or #67 or #70
77	26648	schedule or assessment or test or tests or questionnaire*
83	4454	schedules or instrument or instruments or checklist* or inventory
91	13740	scale or scales or index or measure or measures or measurement
92	38165	#77 or #83 or #91
97	2579	#92 near (rating* or evaluation or evaluating or evaluations)
102	2817	#92 near (valid* or precision or rated or rating)
108	2024	#92 near (evaluated or reliability or reliable or predictability or consistency)
109	5230	#97 or #102 or #108
110	648	"Test-Validity-and-Reliability" in DE
111	0	"Measures-(Instruments)" in DE
112	78	"Rating-Scales" in DE
113	709	#110 or #111 or #112
114	5	#71 and #113
115	6631	#71 or #109
116	362	#41 and #115
124	327	((schedule or assessment or test or tests or questionnaire) near (suitable or appropriate)) in ti,ab
132	60	((schedules or instrument or instruments or checklist* or inventory) near (suitable or appropriate)) in ti,ab
142	1	(utility scale* or health utility or utility index or quality adjusted life) in ti,ab

149	0	(qaly* or qalm* or qald* or qale or contingent valuation) in ti,ab
161	8	(human capital or global health ind* or hrqol* or hr qol or healthy life equivalen*) in ti,ab
169	7	(healthy year equivalen* or hye* or multi attribute scale*) in ti,ab
178	118	((scale or scales or index or measure or measures or measurement) near (suitable or appropriate)) in ti,ab
179	485	#124 or #132 or #142 or #149 or #161 or #169 or #178
180	29	#41 and #179
181	21	#180 not #116

Mental Health Abstracts, Ageline, NTIS

The following databases were searched using the Dialog Corporation Dialog Service in one sweep and records were also de-duplicated during the same process.

Third search iteration (June 1999):

Set	Items	Description
1	9593	LEARNING DISORDERS!/DE FROM 155
2	59020	MENTAL RETARDATION!/DE FROM 155
3	1150	MENTAL COMPETENCY/DE FROM 155
4	18139	BRAIN INJURIES!/DE FROM 155
5	52420	HEAD INJURIES!/DE FROM 155
6	0	BRAIN-DAMAGE,-CHRONIC!/DE FROM 155
7	120430	CEREBROVASCULAR DISORDERS!/DE FROM 155
8	26076	MULTIPLE SCLEROSIS!/DE
9	5197	(LEARNING?(S)(DISAB? OR DISORDER?))/TI,AB FROM 155
10	17551	(MENTAL?(S)(DISAB? OR RETARD?))/TI,AB FROM 155
11	12659	(COGNITIV?(S)(DISORDER? OR DISAB? OR IMPAIR?))/TI,AB FROM 155
12	2731	(INTELLECT?(S)(DISAB? OR RETARD? OR DISORDER? OR IMPAIR?))/TI,AB FROM 155
13	14875	(HEAD()INJUR? OR BRAIN()INJU?)/TI,AB FROM 155
14	7043	(BRAIN()DAMAG?)/TI,AB FROM 155
15	12453	STROKE/TI FROM 155
16	2203	(CEREBROVASCULAR() ACCIDENT?)/TI,AB FROM 155
17	15659	(LEARNING?(S)(DISAB? OR DISORDER?)) FROM 86,163,6,162,35,94,144,77
18	52009	(MENTAL?(S)(DISAB? OR RETARD?)) FROM 86,163,6,162,35,94,144,77
19	20726	(COGNITIV?(S)(DISORDER? OR DISAB? OR IMPAIR?)) FROM 86,163,6,162,35,94,144,77
20	4974	(INTELLECT?(S)(DISAB? OR RETARD? OR DISORDER? OR IMPAIR?)) FROM 86,163,6,162,35,94,144,77
21	15872	(HEAD()INJUR? OR BRAIN()INJU?) FROM 86,163,6,162,35,94,144,77
22	8179	(BRAIN()DAMAG?) FROM 86,163,6,162,35,94,144,77
23	31943	STROKE FROM 86,163,6,162,35,94,144,77
24	1149	(CEREBROVASCULAR() ACCIDENT?) FROM 86,163,6,162,35,94,144,77
25	424320	S1:S24
26	16650	ACTIVITIES OF DAILY LIVING/DE FROM 155
27	4178	HEALTH STATUS INDICATORS!/DE FROM 155
28	19256	QUALITY OF LIFE!/DE FROM 155
29	12435	HEALTH STATUS/DE FROM 155
30	258	KARNOFSKY PERFORMANCE STATUS/DE FROM 155
31	28861	(HEALTH()STATUS OR (QUALITY(3W)LIFE) OR ADL OR QOL OR HEALTH()STATE? ?)/TI,AB FROM 155
32	7182	(SHORT()FORM OR SF OR MOS OR DISABILITY()RATING()SCALE OR EUROQOL)/TI,AB FROM 155
33	736	((QUALITE (3W) VIE) OR HEALTH()PROFILE OR QWB OR (QUALITY(3W)WELLBEING))/AB,TI FROM 155
34	42325	(HEALTH()STATUS OR (QUALITY(3W)LIFE) OR ADL OR QOL OR HEALTH()STATE? ?) FROM 86,163,6,162,35,94,144,77
35	45960	(SHORT()FORM OR SF OR MOS OR DISABILITY()RATING()SCALE OR EUROQOL)FROM 86,163,6,162,35,94,144,77
36	8454	((QUALITE (3W) VIE) OR HEALTH()PROFILE OR QWB OR (QUALITY(3W)WELLBEING)) FROM 86,163,6,162,35,94,144,77
37	156362	S26:S36
38	1635915	(SCALE OR SCALES OR INDEX OR MEASURE OR MEASURES OR MEASUREMENT)/TI,AB
39	1949861	(SCHEDULE OR ASSESSMENT

		OR TEST OR TESTS OR QUESTIONNAIRE?)/TI,AB
40	284572	(SCHEDULES OR INSTRUMENT OR INSTRUMENTS OR CHECKLIST? OR INVENTORY)/TI,AB
41	3374758	S38:S40
42	263263	S41(S)(RATING OR EVALUATING OR EVALUATION)
43	201835	S41(S) (VALID? OR PRECISION OR RATED OR RATING)
44	271412	S41(S) (EVALUATED OR RELIABILITY OR RELIABLE OR PREDICTABILITY OR CONSISTENCY)
45	280040	(REPRODUCIBILITY OF RESULTS/DE OR FEASIBILITY STUDIES/DE OR EVALUATION STUDIES!/DE) FROM 155
46	27658	(PREDICTIVE VALUE OF TESTS/DE OR FACTOR ANALYSIS/DE OR COMPARATIVE STUDY/DE) FROM 155
47	870215	S42:S46
48	1899	S25 AND S37 AND S47
49	93364	MS/TI,AB
50	19452	MULTIPLE()SCLEROSIS/TI
51	1853	(QL OR COOP()CHART? OR MCMASTER()HEALTH()INDEX OR NOTTINGHAM()HEALTH() PROFILE OR NHP OR SF36)/TI,AB
52	2242	(SICKNESS()IMPACT()PROFILE OR SIP)/TI,AB
53	847	(NHP OR UTILITY()SCALE? ? OR HEALTH()UTILITY)/TI,AB
54	780	(UTILITY()INDEX? OR QUALITY() ADJUSTED()LIFE)/TI,AB
55	564	(QALY? OR QALM? OR QALD? OR QALE)/TI,AB
56	139	(QUALITY()ADJUSTED()LIFE() EXPECTANCY)/TI,AB
57	2647	(CONTINGENT()VALUATION OR HUMAN()CAPITAL)/TI,AB
58	5	(GLOBAL()HEALTH()INDEX OR GLOBAL()HEALTH() INDICES)/TI,AB
59	299	(HRQOL? OR HR()QOL)/TI,AB
60	0	(HEALTHY()LIFE() EQUIVALENT?)/TI,AB
61	2	(HEALTHY()YEAR() EQUIVALEN?)/TI,AB
62	1051	HYE?/TI,AB
63	1	(MULTI()ATTRIBUTE() SCALE?)/TI,AB
64	114322	S41(S)(SUITABLE OR APPROPRIATE)
65	1467302	(RATING OR EVALUATING OR EVALUATION OR VALID? OR PRECISION OR RATED)/TI,AB
66	812922	(EVALUATED OR RELIABILITY OR RELIABLE OR PREDICTABILITY OR CONSISTENCY)/TI,AB
67	511471	S25 OR S49 OR S50
68	162762	S37 OR S51:S63
69	2386999	S45:S46 OR S64:S66
70	2656	S67 AND S68 AND S69
71	757	S70 NOT S48
72	707	RD S71 (unique items)

Health and Psychosocial Instruments (HAPI) (Ovid search service)

March 1999:

Set	Documents	Search
1	0	(mental adj reatrd).mp.
2	292	(mental adj retard$).mp.
3	270	(mentally adj retard$).mp.
4	285	(learning adj disab$).mp.
5	257	(learning adj disorder$).mp.
6	14	(mental$ adj disab$).mp.
7	88	(cognitiv$ adj (disord$ or disab$ or impair$)).mp.
8	7	(intellect$ adj (impair$ or disab$ or retard$)).mp.
9	51	stroke.mp.
10	215	(cerebrovascular adj accident$).mp.
11	167	(multiple adj sclerosis).mp.
12	247	(brain adj (damag$ or injur$)).mp.
13	0	head ajd injur$.mp.
14	210	(head adj injur$).mp.
15	0	2 3 4 5 6 7 8 9 110 11.mp.
16	749	2 or 3 or 4 or 5 or 6 or 7 or 8
17	1543	16 or 9 or 10 or 11 or 12 or 14
18	0	(internal adj coconisstency).mp.
19	3223	internal consistency y.ry.
20	240	parallel forms y.ry.
21	1797	test-retest y.ry.
22	879	inter-rater y.ry.
23	2248	content y.va.
24	1697	criterion y.va.
25	0	consturct y.va.
26	0	co s.mp.
27	2814	construct y.va.
28	4981	19 or 20 or 21 or 22 or 23 or 24 or 27
29	114	17 and 28

Econlit (searched using the SilverPlatter interface)

First search iteration (June 1999):

No.	Records	Request
1	17	stroke
2	0	cerebrovascular

3	1184	accident*
4	0	cerebrovascular accident*
5	3985	learning
6	203	disorder*
7	0	learning disorder*
8	3985	learning
9	2655	disabilit*
10	0	learning disabilit*
11	403	mental*
12	2812	disab*
13	6	mental* disab*
14	403	mental*
15	203	disorder*
16	7	mental* disorder*
17	403	mental*
18	180	retard*
19	10	mental* retard*
20	492	cognitiv*
21	164	impair*
22	4	cognitiv* impair*
23	492	cognitiv*
24	203	disorder*
25	0	cognitiv* disorder*
26	492	cognitiv*
27	253	disable*
28	0	cognitiv* disable*
29	492	cognitiv*
30	2812	disab*
31	0	cognitiv* disab*
32	1561	intellect*
33	2812	disab*
34	1	intellect* disab*
35	1561	intellect*
36	164	impair*
37	0	intellect* impair*
38	1561	intellect*
39	206	disord*
40	0	intellect* disord*
41	1561	intellect*
42	180	retard*
43	0	intellect* retard*
44	591	head
45	617	injur*
46	3	head injur*
47	135	brain
48	617	injur*
49	0	brain injur*
50	135	brain
51	1118	damag*
52	1	brain damag*
53	403	mental*
54	492	competen*
55	0	mental* competen*
56	4793	multiple
57	20	sclerosis
58	6	multiple sclerosis
59	490	ms
60	62	ms in ti,ab
61	110	#1 or #4 or #7 or #10 or #13 or #16 or #19 or #22 or #25 or #28 or #31 or #34 or #37 or #40 or #43 or #46 or #49 or #52 or #55 or #60
62	18322	health
63	3125	status
64	212	health status
65	7058	quality
66	11992	life
67	1084	quality with life
68	8	adl
69	4	qol
70	18322	health
71	35138	state
72	48	adl or qol or health state
73	8351	short
74	6824	form
75	13	sf
76	6	mos
77	372	rating
78	5535	scale
79	1	euroqol
80	27	short form or sf or mos or rating scale or euroqol
81	34	qualite
82	18322	health
83	760	profile
84	0	qwb
85	7058	quality
86	23	wellbeing
87	37	qualite or health profile or qwb or (quality with wellbeing)
88	1	ql
89	14	coop
90	637	chart*
91	18322	health
92	20610	index
93	18322	health
94	760	profile
95	1	nhp
96	0	sf36
97	75	sickness
98	18561	impact
99	7	sip
100	15	ql or coop chart* or health index or health profile or nhp or sf36 or sickness impact or sip
101	5811	utility
102	5846	scale*
103	3	utility scale*
104	18322	health
105	5811	utility
106	1	health utility
107	7058	quality
108	1203	adjusted
109	40	qaly*

110	108	quality adjusted or qaly*
111	0	qalm*
112	0	qald*
113	222	dalY*
114	2583	disability
115	1203	adjusted
116	3752	contingent
117	2902	valuation
118	904	qalm* or qald* or dalY* or disability adjusted or contingent valuation
119	16770	human
120	59974	capital
121	18322	health
122	20610	index
123	179	healthy
124	11992	life
125	179	healthy
126	11992	life
127	48	hye*
128	2	multi
129	366	attribute
130	8010	human capital or health index or healthy life or healthy life or hye* or multi attribute
131	10269	#64 or #67 or #72 or #80 or #87 or #100 or #103 or #106 or #110 or #118 or #130
132	8	#61 and #131

EMBASE (searched using the SilverPlatter interface)

Third iteration (July 1999):

No.	Records	Request
1	202	explode "learning-disorder"/ all subheadings
2	1199	explode "mental-deficiency"/ all subheadings
3	762	"cognitive-defect"/ all subheadings
4	1145	explode "brain-injury"/ all subheadings
5	418	"head-injury"/ all subheadings
6	418	"head-injury"/ all subheadings
7	418	"head-injury"/ all subheadings
8	1511	"stroke"/ all subheadings
9	768	"multiple-sclerosis"/ all subheadings
10	5656	#1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9
11	2035	learning
12	2453	disab*
13	2035	learning
14	18564	disorder*
15	107	(learning disab* or learning disorder*) in ti,ab
16	441	(mental* disab* or mental* retard*) in ti,ab
17	408	(cognitive* disab* or cognitiv* disorder* or cognitiv* impair*) in ti,ab
18	89	(intellect* disab* or intellect* retard* or intellect* disord* or intellect* impair*) in ti,ab
19	755	(head injur* or brain injur*) in ti,ab
20	272	brain damag* in ti,ab
21	751	stroke in ti
22	80	cerebrovascular accident* in ti,ab
23	2777	#15 or #16 or #17 or #18 or #19 or #20 or #21 or #22
24	1177	explode "daily-life-activity"/ all subheadings
25	963	explode "health-status"/ all subheadings
26	2150	explode "quality-of-life"/ all subheadings
27	4005	#24 or #25 or #26
28	2272	(health status or (quality near life) or adl or qol or health state*) in ti,ab
29	498	(short form or sf or mos or disability rating scale or euroqol) in ti,ab
30	65	((qualite near vie) or health profile or qwb or (quality near wellbeing)) in ti,ab
31	69	(nhp or ql or coop chart* or mcmaster health index or nottingham health profile) in ti,ab
32	68	(sickness impact profile or sf36 or sip or utility()scale*) in ti,ab
33	54	(health utility or utility index* or quality adjusted life) in ti,ab
34	40	(qaly* or qalm* or qald* or qale or quality adjusted life expect*) in ti,ab
35	18	(contingent valuation or human capital or global health index or global health indices) in ti,ab
36	36	(hrqol or hr qol or healthy life equivalent* or healthy year equivalen*) in ti,ab
37	17	(hye* or multi attribute scale*) in ti,ab
38	5045	#27 or #28 or #29 or #30 or #31 or #32 or #33 or #34 or #35 or #36 or #37
39	253	#10 and #38
40	901	"reproducibility"/ all subheadings
41	15	"prediction-and-forecasting"/ all subheadings
42	2368	"prediction"/ all subheadings
43	14050	explode "evaluation-and-follow up"/ all subheadings

44	1074	"reliability"/ all subheadings
45	17641	#40 or #41 or #42 or #43 or #44
46	302	((scale or scales or index or measure or measures or measurement) near (suitable or appropriate)) in ti,ab
47	406	((schedule or assessment or test or tests or questionnaire) near (suitable or appropriate)) in ti,ab
48	67	((schedules or instrument* or checklist* or inventory) near (suitable or appropriate)) in ti,ab
49	17272	(rating or evaluating or evaluation or valid* or precision or rated) in ti,ab
50	16309	(evaluated or reliability or reliable or predictability or consistency) in ti,ab
51	42609	#45 or #46 or #47 or #48 or #49 or #50
52	111	#39 and #51

Social Scisearch

This database was searched using the Dialog Corporation Dialog service and records were de-duplicated against MEDLINE during the search.

