

**Patient-reported
Health Instruments
Group**

**Instruments for
Diabetes:
a review**

**Andrew M. Garratt
Louise J. Schmidt
Anne E. Mackintosh
Ray Fitzpatrick**

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



health

Outcomes

indicators

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

INSTRUMENTS FOR DIABETES: A REVIEW

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

Andrew M. Garratt
Louise J. Schmidt
Anne E. Mackintosh
Ray Fitzpatrick

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

July 2000

This report should be referenced as follows:

Garratt AM, Schmidt LJ, Mackintosh AE, Fitzpatrick R. *Instruments for Diabetes: a Review Report* from the Patient-reported Health Instruments Group (formerly the Patient-Assessed Health Outcomes Programme) to the Department of Health, July 2000.

Copies of this report can be obtained from:

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

tel: +44 (0)1865 227157
e-mail: kirstie.haywood@uhce.ox.ac.uk

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

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

LIST OF CONTENTS

Frontispiece	1
List of contents	2
Executive Summary	3
Chapter 1: INTRODUCTION	6
a) Diabetes	6
b) Patient-reported health outcome measures	8
Chapter 2: METHODS	12
a) Search strategy	12
b) Inclusion criteria	13
c) Data extraction	14
Chapter 3: RESULTS	18
a) Search outcome	18
b) Patient characteristics	20
c) Instrument content	22
d) Instrument development	31
e) Instrument evaluation	33
Chapter 4: INSTRUMENT REVIEWS	35
Chapter 5: CONCLUSIONS	99
Reference list	102
Appendix: measures excluded from the review	111

Executive Summary

Background

This report presents a review of diabetes-specific measures of symptoms and health-related quality of life. There has been considerable research into the effects of diabetes on quality of life, yielding numerous instruments designed for use as health outcome measures. This review will provide users of such measures with information guiding the selection and evaluation of these instruments for future application, such as in clinical trials.

Research Aims

1. to identify patient-reported health outcome measures specific to diabetes;
2. to extract and assess the evidence relating to the development and evaluation of the instruments, and make recommendations as to their application.

Methods

Electronic databases from 1990-1999 were searched using keywords relevant to the development and testing of instruments specific to diabetes. Several other sources, including conference abstracts, were also searched. The names of instruments were then used in a second search strategy. Instrument authors were sent a letter requesting additional information relating to development, evaluation and scoring.

After receiving replies from authors and retrieving published papers, the following information was extracted relating to the development and evaluation of the instruments:

- the purpose of the instrument, including the underlying phenomena being measured and the proposed application;
- instrument development and scoring;
- patient samples in which the instrument was developed and tested;
- measurement properties of reliability, validity and responsiveness;
- response rates and missing data.

Key Findings

- a) Twenty instruments met the inclusion criteria for the review. The majority were developed and evaluated in Europe or the United States. Two of the instruments, namely Quality of Life: Status and Change (Hornquist et al., 1993) and the Well-being Questionnaire (Bradley, 1994), are predominantly generic in focus but were included in the review because of their widespread use in patients with diabetes. All but four of the instruments have been evaluated in patients with both type I and type II diabetes.

- b) The content of the instruments covers ten quality of life or symptom domains but was largely concentrated in the domains of diabetic symptoms, global health and quality of life, personal constructs, physical functioning, psychological well-being, social well-being, and treatment and quality of life.
- c) Most of the disease-specific instruments have one published evaluation relating to the measurement properties of reliability and validity. Very few instruments have undergone a formal evaluation of responsiveness to change.
- d) The Diabetes Quality of Life Measure (Jacobson, 1997) has the largest number of published evaluations, and has evidence for its reliability and validity across four nationalities of diabetic patients.
- e) Three of the strictly disease-specific instruments have been developed and evaluated in the UK, namely the Audit of Diabetes Dependent Quality of Life (Bradley, Todd et al.1999), the Diabetes Health Profile (Meadows, Steen et al., 1996), and the Newcastle Diabetes Symptom Questionnaire (McColl, Steen et al., 1995).
- f) The disease-specific instruments developed in the UK have not been as extensively evaluated as the Diabetes Quality of Life Measure but have undergone a more rigorous approach to development. The evidence for the measurement properties of the three instruments compares favourably with that for the other instruments reviewed.
- g) The three UK instruments are based on different approaches to the measurement of patient-reported health outcomes. The Diabetes Health Profile produces a profile of diabetes-specific health, and is based on the traditional psychometric approach to instrument development. The Audit of Diabetes Dependent Quality of Life incorporates an individualised importance weighting for each of 13 items. The Newcastle Diabetes Symptoms Questionnaire is a more narrowly focused measure of diabetes symptoms.