Second iteration (Sept. 1999):

s Learning Disorders!/de from 155
s Mental Retardation!/de from 155
s Mental Competency/de from 155
s Brain Injuries!/de from 155
s Head Injuries!/de from 155
s Brain-Damage,-Chronic!/de from 155
s Cerebrovascular Disorders!/de from 155
s Multiple Sclerosis!/de
s (learning?(s)(disab? or disorder?))/ti,ab from 155
s (mental?(s)(disab? or retard?))/ti,ab from 155
s (cognitiv?(s)(disorder? or disab? or impair?))/ti,ab from 155
s (intellect?(s)(disab? or retard? or disorder? or impair?))/ti,ab from 155
s (head()injur? or brain()inju?)/ti,ab from 155
s (brain()damag?)/ti,ab from 155
s stroke/ti from 155
s (cerebrovascular()accident?)/ti,ab from 155
s (learning?(s)(disab? or disorder?)) from 7,434
s (mental?(s)(disab? or retard?)) from 7,434
s (cognitiv?(s)(disorder? or disab? or impair?)) from 7,434
s (intellect?(s)(disab? or retard? or disorder? or impair?)) from 7,434
s (head()injur? or brain()inju?) from 7,434
s (brain()damag?) from 7,434
s stroke from 7,434
s (cerebrovascular()accident?) from 7,434
s (learning(6w)difficult?)/ti,ab
s imbecility/ti,ab
s (mental(6w)handicap?)/ti,ab
s ms/ti,ab
s multiple()sclerosis/ti
s s1:s29
s Activities of Daily Living/de from 155
s Health Status Indicators!/de from 155
s Quality of Life!/de from 155
s Health Status/de from 155
s Karnofsky Performance Status/de from 155
s (health()status or (quality(3w)life) or adl or qol or health()state? ?)/ti,ab from 155
s (short()form or sf or mos or disability()rating()scale or euroqol)/ti,ab from 155
s ((qualite (3w) vie) or health()profile or qwb or (quality(3w)wellbeing))/ab,ti from 155
s (health()status or (quality(3w)life) or adl or qol or health()state? ?) from 7,434
s (short()form or sf or mos or disability()rating()scale or euroqol)from 7,434
s ((qualite (3w) vie) or health()profile or qwb or (quality(3w)wellbeing))from 7,434
s (ql or coop()chart? Or mcmaster()health()index or nottingham()health()profile or nhp or sf36)/ti,ab
s (sickness()impact()profile or sip)/ti,ab
s (nhp or utility()scale? ? or health()utility)/ti,ab
s (utility()index? or quality()adjusted()life)/ti,ab
s (qaly? or qalm? or qald? or qale)/ti,ab
s (quality()adjusted()life()expectancy)/ti,ab
s (contingent()valuation or human()capital)/ti,ab
s (global()health()index or global()health()indices)/ti,ab
s (hrqol? or hr()qol)/ti,ab
s (healthy()life()equivalent?)/ti,ab
s (healthy()year()equivalen?)/ti,ab
s hye?/ti,ab
s (multi()attribute()scale?)/ti,ab
s (disease()management()profile)/ti,ab
s (FLIC or (health()assessment()questionnaire))/ti,ab not arthritis/ti
ss s31:s56
s s30 and s57

HealthSTAR

This was searched using the Dialog Corporation Dialog service and records were de-duplicated against MEDLINE during the search.

June 1999:

Set	Items	Description
1	14097	LEARNING DISORDERS!/DE
2	37654	MENTAL DEFICIENCY!/DE FROM 73
3	6754	COGNITIVE DEFECT/DE FROM 73

4	29166	BRAIN INJURY!/DE FROM 73
5	13371	HEAD INJURY/DE FROM 73
6	79529	MENTAL RETARDATION!/DE FROM 155,151
7	2441	MENTAL COMPETENCY/DE FROM 155,151
8	25450	BRAIN INJURIES!/DE FROM 155,151
9	73595	HEAD INJURIES!/DE FROM 155,151
10	0	BRAIN-DAMAGE,-CHRONIC!/DE FROM 155,151
11	171142	CEREBROVASCULAR DISORDERS!/DE FROM 155,151
12	37149	MULTIPLE SCLEROSIS!/DE
13	27295	STROKE/DE FROM 73
14	12205	(LEARNING?(S)(DISAB? OR DISORDER?))/TI,AB
15	42800	(MENTAL?(S)(DISAB? OR RETARD?))/TI,AB
16	34361	(COGNITIV?(S)(DISORDER? OR DISAB? OR IMPAIR?))/TI,AB
17	7648	(INTELLECT?(S)(DISAB? OR RETARD? OR DISORDER? OR IMPAIR?))/TI,AB
18	37001	(HEAD()INJUR? OR BRAIN()INJU?)/TI,AB
19	16462	(BRAIN()DAMAG?)/TI,AB
20	56649	STROKE/TI OR MULTIPLE()SCLEROSIS/TI
21	5769	(CEREBROVASCULAR() ACCIDENT?)/TI,AB
22	542589	S1:S21
23	33280	ACTIVITIES OF DAILY LIVING/DE FROM 155,151
24	9129	HEALTH STATUS INDICATORS!/DE FROM 155,151
25	57615	QUALITY OF LIFE!/DE
26	25826	HEALTH STATUS/DE FROM 155,151
27	518	KARNOFSKY PERFORMANCE STATUS/DE FROM 155,151
28	82766	(HEALTH()STATUS OR (QUALITY(3W)LIFE) OR ADL OR QOL OR HEALTH()STATE? ?)/TI,AB
29	16623	(SHORT()FORM OR SF OR MOS OR DISABILITY()RATING()SCALE OR EUROQOL)/TI,AB
30	2226	((QUALITE (3W) VIE) OR HEALTH()PROFILE OR QWB OR (QUALITY(3W)WELLBEING))/AB,TI
31	135405	(HEALTH()STATUS OR (QUALITY(3W)LIFE) OR ADL OR QOL OR HEALTH()STATE? ?)
32	17075	(SHORT()FORM OR SF OR MOS OR DISABILITY()RATING()SCALE OR EUROQOL)
33	2381	((QUALITE (3W) VIE) OR HEALTH()PROFILE OR QWB OR (QUALITY(3W)WELLBEING))
34	17853	DAILY LIFE ACTIVITY!/DE FROM 73
35	14066	HEALTH STATUS!/DE FROM 73
36	195732	S23:S35
37	1167488	(SCALE OR SCALES OR INDEX OR MEASURE OR MEASURES OR MEASUREMENT)/TI,AB
38	1552056	(SCHEDULE OR ASSESSMENT OR TEST OR TESTS OR QUESTIONNAIRE?)/TI,AB
39	152136	(SCHEDULES OR INSTRUMENT OR INSTRUMENTS OR CHECKLIST OR INVENTORY)/TI,AB
40	2518480	S37:S39
41	226009	S40(S)(RATING OR EVALUATING OR EVALUATION)
42	163412	S40(S) (VALID? OR PRECISION OR RATED OR RATING)
43	278208	S40(S) (EVALUATED OR RELIABILITY OR RELIABLE OR PREDICTABILITY OR CONSISTENCY)
44	530120	(REPRODUCIBILITY OF RESULTS/DE OR FEASIBILITY STUDIES/DE OR EVALUATION STUDIES!/DE) FROM 155,151
45	54364	(PREDICTIVE VALUE OF TESTS/DE OR FACTOR ANALYSIS/DE OR COMPARATIVE STUDY/DE) FROM 155,151
46	7357	REPRODUCIBILITY/DE FROM 73
47	549	PREDICTION AND FORECASTING/DE FROM 73
48	94978	"EVALUATION AND FOLLOW UP"!/DE FROM 73
49	9595	RELIABILITY/DE FROM 73
50	1136932	S41:S49
51	3911	S22 AND S36 AND S50
52	2569	(QL OR COOP()CHART? OR MCMASTER()HEALTH()INDEX OR NOTTINGHAM()HEALTH()PROFILE OR NHP OR SF36)/TI,AB
53	2112	(SICKNESS()IMPACT()PROFILE OR SIP)/TI,AB
54	1080	(NHP OR UTILITY()SCALE? ? OR HEALTH()UTILITY)/TI,AB
55	1595	(UTILITY()INDEX? OR QUALITY()ADJUSTED()LIFE)/TI,AB
56	1193	(QALY? OR QALM? OR QALD? OR QALE)/TI,AB

57	261	(QUALITY()ADJUSTED()LIFE()EXPECTANCY)/TI,AB
58	494	(CONTINGENT()VALUATION OR HUMAN()CAPITAL)/TI,AB
59	11	(GLOBAL()HEALTH()INDEX OR GLOBAL()HEALTH()INDICES)/TI,AB
60	617	(HRQOL? OR HR()QOL)/TI,AB
61	0	(HEALTHY()LIFE()EQUIVALENT?)/TI,AB
62	10	(HEALTHY()YEAR()EQUIVALEN?)/TI,AB
63	605	HYE?/TI,AB
64	3	(MULTI()ATTRIBUTE()SCALE?)/TI,AB
65	14506	S41(S)(SUITABLE OR APPROPRIATE)
66	1044116	(RATING OR EVALUATING OR EVALUATION OR VALID? OR PRECISION OR RATED)/TI,AB
67	804495	(EVALUATED OR RELIABILITY OR RELIABLE OR PREDICTABILITY OR CONSISTENCY)/TI,AB
68	198532	S36 OR S52:S64
69	2070407	S44:S45 OR S65:S67
70	4530	S68 AND S69 AND S22
71	4868	S51 OR S70
72	108233	S40(S)(SUITABLE OR APPROPRIATE)
73	2140102	S44:S45 OR S64:S67 OR S72
74	4724	S68 AND S22 AND S73
75	5055	S51 OR S74
76	199	S75/1900-1980
77	124	RD S76 (unique items)
78	3545	S75 FROM 155,151
79	1833	RD S78 (unique items)

Appendix 3

Data extraction form

ID | Ref ID | Study | Country

QOL scale

Method of administration

Time of administration

Scale components

Disease status | Cognitively impaired included

Level of cognitive impairment

Population characteristics

Study design

Sampling procedure

Aim(s) of study

Number of participants

Number of participants in final analysis

Other details

Validation results

Conclusion about QoL scale

Overall study conclusions

Comments

Appendix 4

Details of studies included in the review

TABLE 2

Study	Country	QOL measure	Method of administration	Disease status	Study design	No. of participants	Cognitive impairment included?
Ahlsio *et al.*, 1984[55]	Sweden	VAS-QoL	Interview by means of structured questionnaire	Stroke	Cohort	96	Probable
Anderson *et al.*, 1996[15]	Australia	SF-36	Face-to-face interview	Stroke	Cohort	90	Probable
Baird *et al.*, 1987[31]	USA	SIP	Not stated	Stroke	Cross-sectional	95	Yes
Burton & Volpe 1994[32]	USA	SIP	Mailed questionnaire	Brain injury	Cross-sectional	19	Yes
Chow, 1997[59]	Hong Kong	Ho's QoL Scale	Not stated	Stroke	Cross-sectional	109	Probable
Cummins 1997[22]	USA	ComQol-I5	Face-to-face interview	LD	Cross-sectional	430	Yes
Cummins *et al.*, 1997[21]	Canada	ComQol-I5	Face-to-face interview	LD	Cohort	59	Yes
Dikmen *et al.*,1995[30]	USA	SIP	Written questionnaire	Brain injury	Cohort	410	Yes
Dorman *et al.*, 1997[13]	UK	EQ-5D	Written questionnaire with nurse present	Stroke	Cohort	146	Probable
Dorman *et al.*, 1997[27]	UK	EQ-5D	Mailed questionnaire	Stroke	Cohort	2253	Probable
Dorman *et al.*, 1997[27]	UK	SF-36	Mailed questionnaire	Stroke	Cohort	2253	Probable
Dorman *et al.*, 1997[37]	UK	EQ-5D	Written or face-to-face questionnaire	Stroke	Cohort	130	Probable
Dorman *et al.*, 1998[26]	UK	EQ-5D	Mailed questionnaire	Stroke	Cohort	2253	Probable
Dorman *et al.*, 1998[26]	UK	SF-36	Mailed questionnaire	Stroke	Cohort	2253	Probable
Fischer *et al.*, 1999[23]	USA	MS-QLI	Written questionnaire	MS	Cohort	300	Yes
Fischer *et al.*, 1999[23]	USA	SIP	Written questionnaire	MS	Cohort	300	Yes
Fischer *et al.*, 1999[23]	USA	SF-36	Written questionnaire	MS	Cohort	300	Yes
Fleming & Strong 1999[58]	Australia	PCRS	Face-to-face interview	Brain injury	Cohort	53	Yes
Gompertz *et al.*, 1994[51]	UK	NHP	Mailed questionnaire	Stroke	Cohort	158	Probable
Graham *et al.*, 1997[45]	UK	CQOL	Face-to-face interview	LD (mothers)	Cross-sectional	26	Yes

continued

TABLE 2 contd

Study	Country	QOL measure	Method of administration	Disease status	Study design	No. of participants	Cognitive impairment included?
Gulick, 1997[40]	USA	LSS	Mailed questionnaire	MS	Cross-sectional	153	Probable
Harwood *et al.*, 1994[48]	UK	NHP	Mailed questionnaire	Stroke	Cross-sectional	82	Probable
Hawkins *et al.*, 1995[44]	USA	LSI	Face-to-face interview	LD	Cohort	48	Yes
Jones *et al.*, 1997[41]	UK	SF-36	Written questionnaire	LD	Cross-sectional	71	Yes
Kaitaro *et al.*, 1995[54]	Finland	Quality of Life Questionnaire (Viitanen)	Face-to-face interview	Brain injury	Cross-sectional	19	Yes
Koskinen, 1998[20]	Finland	Life Satisfaction – Viitanen *et al.*	Face-to-face interview	Brain injury	Cohort	15	Yes
Kreuter *et al.*, 1998[43]	Sweden	VAS + Global QOL	Written questionnaire	Brain injury	Cross-sectional	92	Yes
Lam *et al.*, 1994[49]	Hong Kong	COOP/WONCA	Face-to-face interview	Stroke	Cross-sectional	84	Yes
Lehman, 1988[57]	USA	QOL Interview	Face-to-face interview	LD	Cohort	41	Yes
Lehman *et al.*, 1986[56]	USA	QOL Interview	Face-to-face interview	LD	Cross-sectional	191	Yes
Ormel *et al.*, 1998[19]	The Netherlands	SWB	Face-to-face or telephone interview	Stroke, MS, Cognitive Impairment	Cohort	5279	Yes
Ormel *et al.*, 1998[19]	The Netherlands	Life Satisfaction – Cantril's ladder	Face-to-face or telephone interview	Stroke, MS, Cognitive Impairment	Cohort	5279	Yes
Ouelette-Kuntz, 1990[46]	Canada	QUOLIS	Face-to-face interview with proxy	LD	Cohort	10	Yes
Pickard *et al.*, 1999[28]	Canada	SF-36	Telephone interview	Stroke	Cohort	161	Yes
Powell *et al.*, 1998[25]	UK	BICRO-39	Written questionnaire	Brain injury	Cross-sectional	235	Probable
Smith, 1992[24]	UK	SIP	Face-to-face interview	Brain injury	Cohort	20	Yes

continued

TABLE 2 contd

Study	Country	QOL measure	Method of administration	Disease status	Study design	No. of participants	Cognitive impairment included?
Smith *et al.*, 1998[35]	Canada	SIP	Face-to-face interview and proxies	Brain injury	Cohort	43	Probable
Sneeuw *et al.*, 1997[14]	The Netherlands	SIP	Face-to-face interview	Stroke	Cohort	437	Yes
Solari *et al.*, 1998[38]	Italy	MSQOL-54	Face-to-face interview	MS	Cohort	204	Yes
Solari *et al.*, 1999[39]	Italy	MSQOL-54	Written questionnaire	MS	Cross-sectional	204	Yes
Stuifbergen, 1995[52]	USA	QoL Index – MS version	Written questionnaire	MS	Cross-sectional	61	Yes
Stuifbergen, 1995[53]	USA	QoL Index – MS version (Ferrans & Powers, 1985)	Written questionnaire	MS	Cohort	61	Yes
Temkin *et al.*, 1988[33]	USA	SIP	Face-to-face interview	Head injury	Cohort	84	Yes
Turner *et al.*, 1997[50]	USA	Oregon QLQ Interviewer Rating Version	Face-to-face interview	LD	Case-control	50	Yes
Vandergriff & Chubon, 1994[42]	USA	LSS	Face-to-face interview	LD	Cross-sectional	120	Yes
Vickrey *et al.*, 1997[29]	USA	SF-36	Mailed questionnaire	MS	Cohort	171	Probable
Vickrey *et al.*, 1997[29]	USA	QoLQ for MS	Telephone questionnaire	MS	Cohort	171	Probable
Visser *et al.*, 1995[34]	The Netherlands	SIP	Written questionnaire	Stroke	Cross-sectional	53	Probable
Visser *et al.*, 1995[34]	The Netherlands	NHP	Written questionnaire	Stroke	Cross-sectional	53	Probable
Wilkinson *et al.*, 1997[47]	UK	SF-36	Face-to-face interview	Stroke	Cohort	96	Yes
Wilkinson *et al.*, 1997[47]	UK	NHP	Face-to-face interview	Stroke	Cohort	96	Yes
Zeldow & Pavlou, 1988[36]	USA	SIP	Face-to-face interview	MS	Cross-sectional	81	Probable

Appendix 5

Details of studies excluded from the review

TABLE 3

Study	Country	QOL measure	Method of administration	Disease status	Study design	No. of participants	Cognitive impairment included?
Cella *et al.*, 1996	USA	FAMS	Mailed questionnaire and written questionnaire	MS	Cohort	433	Unclear
Corrigan *et al.*, 1997	USA	SF-36	Written questionnaire	Brain injury	Cross-sectional	95	Unclear
Corrigan *et al.*, 1997	USA	SIP	Written questionnaire	Brain injury	Cross-sectional	95	Unclear
De Haan *et al.*, 1993	The Netherlands	SIP	Not reported	Stroke	Cohort	87	Unclear
Di Fabio *et al.*, 1997	USA	MSQOL-54	Written questionnaire	MS	Cohort	31	Unclear
Duncan *et al.*, 1997	USA	SF-36	Face-to-face or telephone interview	Stroke	Cross-sectional	304	Unlikely
Ebrahim *et al.*, 1986	UK	NHP	Mailed questionnaire	Stroke	Cross-sectional	153	Unlikely
Freeman *et al.*, 1996	UK	SF-36	Written questionnaire	MS	RCT	50	No
Gage *et al.*, 1996	USA	Utility Assessment	Written questionnaire	Stroke	Cross-sectional	20	Unlikely
Gompertz *et al.*, 1993	UK	NHP	Mailed questionnaire	Stroke	Cohort	21	Unlikely
Granger *et al.*, 1993	USA	SIP	Face-to-face interview	Stroke	Cross-sectional	18	Unclear
Hale *et al.*, 1998	South Africa	SSQ	Face-to-face interview	Stroke	Cohort	19	Unclear
Harper *et al.*, 1986	Canada	Quality of Life – Questionnaire	Written questionnaire	MS	Cross-sectional	256	Unclear
Hutchinson & Hutchinson, 1995	Ireland	Functional Limitations Profile (=SIP)	Face-to-face interview	MS	Cohort	50	Unclear
Indredavik *et al.*, 1998	Norway	VAS	Face-to-face interview	Stroke	RCT	62	Unclear
Indredavik *et al.*, 1998	Norway	NHP	Face-to-face interview	Stroke	RCT	62	Unclear
Jonsson *et al.*, 1996	Denmark	Danish LLQ	Written questionnaire	MS	Cross-sectional	43	Unlikely
King, 1996	USA	Quality of Life Index-Stroke version	Face-to-face interview	Stroke	Cross-sectional	86	No

continued

TABLE 3 contd

Study	Country	QOL measure	Method of administration	Disease status	Study design	No. of participants	Cognitive impairment included?
Lankhorst *et al.*, 1996	The Netherlands	DIP	Written questionnaire	MS	Cohort	73	Unclear
Mathias *et al.*, 1997	USA	HUI	Face-to-face interview	Stroke	Cohort	33	Unclear
O'Mahony *et al.*, 1998	UK	SF-36	Mailed questionnaire	Stroke	Cohort	73	Unclear
Parkin *et al.*, 1998	UK	MSQOL-54	Mailed questionnaire	MS	Cohort	102	Unclear
Pfennings *et al.*, 1999	The Netherlands	DIP	Written questionnaire	MS	Cohort	162	Unclear
Pfennings *et al.*, 1999	The Netherlands	SF-36	Written questionnaire	MS	Cohort	162	Unclear
Pfennings *et al.*, 1999	The Netherlands	COOP/ WONCA charts	Written questionnaire	MS	Cohort	162	Unclear
Pfennings *et al.*, 1999	The Netherlands	SF-36	Written questionnaire	MS	Cross-sectional	456	Unclear
Pfennings *et al.*, 1999	The Netherlands	SF-36 (Dutch version)	Written questionnaire	MS	Cross-sectional	87	Unlikely
Pfennings *et al.*, 1998	The Netherlands	DIP	Written questionnaire	MS	Cohort	187	Unclear
Pfennings *et al.*, 1998	The Netherlands	SF-36	Written questionnaire	MS	Cohort	187	Unclear
Provinciali *et al.*, 1999	Italy	FAMS	Not stated	MS	Cross-sectional	76	No
Robnett & Gliner, 1995	USA	QUAL-OT	Mailed questionnaire	MS	Cross-sectional	133	Unlikely
Rothwell *et al.*, 1997	UK	EQ-5D	Face-to-face interview	MS	Cross-sectional	42	Unclear
Rothwell *et al.*, 1997	UK	SF-36	Face-to-face interview	MS	Cross-sectional	42	Unclear

continued

TABLE 3 contd

Study	Country	QOL measure	Method of administration	Disease status	Study design	No. of participants	Cognitive impairment included?
Rudick *et al.*, 1992	USA	Farmer QoL Index	Face-to-face interview	MS	Cross-sectional	68	Unclear
Schuling *et al.*, 1993	The Netherlands	SIP	Face-to-face interview	Stroke	Cross-sectional	89	Unclear
Schuling *et al.*, 1993	The Netherlands	SIP	Face-to-face interview	Stroke	Cross-sectional	80	Unclear
Schwartz *et al.*, 1997	USA	Q-TWiST	Mailed questionnaire	MS	Cohort	79	Unclear
Schwartz *et al.*, 1999	USA	Health Status Questionnaire (= SF36 + 3)	Mailed questionnaire	MS	Cross-sectional	274	Unclear
Schwartz *et al.*, 1999	USA	QWB Index	Mailed questionnaire	MS	Cross-sectional	274	Unclear
Segal & Schall, 1994	USA	SF-36	Face-to-face interview	Stroke	Cohort	38	Unlikely
Temkin *et al.*, 1989	USA	Modified SIP	Written questionnaire	Brain injury	Cross-sectional	191	Unclear
Van Straten *et al.*, 1997	The Netherlands	SA-SIP30	Face-to-face interview	Stroke	Cohort	319	No
Vazquez-Barquero *et al.*, 1991	Spain	SIP	Written questionnaire	MS	Cross-sectional	50	Unclear
Vickrey *et al.*, 1995	USA	MSQOL-54	Mailed questionnaire	MS	Cross-sectional	183	Unclear
Viitanen *et al.*, 1988	Sweden	Life Satisfaction	Structured interview	Stroke	Cross-sectional	62	No
Yoon, 1997	Korea	SIP	Face-to-face interview	Stroke	Cross-sectional	119	Unclear

RCT, randomised controlled trial; SA-SIP30, Stroke-Adapted SIP (30 items)

Appendix 6

Summary of validity of instruments

TABLE 4 *COMQOL-15*[21,22]

Description: The ComQol-I5 is a multidimensional instrument including seven domains: material well-being, health, productivity, intimacy, safety, place in community, and emotional well-being. Each domain has an objective and a subjective component. The measurement of each objective domain is achieved by obtaining an aggregate score based on the measurement of three objective indices relevant to that domain, for example 'material well-being' is measured by an aggregate score of income, type of accommodation and personal possessions. The measurement of each subjective domain is achieved by obtaining a satisfaction score of that domain, which is weighted by the perceived importance of the domain for the individual.

Validation	Results
Population studied	The scale was validated using the same group of patients (430 patients with LD) in two different studies.[21,22] The studies included more than 50% of respondents with cognitive impairment
Choice of component variables	Not reported
Sensibility	Not reported
Consistency	*Internal consistency (Cronbach's alpha):* Objective scales – material well-being (0.14), health (0.64), productivity (0.35), intimacy (0.45), safety (0.56), community (0.15), emotional well-being (0.42), and for the total objective scale (21 items): 0.56. Subjective scales – for the seven components combined, importance (0.48), satisfaction (0.65) and importance × satisfaction (0.68) *Test–retest reliability:* Generally high reliabilities were obtained at 1–2 weeks, with the exceptions of the 'importance of intimacy' and 'satisfaction with safety' components. Beyond the 2-week interval, the retest correlations became 'somewhat erratic' according to the authors[21,22]
Accuracy	Not reported
Suitability	*Inter-domain correlations:* Correlations between the objective domains were all low, with the highest correlation of 0.25 for 'material well-being × intimacy'. In general there was found to be no significant correlation between objective and subjective domain scores, with the exception of objective health which correlated positively with both health importance ($r = 0.25$; $p < 0.05$) and health satisfaction ($r = 0.33$; $p < 0.001$)
Patient–proxy and inter-rater agreement	Relationships were found between objective scores and the third-party subjective estimates provided by carers. Significant correlations were found in three domains for importance (health = 0.25, $p < 0.05$; safety = 0.23, $p < 0.05$; community = 0.33, $p < 0.01$) and for satisfaction (material well-being = 0.29, $p < 0.05$; health = 0.28, $p < 0.05$; intimacy = 0.33, $p < 0.01$). These correlations did not conform to those of the clients themselves and may indicate that carer third party estimations were influenced by objective cues. In terms of the relative inter-domain hierarchy of importance and satisfaction, only the former achieved a significant degree of congruence between the client and carer estimations ($r = 0.96$, $n = 7$, $p < 0.01$ and $r = 0.55$, $n = 7$ and not stated, respectively). The lack of correspondence for estimations of satisfaction was most marked for the domain of 'community'. In terms of intra-domain data the degree of correspondence was generally slight. For ratings of importance, only the domain of 'intimacy' (0.23; $p < 0.05$) reached significance. For ratings of satisfaction significant agreement was recorded for the domains of 'health' (0.22; $p < 0.05$), 'productivity' (0.33; $p < 0.01$) and 'safety' (0.31, $p < 0.01$)

Conclusion: The studies included more than 50% of respondents with cognitive impairment, and their combined quality score was 0.5. Internal consistency was under 0.60 for the total objective scale and for all subscales except 'health'. Test–retest reliability was generally high within 2 weeks, and reported by the authors as 'somewhat erratic' beyond 2 weeks. Patient–proxy assessments showed generally little correspondence, carer estimations correlated more strongly with objective cues than with patient estimates. No data was reported for the accuracy and sensibility of the scale.

TABLE 5 *SF-36*[15,23,26–29,41,47]

Description: The SF-36 comprises eight health scales: physical functioning (ten items), role limitations-physical (four items), bodily pain (two items), general health (five items), vitality (four items), social functioning (two items), role limitations-emotional (three items) and mental health (five items). Two core dimensions of health: 'physical' and 'mental' can be derived from these eight scales.