Key conclusions and recommendations

1. Three approaches to the measurement of patient-reported health outcomes should be considered for use in diabetes: generic instruments, disease-specific instruments and situation-specific instruments. Generic instruments are useful for comparisons across populations and have particular relevance to economic evaluation. Disease-specific instruments take account of the specific effects of diabetes on quality of life, and are likely to be more responsive to change than generic instruments. Situation-specific instruments are designed to measure the effects of specific interventions on knowledge, attitudes and behaviour.
2. Two instruments developed within the UK, namely the Diabetes Health Profile and the Audit of Diabetes Dependent Quality of Life, are recommended as potential primary outcome measures for clinical and health services research, including clinical trials. Where a clinically relevant and responsive measure of

symptoms is required, it is recommended that the nine-item Newcastle Diabetes Symptoms Questionnaire be used alongside these instruments.

3. Where practical, the disease-specific instruments should be used alongside a validated generic instrument. The Well-being Questionnaire is recommended if a detailed evaluation of psychological well-being is required.
4. None of the instruments included in the review has the recommended levels of reliability for use in individual patients. Where an instrument is required for clinical practice at the group level, for example within clinical audit, it is recommended that the content of the Audit of Diabetes Dependent Quality of Life, the Diabetes Health Profile and the Newcastle Diabetes Symptoms Questionnaire is assessed for relevance.
5. Future quality of life research on diabetes should concentrate on evaluating and, where appropriate, refining existing instruments rather than developing new ones. There is a considerable body of literature relating to the quality of life of diabetes patients to support this work. The Audit of Diabetes Dependent Quality of Life, the Diabetes Health Profile and the Newcastle instrument require further testing for reliability, validity and responsiveness. In particular the instruments should be assessed concurrently for responsiveness to change. This research will inform decisions regarding the selection of instruments for future application.

Chapter 1: INTRODUCTION

a) Diabetes

Diabetes is a group of disorders with common features, the presence of raised blood glucose being the most evident. The three disorders in this group which have the largest impact on health in the UK are:

- Insulin-dependent diabetes mellitus (IDDM) or type 1 diabetes
- Non-insulin dependent diabetes mellitus (NIDDM) or type 2 diabetes
- Gestational diabetes (or diabetes of pregnancy)

This report focuses on the first two types of diabetes and, in common with others (Home, Coles et al., 1999), excludes the specific consideration of pregnancy associated diabetes. Gestational diabetes can be regarded as an extreme form of impaired glucose tolerance in some pregnant women, in response to the metabolic strain of pregnancy.

The prevalence of clinically diagnosed diabetes of all types in England is often estimated to be around 2%, but the true figure is likely to be higher since many patients, particularly those with NIDDM, are undiagnosed. The prevalence of diabetes is higher among lower socio-economic groups and among ethnic minorities (ibid.).

The onset of IDDM often occurs during childhood or adolescence and is associated with a genetic predisposition to the disease, which is ignited by certain as yet unknown environmental factors. NIDDM tends to develop during adulthood but also seems to have a genetic component, as evidenced by the frequent finding in new patients of a family history of the disease.

IDDM can be managed by balancing injected insulin, carbohydrate intake, and energy expenditure. Whereas two injections a day were commonly used by IDDM patients in the 1980s and early 1990s, recent evidence has suggested that more intensive therapy (up to five injections per day) is associated with reduced risk of chronic complications. There is, however, an increased risk of severe hypoglycaemic episodes with this regime (Bradley, Pierce et al., 1998).

People with NIDDM produce insulin, but the amount is usually insufficient to maintain blood glucose levels within their normal range. Dietary management is usually the first course of treatment for NIDDM, supplemented with oral hypoglycaemic agents where necessary. Should neither of these treatments prove effective, insulin is normally recommended (ibid.).

NIDDM has been referred to as “mild” diabetes but the consequences can be no less serious than for patients with IDDM. People with NIDDM are two or three times more likely to develop cardiovascular disease than people without diabetes, and 75% die of cardiovascular disease. The life expectancy of people with NIDDM is, on average, reduced by a decade (ibid.).

Having diabetes, as with many other chronic illnesses, can have a profound effect on patients’ quality of life in terms of social and psychological well-being, as well as

physical ill-health. Diabetes is one of the most psychologically demanding of the chronic diseases, with psychosocial factors pertinent to nearly every aspect of the disease and its treatment (Cox, Blount et al., 1996). In a longitudinal study, the psychosocial impact of diabetes was found to be one of the five best predictors of mortality in diabetic patients – more reliable than many clinical and physiological variables (Davis, Hess et al., 1987).

Following diagnosis, many patients with diabetes experience psychological problems including social withdrawal, depression and anxiety. As the disease progresses, psychosocial problems often occur secondary to the onset of complications (including neuropathy, retinopathy and nephropathy), although several studies have reported an increased prevalence of depression and anxiety among patients with IDDM and NIDDM irrespective of the presence of complications or loss of function (Cox and Gonder-Frederick, 1992). Depression among patients can in turn impact negatively on diabetes management, in terms of its association with poorer glucose regulation and decreased adherence to the treatment regime.

However, it is not only the disease of diabetes which can have a detrimental effect on quality of life but also the treatment itself. The control of symptoms of high blood sugar (dizziness, polyuria, weight loss), which left untreated can lead to coma and death, imposes restrictions on patients' quality of life as it often involves a prescriptive routine of diet, exercise, self-monitoring of blood and self-medication (Hanestad and Albrektsen, 1991). In fact, increasing interest in the measurement of quality of life has followed the recognition that improved glycaemic control can impact negatively on social and emotional well-being (Eiser and Tooke, 1995).

The adoption of such a strict treatment regime has implications for many aspects of life including employment, home life and social life. Besides the effect on psychological well-being, diabetes can also lead to the patient feeling stigmatised by the disease and seeking ways to avoid others becoming aware of its existence. This again can have negative consequences for the control of diabetes and prevention of long-term complications, by reducing adherence to the treatment regime.

Under the auspices of the World Health Organisation and the International Diabetes Federation, representatives of Government Health Departments and patient organisations from all European countries met and agreed the Saint Vincent Declaration in 1989. One of the general goals was sustained improvement in health experience, and a life approaching normal expectations of quality and quantity. These goals create the need for valid and reliable instruments measuring the quality of life of people with diabetes. Without such instruments, it is not possible to ascertain whether steps taken to improve quality of life actually achieve this result.

It is increasingly recognised that instruments taking account of patients' concerns can complement traditional biomedical measures of outcome (Jenkinson, 1995). The use of such instruments is particularly important in chronic conditions such as diabetes, where a major objective of management is to arrest or reverse decline in function and quality of life. There is evidence that assessments made by patients of their health differ from those made on their behalf by health professionals (Slevin et al., 1988; Woodend, Nair et al., 1997). Therefore, valid measures of health outcomes and

quality of life should be patient-reported; this can take the form of self-completed or interviewer-administered questionnaires.

b) Patient-reported health outcome measures

Patient-reported health outcome measures can be either generic (i.e. they can be used for different diseases, as well as in healthy populations) or disease-specific (developed for use in a specific patient group). Generic instruments are designed to measure aspects of health which are of universal importance, and are therefore suitable for comparisons between different groups of patients or with healthy populations. They can take account of the influence of comorbidity on health and have the potential to capture the side-effects of an intervention. This makes them potentially useful for assessing the impact of new health-care technologies when the therapeutic effects are uncertain.