Validation	Results
Population studied	Overall seven studies included 3142 participants. Four studies included 2600 stroke patients[15,26–28,47] two studies included 471 patients with MS[23,29] and one study included 71 patients with LD.[41] Most studies included less than 20% of respondents with cognitive impairment. 20–50% of respondents were cognitively impaired in two studies.[23,47] Only one study included more than 50% with cognitive impairment[41]
Choice of component variables	Not applicable
Sensibility	*Ease of use:* One study among stroke patients found the mean time to complete the assessment was 8 minutes.[15] Completion rates among MS patients were found to range from 96% on 'role limitations' to 100% on 'social function'.[29] Another study compared response rates for SF-36 and EQ-5D.[27] Response and 'response with no missing data' were 60% SF-36 first mailing and 75% for the SF-36 after second mailing (OR = 1.31; 95% CI, 1.1 to 1.6; $p = 0.002$ and OR = 1.35; 95% CI, 1.1 to 1.6; $p = 0.003$, respectively). Taking into account the missing data the overall response rate for SF-36 was 55% (OR = 1.64; 95% CI, 1.4 to 1.9; $p < 0.0001$), with 50% of the questionnaires completed by patients rather than their carers. 50% of the respondents reported dependency in activities of daily living *Skewness:* Among 171 MS-patients the 'role limitations-physical' scale showed a 'floor' effect, and the 'role limitations-emotional' scale showed a 'ceiling' effect, with nearly half of the respondents scoring the lowest (48.5%) or highest (47.9%) score[29]
Consistency	*Internal consistency:* Among 300 MS patients the Cronbach's alpha scores for the SF-36 subscales alphas were good to excellent (0.75–0.94) with the one exception of 'social function'.[23] Cronbach's alpha scores ranged from 0.79 for 'social function' to 0.96 for 'physical function' in another group of 171 MS patients.[23] Among 849 stroke patients the Cronbach's alpha score were 0.95 for 'physical functioning', 0.94 for 'physical role functioning', 0.80 for 'social functioning', 0.87 for 'bodily pain', 0.86 for 'mental health', 0.96 for 'emotional role functioning', 0.83 for 'general health' and 0.81 for 'vitality'.[27] In another study among stroke patients, the Cronbach's alpha scores ranged between 0.6 for 'vitality' and 0.9 for 'physical functioning', 'bodily pain' and 'role limitations emotional'.[15] Finally, among 71 patients with LD, the Cronbach's alpha scores and item total biserial correlations for each subscale were largely acceptable with the exception of the 'social functioning' scale where both the mean item-total correlation and alpha (0.33) were low[41] *Test–retest reliability:* ICCs ranged from 0.30 for 'mental health', and 0.60 for 'emotional role functioning', to 0.89 for 'physical role functioning' in a study of 2253 stroke patients.[26] Among 171 MS patients ICCs ranged from 0.64 for 'social function' to 0.96 for 'physical function'[29] *Factor structure:* Among 71 patients with LD, the SF-36 subscale scores were subjected to a principle components factor analysis followed by a Varimax rotation replicating the analysis.[41] A two-factor solution was obtained and the factors could be clearly interpreted as a physical health factor and a mental health factor with 'vitality' and 'social functioning' loading mainly on the mental health factor but having a reasonable loading on the physical health factor. The physical health factor consisted of 'physical functioning' (factor score = 0.59), 'role-physical' (0.78), 'bodily pain' (factor score = 0.75) and 'general health' (factor score = 0.83). The mental health factor consisted of 'vitality' (factor score = 0.62), 'social functioning' (factor score = 0.58), 'role-emotional' (factor score = 0.75) and 'mental health' (factor score = 0.90)

continued

TABLE 5 contd *SF-36*[15,23,26–29,41,47]

Validation	Results
Accuracy	*Concurrent validity:* Among 90 stroke patients divided into two groups in terms of whether their Barthel Index scores were high or low. Patients with high scores on the Barthel Index were associated with higher scores on the SF-36, especially with regards to the 'physical functioning' and 'general health' subscales.[15] Similarly, higher scores on the GHQ 28 were associated with higher SF-36 scores, especially for the subscales 'social functioning', 'role limitations-emotions' and 'mental health'
Suitability	*Construct validity:* Correlations between the SF36 subscales and impairment measures were weak. Among 300 MS patients with the one exception of the 'physical' subscale which correlated strongly with the EDSS and moderately with quantitative measures (walk-test and 9-hole peg test)[23] Correlations between SF-36 and the Barthel Index were not significantly different if the modified version instead of the original SF-36 questionnaire was used.[47] All of the various SF-36 dimensions showed significant correlations with the individual subject's scores and their Barthel Index score. The correlations were highest for 'physical functioning' (0.810) and lowest for 'role-emotional' (0.217) *Discriminant validity:* One study looked at the sensitivity of SF-36 to important clinical differences, like severity of symptoms, disability, days unable to attend work or school in the past month, and overall QoL.[29] All SF-36 subscales had satisfactory sensitivity, except for the 'role-limitations emotional' scale, which discriminated poorly on all comparisons except overall QoL; the 'pain' scale discriminated poorly for days unable to attend work or school
Patient–proxy and inter-rater agreement	*Patient–proxy agreement:* Test–retest reliability was assessed using ICC for forms completed by the patient, proxies and by both groups. The ICCs were 0.80/0.59/0.74 (respectively) for 'physical functioning', 0.89/0.45/0.67 for 'physical role functioning', 0.79/0.76/0.80 for 'social functioning', 0.81/0.65/0.75 for 'bodily pain', 0.30/0.24/0.28 for 'mental health', 0.60/0.50/0.57 for 'emotional role functioning', 0.81/0.71/0.79 for 'general health' and 0.77/0.55/0.70 for 'vitality'[26] *Inter-rater reliability:* Among 71 patients with LD and their carers, moderate reliability was achieved.[41] However, the 'role-physical' scale and the 'role-emotional' scale were considerably less reliable than the other subscales. The level of reliability evident with the third party version of the scale was such that it would not be useful for comparison across individuals but was within the range required to make group comparisons. Pearson's *r* correlation coefficients for the SF-36 subscales were low for 'social functioning' (0.19) and 'mental health' (0.49), fair for 'general health' (0.60), 'vitality' (0.63), 'physical functioning' (0.76), and 'bodily pain' (0.79), and high for 'role-physical' (0.80) and 'role-emotional' (0.85)

OR, odds ratio; CI, confidence interval; GHQ, General Health Questionnaire

Conclusion: Most of the studies included fewer than 20% of cognitively impaired respondents.[15,26–29] Two studies included 20–50% of respondents with cognitive impairment, but these had low quality scores (0.147 and 0.323). Only one study included more than 50% of respondents with cognitive impairment.[41] This final study had a quality score of 0.3. Results from two of the studies showed good-to-excellent internal consistency for the total scale and for all subscales with one exception ('social function').[23,47] Correlations between SF-36 subscales and impairment measures were weak, with one exception (SF-'physical' correlated strongly with the EDSS and moderately with quantitative measures). One study showed that the internal consistency and item total biserial correlations for each subscale are largely acceptable with the exception of the 'social functioning' scale where both the mean item-total correlation and Cronbach's alpha (0.33) were low.[41] Through factor analysis on the subscales, a two-factor solution was obtained and the factors can clearly be interpreted as a physical health factor and a mental health factor. Inter-rater reliability between patients with LD and their carers was moderate; however, the 'role-physical' scale and the 'role-emotional' scale were considerably less reliable.

TABLE 6 *LSI*[44]

Description: The index consists of 36 items, including five dimensions: living arrangement, relationships and social activities, job status, health, and general happiness.

Validation	Results
Population studied	48 patients with LD; more than 50% with cognitive impairment
Choice of component variables	Not applicable
Sensibility	Not reported
Consistency	*Internal consistency:* Cronbach's alpha scores were acceptable to good for most of the subscales (living arrangement (0.79), job status (0.64), health (0.75), general happiness (0.84)), and the total index (0.87)), but poor for 'relationships and social activities' (0.48)
Accuracy	Not reported
Suitability	*Convergent validity:* As predicted the LSI showed significant negative correlations with 'leisure interests' (r = –0.38) and 'leisure-constraints' (r = –0.45) and no significant relation with broad cognitive functioning (r = –0.30). Contrary to predictions, the LSI showed no significant negative correlations with 'leisure-activity participation' (r = 0.02) and 'leisure-preference' (r = –0.06)
Patient–proxy and inter-rater agreement	Not reported

Conclusion: The LSI was validated in one study, which included 48 patients with LD.[44] More than 50% of the respondents had cognitive impairment, and the quality score was 0.3. Internal consistency was acceptable to good for most scales, but poor for 'relationships and social activities'. The convergent validity results were somewhat unclear.

TABLE 7 *LSS*[40,42]

Description: The LSS is a 20-item Likert-type rating scale developed by Chubon for use in assessing perceived life quality by persons with disabilities and chronic illness, as well as the general population. The items are clear and concise, with readability estimated to be in the fourth grade to fifth grade range. Although designed to be self-administered, provisions are included for oral administration to persons who have reading or other limitations that prohibit them from completing the written self-report form. The scale was developed around ten commonly accepted QoL domains (including work, leisure, health, love/affection, self-esteem) and ten additional areas (including stress, mobility, autonomy, energy level, social support, mood/affect, and public support). These were derived through a critical incidents-based study of adults with various chronic illnesses and/or disability including MS.[40] Items were rated on a six-point scale: disagree very strongly (1), to agree very strongly (6); a one- to seven-point scale has also been used.[42]

Validation	Results
Population studied	153 MS patients[40] and 120 patients with LD.[42] Fewer than 20% of the MS patients had cognitive impairment, but more than 50% of the patients with LD had cognitive impairment
Choice of component variables	Not applicable
Sensibility	Not reported
Consistency	*Internal consistency:* Among patients with LD the Cronbach's alpha score was 0.74 for the high-IQ group and 0.61 for the low-IQ group.[42] These scores were lower than those reported for non cognitive disabilities, but were sufficient to render the results usable in interpreting the group data In a sample of 153 MS patients Cronbach's alpha was 0.89, based on 19 of 20 items used in the study.[40] One item pertaining to public support was deleted owing to non-response by approximately half of the participants

continued

TABLE 7 contd *LSS*[40,42]

Validation	Results
Accuracy	Not reported
Suitability	*Convergent validity:* One of the studies found moderate correlations with the Beck Depression Inventory (BDI; $r = -0.61$), the LSI-A ($r = 0.77$) and the Index of Well-Being ($r = 0.65$).[40] Low correlations with the RCAS ($r = 0.35$, $p < 0.01$), the ABS ($r = 0.35$, $p < 0.01$), and the ABS Part II ($r = 0.04$; p = not significant) were also reported.[42] However the LSS correlated significantly with the PASSING scores ($r = 0.91$; $p = 0.01$)
Patient–proxy and inter-rater agreement	Not reported

Conclusion: The LSS was validated in two studies, which included 153 MS patients[40] and 120 patients with LD.[42] Fewer than 20% of the MS patients had cognitive impairments, but more than 50% of the patients with LD had cognitive impairments. The quality for the LD study was 0.3, and the internal consistency was 0.74 for the high-IQ group and 0.61 for the low-IQ group. This showed that the internal consistency was lower but sufficient for people with cognitive impairment. The convergent validity showed that the LSS is sensitive to living conditions (PASSING), and less sensitive to mental deficiencies (ABS) and independence (RCAS).

TABLE 8 *CQOL*[45]

Description: The CQOL consists of 15 domains (getting about and using hands, doing things for self, soiling or wetting, school, out of school activities, friends, family relationships, discomfort due to bodily symptoms, worries, depression, seeing, communication, eating, sleep and appearance).

Validation	Results
Population studied	26 mothers of patients with LD.[45] More than 50% of the patients had cognitive impairment
Choice of component variables	Not reported
Sensibility	Not reported
Consistency	*Internal consistency:* Cronbach's alpha for combined scales measuring function, satisfaction and upset for all mothers ($n = 75$), including mothers of children with CPD and PD, were 0.81, 0.86 and 0.86, respectively *Test–retest reliability:* The correlation of the combined function scores was 0.83. Individual correlations for the mothers' scores ranged from 0.11 to 1.00 (based on 19 mothers of patients with LD and 15 mothers of children with CPD), with the majority between 0.4 and 0.7. The mothers of children with LD obtained the highest level of reliability
Accuracy	Not reported
Suitability	*Convergent validity:* The correlation between the mothers' total function score and the Child's Global Adjustment Score was 0.64 ($p = 0.01$) *Discriminant validity:* The mothers of LD children rated their children as functioning less well than either of the other two groups (mothers of CPD children and mothers of PD children) on 12 of the 15 items (with the exception of 'family relationships', 'depression' and 'physical discomfort'). The ratings were significantly worse on six of the items
Patient–proxy and inter-rater agreement	Not reported

Conclusion: The CQOL was validated in one study, which included 26 mothers of patients with LD.[45] More than 50% of the patients had cognitive impairment and the quality score was 0.25. Results showed that internal consistency for the combined scales was good, and test–retest reliability was acceptable. Convergent and discriminant validity results were as predicted.

TABLE 9 *QUOLIS*[46]

Description: The QUOLIS consists of 78 statements or indicators, grouped according to 12 domains which are areas of support: health services, family or guardianship, income maintenance, education or employment, housing and safety, transportation, social or recreational, religious or cultural, case management, advocacy, counselling and aesthetics. Each indicator is then rated on four counts, that is according to four dimensions: support, access, participation, and contentment.

Validation	Results
Population studied	Ten patients with LD, more than 50% of whom had cognitive impairment[46]
Choice of component variables	Not reported
Sensibility	Not reported
Consistency	*Internal consistency:* Cronbach's alpha scores for the contentment dimension ranged from 0.79 to 0.99
Accuracy	Not reported
Suitability	Not reported
Patient–proxy and inter-rater agreement	*Inter-rater reliability:* Inter-rater reliability for the contentment dimension ranged from 0.48 to 0.95. Intra-rater agreement showed that 73% (35/48) of the scores were acceptable. For the 13 scores lacking agreement refinements were suggested. Inter-rater agreement showed that 65% (31/48) of the scores were acceptable

Conclusion: The QUOLIS was validated in one study, which included only ten patients with LD.[46] More than 50% of the patients had cognitive impairment, and the quality score was 0.25. Results showed that internal consistency was good, and inter-rater agreement was acceptable for 73% of the scores.

TABLE 10 *VAS-QOL*[43,55]

Description: Global QoL was assessed using a VAS with endpoints labelled as 'very low' and 'very high'.[43] The scale was transformed to match a 0–100 scale. Another study used two VAS, which had bars graded as 'worst possible' at the bottom and 'best possible' at the top.[55] Patients were asked to mark on the left bar the QoL before the stroke and on the right bar the present QoL. The difference in height before and after stroke was calculated.

Validation	Results
Population studied	96 stroke patients less than 20% of whom were cognitively impaired[55] and 92 patients with traumatic brain injury (TBI), more than 50% of whom were cognitively impaired[43]
Choice of component variables	Not applicable
Sensibility	Not reported
Consistency	Not reported
Accuracy	Not reported
Suitability	*Convergent validity:* One of the studies validated changes in QoL marked on the bars against answers on three questions about the total life situation. Increasing degree of change as expressed verbally was thus found to correspond to increasing change of QoL calculated from the graphs (from a minimum of –1.8 to a maximum of –44.8)[55] Mood (HADS; $r = –0.56$; $p < 0.001$), physical and social functioning (SIP; $r = –0.48$; $p < 0.001$) and severity of the disability (GOS; $r = 0.48$; $p < 0.001$) were significantly correlated to the TBI persons' global QoL, as was the time since injury ($r = 0.27$; $p < 0.01$).[43] Partial and multiple correlations determined that for TBI patients the most important predictors were mood, physical and social functioning, severity of disability and time since injury explaining 46% of the variance

continued

TABLE 10 contd *VAS-QOL*[43,55]

Validation	Results
Patient–proxy and inter-rater agreement	Not reported

TBI, traumatic brain injury; HADS, Hospital Anxiety and Depression Scale; GOS, Glasgow Outcome Scale

Conclusion: The VAS-QoL was validated in two studies, which included 96 stroke patients[55] and 92 patients with brain injury.[43] Less than 20% of the stroke patients had cognitive impairment, while more than 50% of the brain injury patients were cognitively impaired. The quality scores for the studies were 0.155 and 0.2.43 The latter study found strong correlations between the VAS-QoL and mood, physical and social functioning and severity of disability, and moderate correlation with the time since injury.

TABLE 11 *SWB*[19]

Description: This is an eight-item scale for measuring physical and social well-being. Factor analysis showed one-dimensionality, and thus ratings were summed.

Validation	Results
Population studied	5279 patients including 128 stroke patients (or consequences of stroke), ten MS patients, and 300 cognitively impaired patients.[19] More than 50% of the participants had cognitive impairment
Choice of component variables	Not applicable
Sensibility	Not reported
Consistency	*Internal consistency:* Cronbach's alpha scores ranged from satisfactory (i.e. 0.75) to very good (i.e. 0.91)
Accuracy	Not reported
Suitability	Not reported
Patient–proxy and inter-rater agreement	Not reported

Conclusion: The SWB was validated in one study, which included 5279 patients with stroke, MS and cognitive impairment.[19] Although the study included more than 50% of respondents with cognitive impairments, the quality score was only 0.1, showing that there was very limited validity information. The results showed that the internal consistency of the scale in people with cognitive impairment was good.

TABLE 12 *Life Satisfaction (Cantril's ladder)*[19]

Description: This scale asks the question: 'Here is a picture of a ladder, suppose that we say the top of the ladder represents the best possible life for you and the bottom represents the worst possible life for you. Where on the ladder do you feel you personally stand at the present time?' The ladder has ten steps (range 1–10).

Validation	Results
Population studied	5279 patients including 128 stroke patients (or consequences of stroke), ten MS patients, and 300 cognitively impaired patients[19]
Choice of component variables	Not applicable
Sensibility	Not reported
Consistency	*Internal consistency:* The Cronbach's alpha scores ranged from satisfactory (i.e. 0.75) to very good (i.e. 0.91)

continued

TABLE 12 contd *Life Satisfaction (Cantril's ladder)*[19]

Validation	Results
Accuracy	Not reported
Suitability	Not reported
Patient–proxy and inter-rater agreement	Not reported

Conclusion: Cantril's ladder was validated in one study, which included 5279 patients with stroke, MS and cognitive impairment.[19] Although the study included more than 50% of respondents with cognitive impairment, the quality score was only 0.1, showing that there is very limited validity information. The results show that the internal consistency of the scale in people with cognitive impairment is good.

TABLE 13 *SIP*[14,23,24,30–36]

Description: The SIP consists of 136 items describing the impact of ill health on behaviour in 12 dimensions: sleep and rest, emotional behaviour, body care and movement, household management, mobility, social interaction, ambulation, alertness behaviour, communication, work, recreation and pastimes, and eating. Weighted sum scores are obtained for the overall profile, physical and psychosocial subtotals, and separately for each of 12 dimensions.

Validation	Results
Population studied	1542 participants in total. Five studies included 576 patients with brain injury[24,30,32,33,35] three studies included 585 stroke patients[14,31,34] and two studies included 381 patients with MS.[23,36] 20–50% of the participants were cognitively impaired in three of the studies.[14,23,24] The rest of the studies had less than 20% of respondents whom were cognitively impaired[30–34,36]
Choice of component variables	Not applicable
Sensibility	*Ease of use:* Among 53 stroke patients, it was found that the mean assessment time of 23.5 minutes (SD = 7.2), was significantly longer than the control group (17.3 minutes, SD = 7.1)[34]
Consistency	*Internal consistency:* Cronbach's alpha for the total SIP score ranged from 0.89 to 0.96.[23,32,33] For SIP subscales Cronbach's alphas were lower, ranging from 0.63 to 0.90, the internal consistency of the 'sleep and rest' and 'eating' subscales was relatively poor, based on patients and proxies in one study and based on patients alone in another study[23] *Test–retest reliability:* Using Spearman correlation coefficients among 42 stroke patients, the coefficients were 0.82 for the psychosocial sum, 0.80 for the physical sum the score and 0.86 for the total SIP.[34] The correlation coefficient for the SIP-'sleep' subscale was lowest (0.31) and values for the other subscales were not reported
Accuracy	Not reported
Suitability	*Concurrent validity:* Correlations between SIP-scores and the MMPI-Depression score were examined in 19 brain injury patients.[32] The correlation for the overall SIP score with depression was 0.566, while correlations with depression for the psychosocial dimension and physical dimension were 0.382 and 0.482, respectively. The psychosocial subscales (social introversion $r = 0.338$, communication $r = 0.168$, alertness behaviour $r = 0.171$, emotional behaviour $r = 0.306$) correlated less with depression than the physical subscales (ambulation $r = 0.427$, mobility $r = 0.471$, body care and movement $r = 0.435$). Most of the other SIP subscales correlated strongly with depression (sleep and rest $r = 0.469$, work $r = 0.670$ ($n = 19$), home management $r = 0.599$, recreation and pastimes $r = 0.504$), except for the subscale eating ($r = 0.188$) SIP scores were significantly poorer for stroke compared with controls on 'emotional behaviour' and 'household management', but not on any of the other dimensions[34]

continued

TABLE 13 contd *SIP*[14,23,24,30–36]

Validation	Results
	Correlations between SIP scores and seven neuropsychological measures (mean impairment rating, WAIS Digit Symbol Scale score, DRAT: Trial 3 average cycle time, Deterioration Quotient, Wechsler Memory Quotient, WAIS FSIQ, Trail Making Test – Part B) were reported for 38 stroke patients.[31] The mean impairment rating and the Trail Making Test – Part B scores predicted most successfully the three overall SIP scores. The DRAT: Trial 3 average cycle time was more strongly correlated to the SIP-physical dimension (r = 0.34) than to the psychological dimension (r = 0.03). While the Deterioration Quotient and Wechsler Memory Quotient correlated more strongly to the SIP psychological dimension (r = 0.28 and –0.16 respectively) compared with the physical dimension (r = 0.13 and –0.04) Correlations between SIP-scores and head injury severity were measured and increasing head injury severity was significantly correlated with increasing levels of SIP dysfunction on all subscales except the 'emotional behaviour' scale.[30] The magnitudes of the other correlations were rather modest, ranging from 0.13 to 0.34 Among 300 MS patients SIP subscales correlated at least moderately with one or more objective measures, and the pattern of correlations between SIP-subscales and quantitative measures of function (walk-test and 9-hole peg test) was not entirely predictable[23] Among 43 brain injury patients the total SIP score was significantly related to CIQ-productivity (–0.41) and life satisfaction (–0.44).[35] The SIP-'physical' subscale was significantly related to CIQ 'productivity' (–0.51), and SIP-'psychosocial' was significantly related to life satisfaction (–0.51). None of the other correlations between SIP and CIQ-subscales ('home', 'social' and 'productivity'), social support and life satisfaction were significant
Patient–proxy and inter-rater agreement	*Patient–proxy agreement:* In a study of 437 stroke patients, it was found that the individual patient level, chance-corrected agreement between patient and proxy scores ranged from moderate for the 'eating' subscale (ICC = 0.47) to excellent for 'ambulation' (ICC = 0.80) and 'body care and movement' (ICC = 0.82).[14] Good to excellent was noted for physical (ICC = 0.85), psychosocial (ICC = 0.61) and total SIP (ICC = 0.77). ICCs for the other subscales ranged between 0.52 and 0.69 *Validity of proxy ratings:* The mean proxy-rated SIP scores of the physical and psychosocial dimensions and total SIP were found to be significantly associated with patients' Rankin grades.[14] However, proxy scores were systematically higher than patient scores *Correlation between patients and 'close others':* Among 20 brain injury patients it was found that correlations were high for all categories (ambulation (0.859), body care and movement (0.720), mobility (0.825), recreation (0.660), social interaction (0.646), alertness (0.745), eating (0.654), communication (0.970), work (0.965) and overall percentage of dysfunction (0.905)). The only exceptions were for 'emotion' (0.395), 'sleep and rest' (0.529) and 'household management' (0.583).[24] It was suggested that 'emotion' and 'sleep and rest' were the most difficult categories for relatives to assess. A comparison of mean scores showed that the difference between patients' and relatives' reports was not significant on any category except 'alertness'. In this category patients reported a mean of 33% dysfunction as opposed to 24% reported by relatives. This difference was significant at the 5% level

SD, standard deviation; MMPI, Minnesota Multiphasic Personality Inventory; WAIS FSIQ, Wechsler Adult Intelligence Scale – Full Scale Intelligence Quotient; DRAT, Discriminant Reaction Averaging Time Test; CIQ, Community Integration Questionnaire

Conclusion: The SIP was validated in ten studies, which included a total of 1542 respondents.[14,23,24,30–34,36] However, the level of validation was limited in people with cognitive impairment. Only three studies included 20–50% of respondents with cognitive impairment[14,23,24] and the remaining studies all included less than 20% cognitively impaired persons.[30–34,36] The quality scores for the 20–50% studies ranged from 0.3 to 0.4, showing that the level of validation was limited. These three studies reported that internal consistency for the total SIP-score was generally good to excellent. For certain SIP subscales the internal consistency was relatively poor (i.e. 'sleep and rest' and 'eating' subscales). Correlations between the SIP-subscales and quantitative measures of function were not entirely predictable and agreement between patient and proxy scores ranged from moderate to excellent. However, proxy scores were systematically higher than patient scores. The 'emotion' and 'sleep and rest' subscales were the most difficult categories for relatives to assess.

TABLE 14 *MS-QLI*[23]

Description: The MS-QLI is a modular MS-specific HRQL instrument consisting of a widely used generic measure, the Health Status Questionnaire (SF-36), supplemented by nine symptom-specific measures, covering: fatigue, pain, bladder function, bowel function, emotional status, perceived cognitive function, visual function, sexual satisfaction, and social relationships.

Validation	Results
Population studied	Validity was assessed in one study including 300 MS patients, 20–50% of whom were cognitively impaired[23]
Choice of component variables	*Content validity:* Three expert panels (MS patients and families, neurologists specialising in MS, and allied health professionals skilled in assessing and treating MS) reviewed a set of candidate measures. Consequently, two items ('bowel and bladder function', and 'caregiving') were added in addition to a new domain ('sexual function')
Sensibility	*Ease of use*: It took approximately 45 minutes to complete the full questionnaire consisting of 137 items and 80 abbreviated scales
Consistency	*Internal consistency:* Cronbach's alpha scores for all of the symptom-specific scales (including 'fatigue impact', 'perceived deficits', 'mental health' and 'social support') were good to excellent (0.77–0.97). The scores were high for the generic HRQL summary scales (0.89–0.95), and good to excellent (0.75–0.94) for the SF-36 subscales with the exception of 'social function'
Accuracy	Not reported
Suitability	*Construct validity:* The construct validity of the generic HRQL (SF-36) was supported at both the summary and subscale levels, e.g. SF-36-'physical' and SIP-'physical' were strongly correlated ($r = -0.62$), as were SF-'mental' and SIP-'psychosocial' ($r = -0.51$). Correlations between dissimilar constructs were weak however ($r < 0.30$). The construct validity of most symptom-specific measures was also supported by modest ($r < 0.40$) correlations for the different constructs of the MS-QLI with 'pure' measures, including measurements on 'bladder', 'bowel', 'sexual' and 'visual functioning', MOS-'pain', and MOS-'social support'. Correlations between the SF-36 subscales and impairment measures were weak, with the exception of SF-'physical' which correlated strongly with the EDSS and moderately with quantitative measures (walk-test and 9-hole peg test)
Patient–proxy and inter-rater agreement	Not reported

Conclusion: The MS-QLI was validated in one study, which included 300 MS patients.[23] 20–50% of respondents were cognitively impaired, and the study's quality score was 0.5, showing that the level of validation was extensive. However, the study findings may not be fully applicable to people with cognitive impairment, as only 20–50% of the study participants were cognitively impaired. The results showed that the internal consistency of the total MS-QLI and the subscales was good, with one exception ('social function'). The scale also appeared to have good construct validity.

TABLE 15 *COOP/WONCA charts*[49]

Description: The Dartmouth COOP/WONCA charts consist of six questions on physical ability, feelings, daily activity, social activity, change in health and overall health, respectively. Each question has five responses, illustrated by drawings, on a scale of 1–5, ranging from 'no limitation' to 'severe limitation' for the first four questions, and from 'much better' to 'much worse' for change of health, and 'excellent' to 'poor' for overall health.

Validation	Results
Population studied	84 stroke patients, 20–50% of whom were cognitively impaired[49]
Choice of component variables	Not applicable
Sensibility	*Ease of use and acceptability:* The interviewer found the charts very easy to administer and most of the patients took less than 5 minutes to complete them. All of the patients were able to complete the charts. Nobody complained about them and many actually liked the drawings. They often pointed to the drawings rather than the words to indicate their responses. The drawings were very helpful to patients who were illiterate. All but one patient said that they understood what the charts were asking. One patient was not sure of the meaning of the charts on physical fitness, feelings and daily activities *Usefulness:* Five doctors saw the 84 patients and found the information obtained by the COOP/WONCA charts extremely useful in 5% of cases, moderately useful in 60%, slightly useful in 33% and useless in 2% of the patients
Consistency	Not reported
Accuracy	Not reported
Suitability	Not reported
Patient–proxy and inter-rater agreement	Not reported

Conclusion: The COOP/WONCA charts were validated in one study, which included 84 stroke patients.[49] Only 20–50% of the respondents were cognitively impaired and so there was little information to indicate the validity of the charts in this patient population. In addition the study only had a quality score of 0.2. However, the results showed that the COOP/WONCA charts were easy to use and well accepted by patients and doctors alike.

TABLE 16 *Oregon QLQ interviewer rating version*[50]

Description: The Oregon QLQ was primarily developed to measure the QoL of mentally ill patients and later to study case management and QoL of mentally ill community patients. The questionnaire consists of 141 items pertaining to services, responsibilities, performances, and satisfaction of client needs in each domain. The items are organised into 21 scales: satisfaction with home, structure and support at home, self and home maintenance, adequacy of income, physical health, meaningful use of time, psychological distress, well-being, interpersonal relations, housing services, home management services, money management services, education services, employment services, social recreational services, physical health, mental health, nutritional services, transportation and protective services. Most items are scored from 1 to 4 (i.e. very inadequate (1) to very adequate (4)).

Validation	Results
Population studied	62 patients with LD, 20–50% of whom had cognitive impairment[50]
Choice of component variables	Not applicable
Sensibility	Not reported

continued

TABLE 16 contd *Oregon QLQ interviewer rating version*[50]

Validation	Results
Consistency	*Internal consistency*: This was assessed using Cronbach's alpha. Ten scales had fair-to-high Cronbach's alphas (defined as above 0.50): home satisfaction, self and home maintenance, physical health, leisure time, physical distress, well-being, interpersonal relations, money services, educational services and employment services. Seven had low Cronbach's alphas: adequacy of income (0.2940), housing services (0.0149), home making services (0.2880), social and recreational services (0.1485), mental health (0.3574), nutritional services (0.3378) and transportation services (–0.2657). 'Structure and support' was not included in the calculations as this was a single-item scale
Accuracy	Not reported
Suitability	Not reported
Patient–proxy and inter-rater agreement	*Inter-rater reliability:* Overall the inter-rater reliability was satisfactory

Conclusion: The Oregon QLQ was validated in one study that included 62 patients with LD, 20–50% of whom had cognitive impairment.[50] The quality score for the study was 0.2, showing that the level of validation was limited. The results showed that the internal consistency was below 0.5 for seven out of 17 scales, and overall the inter-rater reliability was satisfactory.