Specific instruments can be selected reflecting the areas considered by patients or clinicians to be of greatest importance, and can be specific to a particular disease, population, function, condition, or problem. The narrow focus of specific instruments has the potential to make them more responsive to changes in health. The possible complementary role played by generic and specific instruments has led to recommendations that the two approaches be used in conjunction.

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

Health profiles measure health across a number of distinct 'dimensions' such as physical functioning, mental health and role limitations. The items within health profiles are scored and summed to reflect individual dimensions but sometimes also produce a single index. The Diabetes Impact Measurement Scales is a diabetes-specific health profile measuring health across five dimensions including symptoms (those specific to and those less specific to diabetes), diabetes-related morale, social role fulfilment, and well-being (Hammond and Aoki, 1992). These dimensions also sum to produce a DIMS total score.

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

Before an instrument can be recommended for application, the measurement properties of reliability, validity and responsiveness should be assessed. The data quality, scaling assumptions and dimensionality of the instrument should be assessed concurrently. Individual items should be assessed for levels of missing data and response frequencies. Items with relatively large amounts of missing data should be removed from the instrument. Items with large end effects at the floor or ceiling are poor discriminators and should be removed from the instrument.

In deriving items, instrument authors often hypothesise dimensions that may be based on their own experience or theory. These dimensions can be assessed empirically through the statistical techniques of factor analysis and principal component analysis (PCA), which is sometimes referred to as internal validation. PCA groups together items measuring the same underlying construct, and provides evidence for the dimensionality of an instrument. Item-total correlation is usually undertaken after the application of PCA, and assesses the strength of association between an item and the remainder of the dimension.

Reliability is concerned with whether an instrument is internally consistent or reproducible. Tests of internal consistency are appropriate for multi-item scales. Cronbach's alpha assesses the overall level of correlation between items within a scale. It is equivalent to the average level of correlation between all the possible halves of a scale. Internal consistency can be assessed with a single administration of an instrument but is applicable only to multi-item scales. Instruments that do not have multi-items scales must be assessed for test-retest reliability.

Test-retest reliability is designed to take account of variation over time. It assesses the level of association between two sets of instrument scores from the same group of patients on two different occasions. There is no real agreement on the length of time between administrations of test and retest questionnaires, but it should not be so short that patients can recall their previous responses, nor so long that health may have changed. In practice, the use of postal surveys means that two weeks is the minimum time between test and retest. It is common for such studies to include a health transition question to identify patients who do not change between administrations. These questions assess whether there has been any change in the aspects of health that are the focus of the instrument being assessed.

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

Validity is concerned with whether an instrument is measuring what is intended. Both qualitative and quantitative methods can be used to assess the validity of instruments. Validity is not a fixed property ascertainable from a single experiment, and should be assessed in relation to the application of an instrument. New instruments, refined instruments, and instruments being used in a new setting should be tested for validity.

Face and content validity are matters of qualitative judgement as to whether an instrument is suitable for its proposed application. Face validity concerns judgements about individual items after an instrument has been constructed, whereas content

validity is concerned with judgements about how well the domain of interest has been sampled for items. These forms of validity are assessed through inspection of the instrument; together they assess whether the items relate adequately to the domain of interest, and whether the domain is sufficiently covered by the items in their entirety (ibid.).

Patient-reported health outcome measures are concerned with the measurement of variables that are not directly observable, referred to as 'hypothetical constructs'. Drawing on theoretical or empirical work, patient-reported measures can be expected to have quantifiable relationships with other constructs. For example, patients with restricted physical functioning may take more days off work, while patients with severe pain may take more analgesics.