TABLE 17 *NHP*[34,47,48,51]

Description: The NHP consists of 38 items describing health-related behaviour in six dimensions: pain, physical mobility, sleep, emotional reactions, energy and social isolation (part 1); and seven 'yes/no' questions concerning domains of daily life (part 2).

Validation	Results
Population studied	A total of 389 stroke patients in four studies.[34,47,48,51] Less than 20% of the patient were cognitively impaired in two studies[34,51] and in the remaining studies only 20–50% of the participants were cognitively impaired[47,48]
Choice of component variables	Not applicable
Sensibility	*Ease of use:* One study which used the NHP to assess QoL in 53 stroke patients, reported a mean assessment time of 8.6 minutes (SD = 4.0)[34]
Consistency	*Test–retest reliability:* Test–retest reliability was assessed using Spearman's rank correlation coefficients in 42 stroke patients.[34] The correlation coefficients were: pain (0.65), physical mobility (0.68), sleep (0.70), emotional reactions (0.80), energy (0.86) and social isolation (0.86). The correlation coefficient for Part 2 of the scale was 0.88
Accuracy	Not reported
Suitability	*Construct validity:* Strong correlations were reported between the London Handicap scale and the NHP Part 1 overall score (correlation coefficient = –0.42), and Part 1 subscales 'physical mobility' (–0.52), 'energy' (–0.36), 'social isolation' (–0.30) and 'pain' (–0.31).[48] Weak correlations were reported between the London Handicap scale and the NHP Part 2 overall score (–0.28) and the Part 1 subscales 'emotion' (–0.28) and 'sleep' (–0.19)[48] A strong correlation was also identified between the NHP and the EADL in terms of NHP 'physical mobility' (correlation coefficient = –0.66 at 6 months and –0.70 at 12 months), 'energy' (–0.41;–0.31) and 'pain' scores (–0.35 at 6 and 12 months).[51] Strong correlations were also observed between EADL and carer NHP 'energy' (–0.25 at 6 months; –0.25 at 12 months), 'emotion' (–0.26 and –0.30) and 'pain' (–0.28 and –0.18) scores All of the dimensions of NHP apart from 'sleep' (–0.189) had a significant correlation with the Barthel Index score, the correlation coefficients ranging from –0.423 for 'emotion', to –0.840 for 'physical mobility'[47]

continued

TABLE 17 contd *NHP*[34,47,48,51]

Validation	Results
	Discriminant validity: NHP scores were found to be significantly poorer for stroke patients compared with controls for 'energy' and the Part 2 questions, but not on any of the other dimensions[34]
Patient–proxy and inter-rater agreement	Not reported

EADL, Extended Activities of Daily Living

Conclusion: The NHP was validated in four studies including 389 stroke patients. There was very limited information on the validity of the NHP in people with cognitive impairment. In two of the studies less than 20% of the respondents had cognitive impairment.[34,51] In the other two studies 20–50% of respondents were cognitively impaired.[47,48] However, the quality score for both of these studies was only 0.1.

TABLE 18 *BICRO-39*[25]

Description: The BICRO-39 consists of 76 items, categorised into four domains: independent personal functioning (26 items); leisure and work activities (18 items); personal relationships (ten items); psychological and behavioural adjustment (22 items). The WHO handicap dimensions of 'mobility', 'physical independence', 'orientation', 'occupation', and 'social integration' were covered by items within the first three categories. The WHO dimension 'socio-economic self-sufficiency' was not included. The domain of 'psychological and behavioural adjustment' maps only indirectly onto the WHO handicap dimension: 'orientation'; within the International Classification of Impairments, Disabilities and Handicaps (ICIDH), this aspect of functioning is seen as disability of behaviour and communication. Items were rated on a six-point scale (0 to 5). Three forms of the questionnaire were generated: patient pre-form (to be filled in retrospectively); patient post-injury; and carer post-injury (to be completed by carer or significant other).

Validation	Results
Population studied	235 patients with brain injury.[25] Less than 20% of the patients were cognitively impaired
Choice of component variables	Thirty-nine out of the original 76 items were retained as a result of high factor loadings. The shortened version of the scale produced essentially the same results and was more readily administered in clinical practice. The validity results were based on the shortened version and the average of the post-study patient and carer ratings (CP-POST) are reported
Sensibility	*Ease of use:* Of the 235 patients assessed, 182 returned the P-PRE (patient pre-study) form, 223 returned the P-POST (patient post-study) form, 186 carers completed the C-POST (carer post-study) form. One hundred and seventy-four participants completed both the P-POST and C-POST forms; 49 returned the P-POST only and 12 the C-POST only *Skewness:* For CP-POST only the 'psychological' and 'socialising' scales had normal distribution: 'personal care' was skewed to the right; 'self-organisation' and 'mobility' were evenly distributed across the range; and 'productive employment' and 'parent/sib contact' were skewed to the right. For the P-PRE most scales including 'personal care', 'mobility', 'self-organisation', 'psychological' were skewed to the left, 'socialising' was normally distributed
Consistency	*Internal consistency:* were very high for the 'personal care' (0.94), 'mobility' (0.88), 'self-organisation' (0.94) and 'psychological' (0.95) scales, indicating that the items within the scales are highly correlated with one another. The Cronbach's alpha scores were moderate for 'socialising' (0.67), 'parent–sibling contact' (0.70), 'partner–child contact' (0.55) and very low for 'productive employment' (0.30) *Test–retest reliability:* Test–retest reliability was assessed at 1–28 days apart and was based upon 25 P-PRE forms, 23 P-POST and 22 C-POST. For the majority of BICRO-39 subscales the Spearman correlations across the three forms were high (> 0.75) indicating good reliability. For the P-PRE however three of the eight scales did not show good correlation ('self-organisation', 0.53; 'mobility', 0.55; and 'personal care', not done as no variance, $p < 0.01$ for all three)

continued

TABLE 18 contd *BICRO-39*[25]

Validation	Results
Accuracy	Not reported
Suitability	*Construct validity:* BICRO-39 showed significant correlations (n = 95) with the FIM+FAM scales ('personal care', 'mobility', 'self-organisation' and 'psychological'). 'Psychological' also correlated (n = 16) with the HADS and the BICRO-39 correlated (n = 15) with the CIQ as predicted (only weak-to-moderate correlations were predicted) *Discriminant validity:* A subgroup of patients (n = 65) was assessed at intake and then again after a period of treatment (mean = 46.6 weeks, SD = 46.1). Significant improvements occurred in the 'personal care', 'mobility', and 'psychological' scales and 'self-organisation' showed a trend towards improvement ($p < 0.06$)
Patient–proxy and inter-rater agreement	*Patient–proxy agreement:* Spearman correlations showed reasonably good agreement between patients and carers for most of the scales (> 0.67), all correlations being highly significant ($p < 0.001$). Lowest levels of agreement were found for the 'socialising' (0.62) and 'productive employment' (0.63) scales. Wilcoxon's matched pairs signed rank tests showed only a significant difference between patients and carers for the 'psychological' scale, where carers tended to rate problems as very slightly (0.15 scale points) more severely than patients
FIM, Functional Independence Measure; FAM, Functional Assessment Measure	

Conclusion: The BICRO-39 was validated in one study, which included 235 patients with brain injury. Less than 20% of the patients had cognitive impairment and the quality score for the study was high (0.67), indicating that a number of validity tests were carried out. However, the low percentage of people with cognitive impairment in the study, makes it unclear as to whether the findings can be applied to people with cognitive impairment.

TABLE 19 *Modified SIP*[33]

Description: Items particularly relevant to head injuries (i.e. post-concussional symptoms and cognitive impairment) and items that were particularly relevant to this population (such as schoolwork) were added; irrelevant items were excluded on a person-to-person basis and items were re-weighted as judged by a 25-member panel.

Validation	Results
Population studied	84 patients with head injuries, less then 20% of whom were cognitively impaired respondents[33]
Choice of component variables	Not reported
Sensibility	Not reported
Consistency	*Internal consistency:* Cronbach's alpha scores were essentially identical for both the modified and the standard SIP scales at both 1 month and 1 year, ranging from 0.93 to 0.96
Accuracy	*Concurrent validity:* The SIP and the modified SIP were individually significantly related to the severity of neurological injury (Glasgow Coma Scale) and to neuropsychological abilities (Halstead–Reitan Neuropsychological Test Battery and additional measures of attention, speed and memory). However, none of the modifications had a significantly greater correlation than the original unmodified SIP. Individuals were grouped into injury severity groups according to the Glasgow Coma Scale and no modifications significantly improved the ability of SIP to correctly classify the groups The standard SIP and modified SIP were equally correlated with the subjects' self-assessment of overall functioning as measured by a mark on an analogue scale and the Modified Function Status Index. Differences between the two were neither statistically nor practically significant

continued

TABLE 19 contd *Modified SIP*[33]

Validation	Results
Suitability	*Discriminant validity:* The standard and modified SIP scales were excellent discriminators of healthy and injured populations. At 1 month, the standard SIP correctly classified 91%, and the modified SIP 92%. At 1 year these percentages were 78% and 80%, respectively *Predictive ability:* The standard and modified SIP scales were compared in their abilities at 1 month to predict the major role (work, school or homemaking) limitations 1 year after injury. The standard SIP correctly classified 64% and the modified scale 69%
Patient–proxy and inter-rater agreement	Not reported

Conclusion: One study compared the reliability of the SIP with a modified version of the SIP in a study among 84 patients with head injuries.[33] Less then 20% of respondents in this study were cognitively impaired; therefore it is unclear whether the findings apply to people with cognitive impairment.

TABLE 20 *EQ-5D*[13,26,27,37]

Description: The EQ-5D is a generic instrument for the measurement of HRQL. It provides a simple descriptive profile in five dimensions (mobility, self-care, usual activities, pain/discomfort and anxiety/depression), each with three levels. The patient's health state can therefore be classified into 1 of 243 x 3 x 3 x 3 x 3 x 3) theoretically possible health states, each of which has been assigned a utility (i.e. value to the patient). The EQ-5D also includes a VAS on which patients rate their own health between 0 and 100.

Validation	Results
Population studied	Overall 2399 stroke patients in two separate studies.[13,26,27,37] In both studies less than 20% of the respondents were cognitively impaired
Choice of component variables	Not applicable
Sensibility	*Ease of use:* Among 152 stroke patients, 92 patients (61%) completed the questionnaire without help, 54 (35%) were assessed by interview and six (4%) were excluded because of significant difficulties in communication[27] *Response:* Response rates for the SF-36 and EQ-5D questionnaires were compared.[27] Response and 'response with no missing data' were significantly more frequent in patients allocated to the EQ-5D instrument. 66% responded to the first mailing of the EQ-5D (OR = 1.31; 95% CI, 1.1 to 1.6; p = 0.002), 80% to the second mailing (OR = 1.35; 95% CI, 1.1 to 1.6; p = 0.003). When the missing data were taken into account overall 66% of the participants completed the EQ-5D (OR = 1.64; 95% CI, 1.4 to 1.9; $p < 0.0001$). 51% of the EQ-5D questionnaires were completed by the patients rather than by their carers. Respondents to the EQ-5D reported dependency in ADL significantly more than those responding to the SF-36 (58% vs 50%; p = 0.00006)
Consistency	*Test–retest reliability:* For forms completed by the patient.[26] Test–retest reliability was assessed using Kappa statistics. The Kappa statistics for the EQ-5D subscales were 0.85 (95% CI, 0.72 to 0.94) for 'mobility', 0.74 (95% CI, 0.62 to 0.86) for 'self-care', 0.66 (95% CI, 0.54 to 0.78) for 'usual activities', 0.71 (95% CI, 0.59 to 0.83) for 'pain/discomfort', and 0.73 (95% CI, 0.61 to 0.86) for 'anxiety/depression'
Accuracy	Not reported

continued

TABLE 20 contd *EQ-5D*[13,26,27,37]

Validation	Results
Suitability	*Convergent validity:* Increasing dysfunction reported on the EQ-5D scale was associated with lower scores on the standard instruments, which differed significantly from each other ($p < 0.0002$) in all domains. The correlations were as follows: EQ-5D-'mobility' and the Office of Population Censuses and Surveys- 'Locomotion' ($\chi^2 = 54.5$; $p < 0.0001$; $r = 0.61$); EQ-5D-'self-care' and the Barthel Index ($\chi^2 = 65.0$; $p < 0.0001$; $r = –0.64$); EQ-5D-'usual activities' and the Frenchay Activities Index ($\chi^2 = 52.1$; $p < 0.0001$; $r = –0.60$); EQ-5D-'pain/discomfort' and the VAS-'pain' ($\chi^2 = 71.0$; $p < 0.0001$; $r = 0.71$); EQ-5D-'anxiety/depression' and the HADS-'anxiety' ($\chi^2 = 41.3$; $p < 0.0001$; $r = 0.56$); EQ-5D-'anxiety/depression' and the HADS –'depression'($\chi^2 = 16.9$; $p = 0.0002$; $r = 0.35$) Patients' estimates of overall HRQL correlated most closely with the 'depression' subscale of the HADS (–0.54)[13] *Discriminant validity:* Discriminant validity was assessed by comparing outcomes in different groups of stroke patients.[13] For all domains: the worst outcomes were observed in patients with the most extensive corticol strokes. The best outcomes were observed in patients with posterior circulation strokes. In addition, with the exception of the 'psychological functioning' domain, better predicted prognoses were associated with better-reported health status at follow-up The bivariate correlations were also examined between each of the individual domains of the EQ-5D.[13] 'Mobility' correlated best with 'usual activities' ($r = 0.56$) and least well with 'anxiety/depression' ($r = 0.28$). 'Usual activities' correlated best with 'self-care' ($r = 0.61$) and least well with 'anxiety/depression' ($r = 0.35$). 'Self-care' correlated best with 'usual activities' ($r = 0.61$) and least well with 'anxiety/depression' ($r = 0.24$). 'Pain/discomfort' correlated best with 'mobility' ($r = 0.43$) and least well with 'anxiety/depression' ($r = 0.31$). 'Anxiety/depression' correlated best with 'usual activities' ($r = 0.35$) and least well with 'self-care' ($r = 0.24$). Overall HRQL correlated best with 'usual activities' ($r = –0.49$) and least well with 'self-care' ($r = –0.28$)
Patient–proxy and inter-rater agreement	*Patient–proxy agreement:* Test–retest reliability was assessed using Kappa statistics for forms completed by the patient, by proxies, and by both patients and proxies.[26] The Kappa statistics for forms completed by the patient, by proxies, and by both are as follows: 'mobility' – patient 0.85 (0.72 to 0.94), proxy 0.31 (0.00 to 0.66), both 0.80 (0.71 to 0.90); 'self-care' – patient 0.74 (0.62 to 0.86), proxy 0.63 (0.50 to 0.77), both 0.73 (0.65 to 0.81); 'usual activities' – patient 0.66 (0.54 to 0.78), proxy 0.61 (0.47 to 0.76), both 0.68 (0.59 to 0.76); 'pain/discomfort' – patient 0.71 (0.59 to 0.83), proxy 0.61 (0.45 to 0.78), both 0.68 (0.58 to 0.77); 'anxiety/depression' – patient 0.73 (0.61 to 0.86), proxy 0.49 (0.32 to 0.65), both 0.63 (0.53 to 0.73) The percentage of patient–proxy agreement was assessed using Kappa statistics.[37] The results were: 80% agreement, Kappa = 0.60 (95% CI, 0.46 to 0.74) for 'mobility'; 80%, Kappa = 0.60 (95% CI, 0.51 to 0.77) for 'self-care'; 72%, Kappa = 0.56 (95% CI, 0.44 to 0.69) for 'usual activities'; 69%, Kappa = 0.45 (95% CI, 0.30 to 0.59) for 'pain/discomfort'; and 62%, Kappa = 0.30 (95% CI, 0.14 to 0.45) for 'anxiety/depression'. Agreement was better for self-completed forms (67–83%, Kappa, 0.38-0.62) than for interviewed patients (54–84%, Kappa, 0.05-0.62). The mean of the differences between patients' and proxies' estimates of overall HRQL was 2 (95% CI, –38 to 42), which indicated no significant difference between the two. This was confirmed by a factorial analysis of variance

Conclusion: The EQ-5D was validated in two studies including 2399 stroke patients.[13,26,27,37] Less than 20% of the respondents in both studies were cognitively impaired; therefore it is unclear whether the findings apply to people with cognitive impairment.