Construct validity involves comparing instrument scores with other variables. Construct validation is not necessarily concerned with strong relationships, but with whether an instrument measures something in a way predicted by theory or established evidence. Thus, validity testing should be seen not as a once for all exercise, but an ongoing process of accumulating evidence arising from a number of tests.

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

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

Responsiveness refers to the ability of an instrument to measure significant changes in health. This is an important measurement property of any instrument used for measuring outcomes. Responsiveness is assessed by looking at changes in instrument scores for groups of patients whose health is known to have changed. Two approaches are applied in selecting patients for studies of responsiveness: in the first, the patient is given an intervention of known efficacy and the responsiveness of the instrument is assessed for patients who are expected to have improved; in the second, patient or physician ratings of change are used to select patients (Fitzpatrick et al., op. cit.).

Various statistics can be used for quantifying the responsiveness of instruments, many of which are referred to as effect sizes. Responsiveness is rarely defined in terms of importance to patients but, as stated by Fitzpatrick et al. (ibid.), this is a key

consideration in relation to patient-reported measures. Few studies have assessed whether the changes demonstrated by outcome measures are at a level patients consider important.

Quality of life assessment packages with three components have been recommended for the measurement of outcomes in diabetes patients (Jacobson, De Groot et al., 1994; Beaser, Garbus et al., 1996). The first component is a generic instrument allowing comparison with other illness groups. The second is a disease-specific instrument likely to be more responsive to important changes in health. The third is a situation-specific instrument, or questions designed to address specific situations.

Generic instruments used with diabetes patients include the Nottingham Health Profile/NHP (Hunt, McKenna et al., 1981), Short Form 36-item Health Survey/SF-36 (Ware and Sherbourne, 1992), and the UK version of the Sickness Impact Profile/SIP (Bergner, Bobbitt et al., 1981a; Charlton, Patrick et al., 1983). However, both the NHP and the UK version of the SIP have been found to produce skewed responses towards positive health in diabetic patients (Bardsley, Astell et al., 1993; McColl, Steen et al., *op. cit.*). This ceiling effect is likely to limit the responsiveness to change of these instruments, making them unsuitable for use as outcome measures in diabetes.

The SF-36 scales have been described as mostly relevant to people with diabetes (Bradley, 1994). The SF-36 has also been shown to be more sensitive than the NHP to low levels of morbidity in the general population, and more sensitive than the UK version of the Sickness Impact Profile in a diabetic population (Brazier, Harper et al., 1992; McColl, Steen et al., *op. cit.*).

Disease-specific instruments cover the domains of quality of life and psychological well-being. The literature review underpinning this report identified seven publications (Rodin, 1990; Cox and Gonder-Frederick, *op. cit.*; Eiser and Tooke, *op. cit.*; Hörnquist, Wikby et al., 1995; Bradley, 1994 *op. cit.*; Baker and Meadows, 1993; Beaser et al., *op. cit.*; Rubin and Peyrot, 1999) reviewing the impact of diabetes on quality of life, and the implications for its measurement. The reviews concur that patient-reported measures of health outcome for a chronic illness like diabetes should be multi-dimensional, encompassing functional disability, emotional distress, psychiatric illness, and a sense of well-being.

Diabetes-specific instruments have also been developed measuring the constructs of knowledge and cognitive functioning, attitudes and beliefs, management and self-care behaviour, and treatment satisfaction (Meadows, Steen et al., *op. cit.*). To the extent that interventions designed to influence these outcomes have as their desired endpoint improvements in quality of life, these instruments can be described as situation-specific.