TABLE 21 *MSQOL-54*[38,39]

Description: The MSQOL-54 is based on the SF-36 with an additional MS-18 module. The SF-36 consists of two composite scores (physical health and mental health) from eight multi-item scales: physical function, role limitation-physical, bodily pain, general health, vitality, social function, role limitation-emotional and mental health) and in total 36 items. The MS-18 module adds 18 additional items concerning: health distress, sexual function, satisfaction with sexual function, overall quality of life, cognitive function and energy. The overall MSQOL-54 consists of 52 items grouped in 12 scales with an additional two lone items.

Validation	Results
Population studied	204 MS patients, less than 20% of whom were cognitively impaired[38,39]
Choice of component variables	Not reported
Sensibility	*Ease of use:* The mean time taken to complete the MSQOL-54 was 19 minutes (range, 5–60)[38] Seventy-eight MS patients (38%) needed help in reading (n = 17), marking the form (n = 9), reading and marking the form (n = 22), or an explanation of the items (n = 30). Eighteen cognitively impaired MS patients in the same study had a mean completion time of 23 minutes (SD = 9) and 67% needed help. The number of missing items ranged from 5.5% to 33.3%, versus 0.5% to 2.9% for non-impaired MS patients. 50% of the 18 cognitively impaired MS patients had no logical inconsistencies in their forms, and 67% had one or less, versus 66% and 82% of non-impaired MS patients The percentage of missing data was low (0.5–2.9%) at the item level, except for items concerning 'sexual function' and 'satisfaction with sexual function', where high percentages of data were missing (both 19%).[39] No pattern in the missing data was identified; however, the missing data were independent of both clinical and demographic characteristics. Only 66% of completed forms had no logical inconsistencies and in 18% of the forms there were two or more inconsistencies. Over 90% of the patients did not consider that the questionnaire was difficult to understand or that it contained embarrassing questions
Consistency	*Internal consistency:* Internal consistency reliability indices were greater than 0.70 for all scales except 'sexual function'[39]
Accuracy	Not reported
Suitability	*Convergent validity:* Within-scale coefficients were homogeneous and higher than 0.40 in most instances, indicating good convergent validity.[39] In the same study the EDSS and the BDI scores were inversely related to the physical health composite score of the MSQOL-54. However the relationship was not statistically significant for EDSS after adjusting for age, clinical worsening and BDI. The BDI score was inversely predictive of the mental health composite score. Depressive symptoms had a major influence on HRQL and patients with higher BDI scores had lower scores in all MSQOL-54 scales, and the BDI score was a statistically significant predictor of low scores in all scales and composite scales *Discriminant validity:* Greater item scale correlations were found within scales than between scales, indicating good discriminant validity[39]
Patient–proxy and inter-rater agreement	*Patient–proxy agreement:* The level of agreement between MS patients and designated proxies was moderate to substantial (0.47–0.75) for all scales, except 'health perceptions' (0.34), 'social functioning' (0.30) and overall QoL (0.34).[38] Proxies under-reported the level of HRQL in all scales except 'cognitive function'. These differences were significant for the 'physical role limitations', 'bodily pain', 'energy' and 'health perceptions' subscales

Conclusion: The MSQOL-54 was validated in one study, which included 204 MS patients, less than 20% of whom were cognitively impaired.[38,39] Therefore it is unclear whether the findings apply to people with cognitive impairment. A separate analysis for 18 cognitively impaired patients was done to establish the ease of use. The mean time to complete the scale for cognitively impaired patients was 23 minutes (SD = 9) compared with 19 minutes for all patients. Overall 67% of cognitively impaired patients needed help compared with 38%. The number of missing items ranged from 5.5% to 33.3%, versus 0.5% to 2.9% for non-impaired MS patients. No logical inconsistencies were found in 50% of the forms completed by cognitively impaired MS patients (n = 18) and 67% had one or less, versus 66% and 82% of non-impaired MS patients.

TABLE 22 *QoL for MS*[29]

Description: The QoL for MS contains five dimensions: self-selected physical problems (five items), mobility (five items), fatigue (four items), control (three items), and emotional upset (seven items).

Validation	Results
Population studied	171 patients with MS, less than 20% of whom were cognitively impaired[29]
Choice of component variables	Not applicable
Sensibility	*Ease of use:* The completion rate for the QoL for MS was 100% for all five scales. The average length of time to complete the questionnaire by telephone interview was 35 minutes (median = 23; range = 9 to 106) *Skewness:* There were no 'floor' or 'ceiling' effects for the QoL for MS. All the scale's scores were in the middle of the 0 to 100 range, and less than 7% scored either minimum or maximum scores
Consistency	*Internal consistency:* The Cronbach's alpha scores were good for all five scales: 'self-selected physical problems' (0.83), 'mobility' (0.90), 'fatigue' (0.87), 'control' (0.85), and 'emotional upset' (0.89) *Test–retest reliability:* Product–moment correlations and ICC were used and found to be acceptable to good for all five scales: 'self-selected physical problems' (Product–moment correlation = 0.72/ICC = 0.72), 'mobility' (0.90/0.89), 'fatigue' (0.80/0.80), 'control' (0.85/0.85), and 'emotional upset' (0.82/0.82)
Accuracy	Not reported
Suitability	*Discriminant validity:* The sensitivity of different measures to important clinical differences was assessed. The clinical indicators included severity of symptoms (none/mild/moderate/extreme), disability (able to walk/requires a walk-aid/requires a wheelchair), days unable to attend work or school in the past month (0 days/1–15 days/16–30 days), and overall QoL (high/middle/low). All of the MS-subscales had satisfactory sensitivities (i.e. sensitivities were > 3), except for the 'emotional upset' scale, which discriminated poorly on level of ambulation
Patient–proxy and inter-rater agreement	Not reported

Conclusion: The QoL for MS scale was validated in one study, which included 171 MS patients.[29] Less than 20% of the patients had cognitive impairment, and the overall quality score was 0.4. The low percentage of people with cognitive impairment makes it unclear as to whether the findings apply to people with cognitive impairment.

TABLE 23 *PCRS*[58]

Description: Both the person with TBI and a significant other rate the ease with which the person with TBI is able to perform functional activities (including 'daily living', 'interpersonal skills', 'cognitive and memory function' and 'emotional status'). A five-point Likert scale is used to rate how difficult or easy it is to perform 30 different behavioural tasks. The scales range from 'can't do' (point 1) to 'can do with ease' (point 5). The assessor compares the person's and the significant other's perspective to obtain a measure of self-awareness.

Validation	Results
Population studied	53 patients with TBI.[58] Less than 20% of the participants had cognitive impairment
Choice of component variables	Not applicable
Sensibility	Not reported
Consistency	Internal consistency: Cronbach's alpha scores indicated that there were strong internal consistencies for the patients (0.91, *n* = 55), relatives (0.93, *n* = 50) and the therapists (0.95, *n* = 47)

continued

TABLE 23 contd *PCRS*[58]

Validation	Results
	Test–retest reliability: Retest data were not collected for the relatives' version of the PCRS. High test–retest reliabilities were reported for both the patients (0.85, n = 20) and the therapists (0.95, n = 23)
Accuracy	Not reported
Suitability	Not reported
Patient–proxy and inter-rater agreement	*Patient–proxy agreement:* At 3 months post-injury, significant differences ($p < 0.0017$) at the corrected alpha level (Bonferroni correction) between subject and informant responses were reported for 15/30 of the PCRS items. These included 'preparing meals', 'managing finances', 'keeping appointments', 'work activities', 'remembering important things', 'driving a car', 'adjusting to changes', 'handling arguments', 'showing affection', 'recognising if upset', 'scheduling activities', 'understanding new instructions', 'coping with daily responsibilities', 'emotions and activity, and 'controlling laughter'. In all cases the difference was in the direction of participants overestimating their level of competency compared with the proxies. For each item a small number of participants underestimated their level of competency compared with proxies. At 12 months' post-injury, three of the differences between patient and proxy ratings were still significant (Bonferroni correction, $p < 0.0017$). These were 'managing finances', 'driving a car' and 'recognising if upset'

Conclusion: The PCRS was validated in one study, which included 53 TBI patients.[58] Less than 20% of the patients had cognitive impairment, and the quality score for the study was 0.4. The low percentage of people with cognitive impairment makes it unclear whether the findings apply to people with cognitive impairment.

TABLE 24 *QOL-Interview*[56,57]
Description: The QoL-Interview consists of eight objective scales: living situation, frequency of family contacts, frequency of social contacts, number of leisure activities, work, frequency of religious activities, finances, safety and health). The same eight factors are examined in eight subjective scales, which measure general life satisfaction.

Validation	Results
Population studied	232 patients with LD.[56,57] Less than 20% of the participants were cognitively impaired
Choice of component variables	Not applicable
Sensibility	Not reported
Consistency	*Internal consistency:* The scale measuring 'satisfaction with social relations' reliably measured a main overall factor (0.70) as well as the subfactors 'satisfaction with relations within the treatment facility' (0.67) and 'satisfaction with relations outside of the facility' (0.68).[57] Internal consistencies for objective life conditions and life satisfaction, ranged from 0.5 to 0.956 *Test–retest reliabilities:* Ranged from 0.5 to 0.9 for objective life conditions and life satisfaction[56]
Accuracy	Not reported

continued

TABLE 24 contd *QOL-Interview*[56,57]

Validation	Results
Suitability	*Convergent validity:* Among patients with different disorders (mental retardation, schizophrenia, affective disorder, alcoholism, drug abuse, organic brain syndrome or personality disorder) three sets of correlations were compared among the variables.[57] The domain-specific subjective QoL measures correlated most strongly with general life satisfaction, objective QoL measures correlated less strongly with life satisfaction, and demographic and diagnostic variables correlated least strongly with life satisfaction. The dimensions of 'psychopathology', 'depression', 'anxiety', and 'thought disorder', which were assessed concomitantly with 'general life satisfaction', 'depression' and 'anxiety'. These variables consistently showed significant, negative correlations with 'general life satisfaction' across the various patient populations (e.g. 'depression' $r = -0.17$ to -0.56, $p < 0.05$ to < 0.001 and 'anxiety' $r = -0.25$ to -0.33, $p < 0.001$ to < 0.0001). 'Thought disorder' did not correlate with 'life satisfaction' ($r = 0.06$ to -0.14). Therefore concomitant assessment of a respondent's level of psychiatric symptoms, especially depression and anxiety, seemed advisable in this population *Discriminant validity:* Only one subscale ('physical comfort of the living situation') was able to discriminate between four groups of participants (in-patients, community residents and current length-of-stay more or less than 6 months). Using this subscale 58% of the participants were classified into the correct patient groups[56]
Patient–proxy and inter-rater agreement	Not reported

Conclusion: The Qol-Interview was validated in two studies which included 232 patients with LD.[56,57] In both studies less than 20% of the respondents were cognitively impaired; therefore it is unclear whether the findings apply to people with cognitive impairment.

TABLE 25 *Quality of Life-Index MS-Version*[52,53]

Description: Originally the QoL-Index was developed to measure QoL of healthy persons. The QoL-Index (MS-version) consists of 72 items composed of two parts: part 1 measures satisfaction with various domains of life (e.g. health, being able to get around, standard of living, and achieving personal goals), and part 2 measures the importance of the same domains. Items are assessed on a six-point scale ranging from 'very satisfied' to 'very dissatisfied' for part 1 and 'very important' to 'very unimportant' for part 2. Total QoL scores are calculated by weighting each satisfaction response with its paired importance response. High scores are produced by a combination of high importance and high satisfaction.

Validation	Results
Population studied	61 MS patients, less than 20% of whom were cognitively impaired[52,53]
Choice of component variables	Not applicable
Sensibility	Not reported
Consistency	*Internal consistency:* The internal consistency was reported as 0.87 in both publications
Accuracy	Not reported

continued

TABLE 25 contd *Quality of Life-Index MS-Version*[52,53]

Validation	Results
Suitability	*Correlations with disease factors:* The number of cognitive tests that participants failed was not related to their QoL (QoL-Index, $r = 0.08$); however, perceived disturbances of cognition were significantly negatively related to perceived QoL ($r = –0.28$; $p < 0.05$). The following correlations between the QoL-Index and demographic and disease factors were assessed: age ($r = 0.11$), gender ($r = 0.37$; $p < 0.01$), functional disability ($r = –0.48$; $p < 0.01$) and length of illness ($r = 0.06$). Furthermore correlations with financial situation ($r = 0.45$; $p < 0.01$), education ($r = 0.05$), social support ($r = 0.52$; $p < 0.01$), reciprocity ($r = 0.50$; $p < 0.01$) and conflict ($r = –0.28$; $p < 0.05$) were assessed, as well as correlations with perceptual factors: perceived health ($r = 0.57$; $p < 0.01$), self rated abilities ($r = 0.62$; $p < 0.01$), general self-efficacy ($r = 0.41$; $p < 0.01$), demands of illness ($r = –0.48$; $p < 0.01$) and health promoting lifestyle ($r = 0.51$; $p < 0.01$)
Patient–proxy and inter-rater agreement	Not reported

Conclusion: The QoL-Index was validated in one study, which included 61 MS patients, less than 20% of whom were cognitively impaired. Therefore it is unclear whether the findings apply to people with cognitive impairment.

TABLE 26 *Life Satisfaction (Viitanen's Quality of Life Questionnaire)*[20,54]
Description: The life satisfaction scale comprises one global and six domain-specific (self-care ADL, marriage/courtship, sexuality, leisure, togetherness/friends, and togetherness/family) life satisfaction items.

Validation	Results
Population studied	19 patients with brain injury.[20,54] Less than 20% of the participants were cognitively impaired
Choice of component variables	Not applicable
Sensibility	Not reported
Consistency	*Split-half reliability:* One of the studies reported a split-half reliability of 0.92, in 15 patients with brain injury[20]
Accuracy	Not reported
Suitability	*Correlations with functional status:* Correlations were assessed between functional status and QoL.[20] The following correlation coefficients were reported for neurological status (–0.13), physiotherapy (0.22), Barthel Index (0.30), pulses profile, summed score (–0.17), speech/reception (0.72; $p < 0.01$), speech/production (0.71; $p < 0.01$), speech/intelligibility (0.70; $p < 0.01$), verbal IQ (0.49; $p < 0.05$), performance IQ (0.22), whole-scale IQ (0.37) and Wechsler Memory Scale Memory Quotient (0.22)[54]
Patient–proxy and inter-rater agreement	Not reported

Conclusion: Validity was assessed in one study, which included 19 brain-injured patients, less than 20% of whom were cognitively impaired.[20,54] Therefore it is unclear whether the findings apply to people with cognitive impairment.

TABLE 27 *Ho's QoL Scale*[59]

Description: Ho's QoL Scale consists of a nine-point Likert scale composed of 32 items subsumed under six components (health and illness, human relationships and support, living, activities and interests, psychological well-being, philosophical, spiritual, religious well-being, fulfilment, completion of tasks), which have been conceptualised to represent the essence of the Chinese view of QoL.

Validation	Results
Population studied	109 patients with stroke.[59] Less than 20% of the respondents had cognitive impairment
Choice of component variables	Not applicable
Sensibility	Not reported
Consistency	*Internal consistency:* Cronbach's alpha = 0.933
Accuracy	Not reported
Suitability	Not reported
Patient–proxy and inter-rater agreement	Not reported

Conclusion: Ho's QoL Scale was validated in one study, which included 109 patients with stroke. Less than 20% of the respondents had cognitive impairment, and the quality score for the study was only 0.1, showing that the level of validation was limited.

Appendix 7

Quality of included studies

TABLE 28 *Scale validations in populations where more than 50% of the respondents were cognitively impaired*

Scale	Study	Components	Sensibility	Consistency	Accuracy	Suitability	Patient–proxy/interrater	Total score
ComQol-I5	Cummins *et al.*, 1997[21,22]	0	0	2	0	2	2	0.5 (6/12)
SF-36	Jones *et al.*, 1997[41]	–	0	2	0	0	1	0.3 (3/10)
LSI	Hawkins *et al.*, 1995[44]	–	0	1	0	2	0	0.3 (3/10)
LSS	Vandergriff & Chubon, 1994[42]	–	0	1	0	2	0	0.3 (3/10)
CQOL	Graham *et al.*, 1997[45]	0	0	2	0	1	0	0.25 (3/12)
QUOLIS	Ouellette Kuntz, 1990[46]	0	0	1	0	0	2	0.25 (3/12)
VAS-QoL	Kreuter *et al.*, 1998[43]	–	0	0	0	2	0	0.2 (2/10)
SWB	Ormel *et al.*, 1998[19]	–	0	1	0	0	0	0.1 (1/10)
Life Satisfaction (Cantril)	Ormel *et al.*, 1998[19]	–	0	1	0	0	0	0.1 (1/10)

TABLE 29 *Scale validations in populations where 20–50% of the respondents were cognitively impaired*

Scale	Study	Components	Sensibility	Consistency	Accuracy	Suitability	Patient–proxy/interrater	Total score
SIP	Fischer *et al.*, 1999[23]	–	0	1	0	2	0	0.3 (3/10)
	Smith, 1992[24]	–	1	0	0	1	2	0.4 (4/10)
	Sneeuw *et al.*, 1997[14]	–	0	1	0	1	2	0.4 (4/10)
	Total score	–	1	1	0	2	2	0.6 (6/10)
MS-QLI	Fischer *et al.*, 1999[23]	2	1	1	0	2	0	0.5 (6/12)
SF-36	Fischer *et al.*, 1999[23]	–	0	1	0	2	0	0.3 (3/10)
	Wilkinson *et al.*, 1997[47]	–	0	0	0	1	0	0.1 (1/10)
	Total score	–	0	1	0	2	0	0.3 (3/10)
COOP/WONCA	Lam *et al.*, 1994[49]	–	2	0	0	0	0	0.2 (2/10)

continued

TABLE 29 contd *Scale validations in populations where 20–50% of the respondents were cognitively impaired*

Scale	Study	Components	Sensibility	Consistency	Accuracy	Suitability	Patient–proxy/interrater	Total score
Oregon QLQ	Turner, 1997[50]	–	0	1	0	0	1	0.2 (2/10)
NHP	Harwood *et al.*, 1994[48]	–	0	0	0	1	0	0.1 (1/10)
	Wilkinson *et al.*, 1997[47]	–	0	0	0	1	0	0.1 (1/10)
	Total score	–	0	0	0	1	0	0.1 (1/10)

TABLE 30 *Scale validations in populations where less than 20% of the respondents were cognitively impaired*

Scale	Study	Components	Sensibility	Consistency	Accuracy	Suitability	Patient–proxy/interrater	Total score
SF-36	Anderson *et al.*, 1996[15]	–	1	1	0	2	0	0.4 (4/10)
	Dorman *et al.*, 1997,[27] 1998[26]	–	1	2	0	0	1	0.4 (4/10)
	Pickard, 1999[28]	–	0	0	2	1	1	0.4 (4/10)
	Vickrey *et al.*, 1997[29]	–	1	2	1	0	0	0.4 (4/10)
	Total score	–	1	2	2	2	1	0.8 (8/10)
SIP	Baird *et al.*, 1987[31]	–	0	0	0	2	0	0.2 (2/10)
	Burton & Volpe, 1994[32]	–	0	0	0	2	0	0.2 (2/10)
	Dikmen *et al.*, 1995[30]	–	0	0	2	1	0	0.3 (3/10)
	Smith *et al.*, 1998[35]	–	0	1	0	1	0	0.2 (2/10)
	Visser *et al.*, 1995[34]	–	1	1	0	0	0	0.2 (2/10)
	Temkin *et al.*, 1988[33]	0	0	1	1	2	0	0.33 (4/12)
	Zeldow & Pavlou, 1988[36]	–	0	0	0	2	0	0.2 (2/10)
	Total score	–	1	2	2	2	0	0.7 (7/10)
Modified SIP	Temkin *et al.*, 1988[33]	0	0	1	1	2	0	0.33 (4/12)
BICRO-39	Powell *et al.*, 1998[25]	1	1	2	0	2	2	0.67 (8/12)
EQ-5D	Dorman *et al.*, 1997[13,37]	–	0	0	0	2	2	0.4 (4/10)
	Dorman *et al.*, 1997,[27] 1998[26]	–	1	1	0	0	2	0.4 (4/10)
	Total score	–	1	1	0	2	2	0.6 (6/10)
LSS	Gulick, 1997[40]	–	0	2	1	2	0	0.5 (5/10)

continued

TABLE 30 contd *Scale validations in populations where less than 20% of the respondents were cognitively impaired*

Scale	Study	Components	Sensibility	Consistency	Accuracy	Suitability	Patient–proxy/interrater	Total score
MSQOL-54	Solari *et al.*, 1998,[38] 1999[39]	0	2	1	0	2	1	0.5 (6/12)
QoL for MS	Vickrey *et al.*, 1997[29]	–	1	2	0	1	0	0.4 (4/10)
PCRS	Fleming & Strong, 1999[58]	–	0	2	0	0	2	0.4 (4/10)
QoL Interview	Lehman *et al.*, 1986[56]	–	0	2	0	1	0	0.3 (3/10)
	Lehman, 1988[57]	0	0	1	0	2	0	0.25 (3/12)
	Total score	0	0	2	0	2	0	0.4 (4/10)
NHP	Gompertz *et al.*, 1994[51]	–	0	0	0	2	0	0.2 (2/10)
	Visser *et al.*, 1995[34]	–	1	1	0	0	0	0.2 (2/10)
	Total score	–	1	1	0	2	0	0.4 (4/10)
QoL-Index MS	Stuifbergen, 1995[52,53]	–	0	1	0	2	0	0.3 (3/10)
Life Sat (Viitanen)	Koskinen, 1998,[20] Kaitaro *et al.*, 1995[54]	–	0	1	0	2	0	0.3 (3/10)
VAS-QoL	Ahlsio *et al.*,1984[55]	–	0	0	0	1	0	0.1 (1/10)
Ho's QoL Scale	Chow, 1997[59]	–	0	1	0	0	0	0.1 (1/10)

Appendix 8

Information for acquiring instruments

- Comprehensive Quality of Life scale – Intellectual disability (ComQol-I5)
  Robert A Cummins, Ph.D., F.A.P.S.,
  School of Psychology, Deakin University,
  221 Burwood Highway, Melbourne Victoria 3125, Australia.
  Phone: +61 3 9244 6845; Fax: +61 3 9244 6858;
  E-mail: cummins@deakin.edu.au

- Medical Outcome Studies-Short Form-36 (SF-36)
  Webpage: http://www.sf-36.com/

- Life Satisfaction Index (LSI)
  Barbara A Hawkins, ReD, Associate Professor, School of HPER, HPER Building, Indiana University, Bloomington, IN 47405-4801

- Life Situation Survey (LSS)
  Robert A Chubon, Ph.D.,
  University of South Carolina, School of Medicine, Department of Neuropsychiatry and Behavioural Science, Rehabilitation Counseling Program, 3555 Harden Street Ext., Columbia, SC 29203, USA.
  Phone: +1 803 434 4296; E-mail: rchubon@npsy.ceb.sc.edu
  Webpage: http://www.med.sc.edu:94/lss.htm

- Child Quality of Life questionnaire (CQOL)
  Professor Philip Graham, 27 St Albans Road, London NW5 1RG, UK

- Quality of Life scale (QUOLIS)
  Helene Ouellette Kuntz, Department of Community Health and Epidemiology, Queen's University at Kingston, Kingston, Ontario, Canada, K7L 3N6

- Sickness Impact Profile (SIP)
  Johns Hopkins University, 8 Park Plaza #503, Boston, MA 02116-4313, USA (Distributed by the Medical Outcomes Trust).
  Phone: +1 617 426 4046; Fax +1 617 426 4131;
  E-mail: MOTrust@worldnet.att
  Webpage: http://www.outcomes-trust.org/

- Multiple Sclerosis Quality of Life Interview (MS-QLI)
  Carol Estwing Ferrans, PhD, RN, FAAN, Associate Professor, The University of Illinois at Chicago, Department of Medical-Surgery Nursing (M/C 802), College of Nursing, 845 South Damen Avenue, 7th Floor, Chicago, Illinois 60612-7350.
  Phone: +1 312 996 7900; Fax: +1 312 996 4979;
  E-mail: cferrans@uic.edu

Methodology Group

Members

Methodology Programme Director
Professor Richard Lilford
Director of Research and Development
NHS Executive – West Midlands, Birmingham

Chair
Professor Martin Buxton
Director, Health Economics Research Group
Brunel University, Uxbridge

Professor Douglas Altman
Professor of Statistics in Medicine
University of Oxford

Dr David Armstrong
Reader in Sociology as Applied to Medicine
King's College, London

Professor Nicholas Black
Professor of Health Services Research
London School of Hygiene & Tropical Medicine

Professor Ann Bowling
Professor of Health Services Research
University College London Medical School

Professor David Chadwick
Professor of Neurology
The Walton Centre for Neurology & Neurosurgery
Liverpool

Dr Mike Clarke
Associate Director (Research)
UK Cochrane Centre, Oxford

Professor Paul Dieppe
Director, MRC Health Services Research Centre
University of Bristol

Professor Michael Drummond
Director, Centre for Health Economics
University of York

Dr Vikki Entwistle
Senior Research Fellow, Health Services Research Unit
University of Aberdeen

Professor Ewan B Ferlie
Professor of Public Services Management
Imperial College, London

Professor Ray Fitzpatrick
Professor of Public Health & Primary Care
University of Oxford

Dr Naomi Fulop
Deputy Director, Service Delivery & Organisation Programme
London School of Hygiene & Tropical Medicine

Mrs Jenny Griffin
Head, Policy Research Programme
Department of Health
London

Professor Jeremy Grimshaw
Programme Director
Health Services Research Unit
University of Aberdeen

Professor Stephen Harrison
Professor of Social Policy
University of Manchester

Mr John Henderson
Economic Advisor
Department of Health, London

Professor Theresa Marteau
Director, Psychology & Genetics Research Group
Guy's, King's & St Thomas's School of Medicine, London

Dr Henry McQuay
Clinical Reader in Pain Relief
University of Oxford

Dr Nick Payne
Consultant Senior Lecturer in Public Health Medicine
ScHARR
University of Sheffield

Professor Joy Townsend
Director, Centre for Research in Primary & Community Care
University of Hertfordshire

Professor Kent Woods
Director, NHS HTA Programme, & Professor of Therapeutics
University of Leicester

HTA Commissioning Board

Members

Programme Director
Professor Kent Woods
Director, NHS HTA Programme, &
Professor of Therapeutics
University of Leicester

Chair
Professor Shah Ebrahim
Professor of Epidemiology of Ageing
University of Bristol

Deputy Chair
Professor Jon Nicholl
Director, Medical Care Research Unit
University of Sheffield

Professor Douglas Altman
Director, ICRF Medical Statistics Group
University of Oxford

Professor John Bond
Director, Centre for Health Services Research
University of Newcastle-upon-Tyne

Ms Christine Clark
Freelance Medical Writer
Bury, Lancs

Professor Martin Eccles
Professor of Clinical Effectiveness
University of Newcastle-upon-Tyne

Dr Andrew Farmer
General Practitioner &
NHS R&D Clinical Scientist
Institute of Health Sciences
University of Oxford

Professor Adrian Grant
Director, Health Services Research Unit
University of Aberdeen

Dr Alastair Gray
Director, Health Economics Research Centre
Institute of Health Sciences
University of Oxford

Professor Mark Haggard
Director, MRC Institute of Hearing Research
University of Nottingham

Professor Jenny Hewison
Senior Lecturer
School of Psychology
University of Leeds

Professor Alison Kitson
Director, Royal College of Nursing Institute, London

Dr Donna Lamping
Head, Health Services Research Unit
London School of Hygiene & Tropical Medicine

Professor David Neal
Professor of Surgery
University of Newcastle-upon-Tyne

Professor Gillian Parker
Nuffield Professor of Community Care
University of Leicester

Dr Tim Peters
Reader in Medical Statistics
University of Bristol

Professor Martin Severs
Professor in Elderly Health Care
University of Portsmouth

Dr Sarah Stewart-Brown
Director, Health Services Research Unit
University of Oxford

Professor Ala Szczepura
Director, Centre for Health Services Studies
University of Warwick

Dr Gillian Vivian
Consultant in Nuclear Medicine & Radiology
Royal Cornwall Hospitals Trust
Truro

Professor Graham Watt
Department of General Practice
University of Glasgow

Dr Jeremy Wyatt
Senior Fellow
Health Knowledge Management Centre
University College London

Current and past membership details of all HTA 'committees' are available from the HTA website (see inside front cover for details)

Feedback

The HTA programme and the authors would like to know your views about this report.

The Correspondence Page on the HTA website (http://www.ncchta.org) is a convenient way to publish your comments. If you prefer, you can send your comments to the address below, telling us whether you would like us to transfer them to the website.

We look forward to hearing from you.

Copies of this report can be obtained from:

The National Coordinating Centre for Health Technology Assessment,
Mailpoint 728, Boldrewood,
University of Southampton,
Southampton, SO16 7PX, UK.
Fax: +44 (0) 23 8059 5639 Email: hta@soton.ac.uk
http://www.ncchta.org

ISSN 1366-5278