Chapter 2: METHODS

The focus of the review is disease-specific patient-reported health outcome measures for adults with insulin-dependent diabetes mellitus (IDDM) and non-insulin-dependent diabetes mellitus (NIDDM). This includes broad-based measures of symptoms and instruments purporting to measure such constructs as health-related quality of life (HRQL), quality of life and psychosocial functioning.

a) Search strategy

The search strategy was designed to retrieve references relating to patient-reported measures of health outcome and diabetes, including the development and testing of instruments, reviews of such instruments, and conceptual and methodological issues in measurement. The search strategy was not designed to retrieve references relating purely to the application of instruments, but subsequent searches using instrument names and abbreviations produced a number of such references. The search terms were developed by combining terms specific to patient-reported health outcomes (developed previously) with diabetes-specific terms. The search strategy was as follows:

Diabet* and

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

The following search strategy was designed to retrieve symptom-based instruments specific to diabetes:

symptom* near2 (index or indices or instrument* or measure* or questionnaire* or profile* or scale* or score* or status or survey*) and diabet*

Electronic databases including Embase, Medline, Biological Abstracts, Psychlit, AMED, Econlit, Sociological Abstracts, and Cinahl were searched for the period 1990-1999. No criteria were set regarding reference type. The references retrieved produced an electronic database within which the title and abstract of each reference were assessed against the inclusion criteria. The names of identified instruments were used as terms for a further search of the electronic databases.

Original papers were retrieved when references related to the development or evaluation of disease-specific patient-reported health outcome measures. The citation lists of these papers were examined for references to additional potentially relevant instruments. Abstracts relating only to the application of an instrument, within a clinical trial for example, were excluded from the review but retained in the database.

The first and second authors of instruments included in the review were sent a letter asking them if they had produced any further work relating to the instrument, and requesting them to supply a copy of the questionnaire along with any scoring instructions. Reminders were sent at one month.

The following sources were also examined:

- 6th Annual Conference for the International Society for Quality of Life Research, November 1999
- 58th-60th Scientific Session of the American Diabetes Association
- The Outcomes Structured Abstracts Database
- Salek, S. (1998). *Compendium of Quality of Life Instruments*. New York: Wiley.
- Tamburini, M. *Researcher's Guide to the Choice of Quality of Life Assessment in Medicine* - <http://www.qlmed.org/index.html>
- Bowling, A. (1995). *Measuring Disease*. Buckingham: Open University Press.
- Shumaker and Berzon, eds. (1995). *The International Assessment of Health-Related Quality of Life: Theory, Translation, Measurement and Analysis*.
- Spilker, B. (1996). *Quality of Life and Pharmacoeconomics in Clinical Trials* (2nd Ed.) Philadelphia: Lippincott-Raven.
- *Diabetes Health Economics Bibliography 1990-2000*

b) Inclusion criteria

The inclusion criteria were as follows:

- (i) the authors define their instrument as measuring quality of life or health-related quality of life; *or*
- (ii) the instrument appears to have health-related quality of life as a main focus; *or*
- (iii) the instrument is multi-faceted with a significant health-related quality of life component forming a dimension in itself; *or*
- (iv) the instrument is a broad-based measure of symptoms of diabetes; *or*
- (v) the instrument has undergone some form of evaluation for reliability, validity or responsiveness.

Situation-specific instruments were excluded; these were instruments having as their sole focus constructs such as knowledge and cognitive functioning, attitudes and beliefs, management and self-care behaviour, and treatment satisfaction. Instruments measuring diabetic complications were excluded. Instruments for which there was no published development or testing for any of the measurement properties of reliability, validity and responsiveness were also excluded. The review was restricted to instruments providing an English-language validation for use among adult patients only.

c) Data extraction

The content of the instruments was reviewed against the general classification proposed by Fitzpatrick et al. (op. cit.) shown in Table I, which was adapted for diabetes-specific instruments. This was based on copies of instruments included with published papers, or a reply to the request for further information from authors.

Table I: Classification of dimensions for patient-reported health outcome measures

<p>I Physical Function Mobility, dexterity, range of movement, physical activity, activities of daily living, ability to eat whenever and whatever</p> <p>II Symptoms Diabetes-specific symptoms (hypoglycaemia, polyuria, thirst) More general symptoms such as pain, nausea, energy, sleep</p> <p>III Global judgements of health</p> <p>IV Psychological Well-being Anxiety, depression, coping, positive well-being and adjustment, sense of control, self-esteem</p> <p>V Social well-being Family and intimate relations Social contact, integration and social opportunities Leisure activities Sexual activity and satisfaction</p> <p>VI Cognitive functioning Cognition, alertness, concentration, memory, confusion, ability to communicate</p> <p>VII Role activities Employment, household management, financial concerns</p> <p>VIII Personal constructs Satisfaction with bodily appearance Stigma and stigmatising conditions Life satisfaction Spirituality</p> <p>IX Satisfaction with care and flexibility of treatment</p>

Fitzpatrick et al. (op. cit.)

The extraction criteria were based on those used in previous reviews (Wilkin, Hallam et al., 1992; Bowling, 1995; McDowell and Newell, 1996) and followed recommendations relating to the development and testing of instruments (Streiner and Norman, 1995; Fitzpatrick et al., op.cit.). The papers were reviewed independently by two researchers (Andrew Garratt and Louise Schmidt), and information relating to development and testing was extracted under the headings in Table II.

Information extracted under the heading of purpose includes what the instrument purports to be measuring and the applications for which it is designed. What the instrument purports to be measuring could include some underpinning definition of quality of life. Applications include research and clinical practice.

The heading of description includes information relating to the development and scoring of the instrument. Instrument development includes the source of items, the scaling of items, data quality and the use of principal component or factor analysis to assess dimensionality. Any modifications made to the instrument are also described here.

Information extracted under the heading of patients was largely defined by what was documented by the majority of authors, but attention was given to characteristics identified as being associated with quality of life (Rubin and Peyrot, *op.cit.*). Any comparison of responders and non-responders was also included.

Information extracted under the heading of reliability includes the results of tests of internal consistency and test-retest reliability, specifically Cronbach's alpha and test-retest correlation coefficients. Also included is the form of any health transition question used to assess changes in health between test and retest. The reliability coefficient should exceed 0.7 if the instrument is to be used with groups of patients; if the instrument is to be used in individual patients then it should exceed 0.9 (Fitzpatrick et al, *op. cit.*).

The information extracted under the heading of validity includes qualitative and quantitative forms of validation. The methods and results of testing content, face and construct validity are extracted, together with any a priori expected hypotheses relating to the direction and magnitude of coefficients. If expected relationships are not defined, information is extracted relating to variables found to have evidence of a relationship with quality of life in diabetes patients in a recent comprehensive and structured review (Rubin and Peyrot, *op. cit.*).

This review found that:

- the duration and type of diabetes are not consistently associated with quality of life
- intensive treatment does not impair quality of life
- better glycaemic control is associated with better quality of life
- complications of diabetes are the most important disease-specific determinant of quality of life
- quality of life is generally worse among women with diabetes than men with diabetes
- advancing age seems to impact upon quality of life in patients with diabetes, especially as regards physical functioning
- married people with diabetes have a better quality of life than single people with diabetes
- people with diabetes in lower socioeconomic groups tend to have poorer quality of life than those in higher socioeconomic groups

Information extracted under the heading of responsiveness includes both the study design and the statistics used to quantify responsiveness. Responsiveness studies can be based on the use of a treatment of known efficacy, or health transition questions for

judging whether a patient's health has changed. Responsiveness is quantified using effect size statistics such as the standardised response mean. Developmental and evaluative papers incorporating longitudinal studies are included as evidence of instrument responsiveness.

Acceptability refers to patient acceptance of the instrument. The instruments are all administered by means of a self-administered questionnaire. Information extracted under this heading includes response rates, completion rates and levels of missing data.

Table II: Data extracted relating to instrument development and testing

<i>Purpose</i>	<i>Patients</i>	<i>Description</i>	<i>Reliability</i>	<i>Validity</i>	<i>Responsiveness</i>	<i>Acceptability</i>
<p>What does the instrument purport to measure?</p> <p>What are the proposed applications?</p>	<p>Study setting</p> <p>Age</p> <p>Sex</p> <p>Ethnicity</p> <p>Socioeconomic status</p> <p>Marital status</p> <p>Type of diabetes</p> <p>Disease duration</p> <p>Type of treatment</p> <p>Complications</p> <p>Haemoglobin levels</p>	<p>Number of items/dimensions</p> <p>Scoring</p> <p>Item generation</p> <p>End effects</p> <p>Missing data</p> <p>Dimensionality</p> <p>Item-total correlation</p>	<p>Internal consistency (Cronbach's alpha)</p> <p>Test-retest (correlation)</p>	<p>Content and face validity</p> <p>Construct validity</p> <p>Generic/specific instruments</p> <p>Clinical variables: - haemoglobin - complications</p> <p>Sociodemographic variables: - age - sex</p>	<p>Effect size</p>	<p>Response rates</p> <p>Completion times</p> <p>Completion rates</p>

Chapter 3: RESULTS

a) Search outcome

The main search strategy produced 1115 references, which fell to 846 references after duplicates were removed. The application of inclusion criteria produced 171 references. The second search using individual instrument names as keywords produced 1582 references. After excluding duplicates and those not related to diabetes (which arose since we also searched on abbreviations of measures), 405 references remained. The application of inclusion criteria produced 81 references.

From these references, 28 potentially relevant instruments were identified; 20 met the inclusion criteria and were included in the review (Table III). Of the remaining eight instruments, three were excluded because quality of life was not a significant component of the instrument, and five were excluded due to insufficient published material on development and/or evaluation. Several instruments that had undergone subsequent revision were classified according to the main body of work: for example, the WHO-Ten Index was referred to under the Well-being Questionnaire.

The search for symptom-based references did not produce any additional instruments that met the inclusion criteria. The majority of the references related to clinical rather than patient-reported health outcome measures. The search of the other sources (conference abstracts, compendia, etc.) did not produce any additional instruments that met the inclusion criteria.

The review has produced a bibliographic database containing the references relating to the development and testing of the disease-specific instruments reviewed here. The database also includes a number of other references of interest. First, there are a number of references relating to the application of the instruments reviewed, including clinical trials. Secondly, there are several diabetes-specific instruments that measure constructs such as knowledge and cognitive functioning, attitudes and beliefs, management and self-care behaviour, and treatment satisfaction. These situation-specific instruments did not meet the entry criteria for this review but are listed in an appendix. The search strategy was not designed to capture these instruments, but rather measures of quality of life and therefore this reference list should not be considered exhaustive.

Detailed reviews of individual instruments are contained in Chapter 4.

Table III: Instruments included in the review

<i>Name of Instrument</i>	<i>Principal author</i>	<i>No. of evaluative papers</i>	<i>Copy of instrument Yes/No</i>	<i>Authors' reply Yes/No</i>
Appraisal of Diabetes Scale	Carey	1	Y	Y
Audit of Diabetes Dependent Quality of Life	Bradley	1	Y	Y
Diabetes Care Profile	Fitzgerald	4	Y	Y
Diabetes Health Profile	Meadows	2	Y	Y
Diabetes Health Status Questionnaire	Wierenga	1	Y	Y
Diabetes Impact Measurement Scales	Hammond	1	Y	Y
Diabetes Quality of Life	Jacobson	5	Y	N
Diabetes Quality of Life Clinical Trial Questionnaire	Shen	1	Y	Y
Diabetic Quality of Life Questionnaire	Henderson	1	Y	N
Diabetes-Related Knowledge and Quality of Life Questionnaire	Gilden	2	Y	N
Diabetes-Specific Quality-of-Life Scale	Bott	1	Y	Y
Diabetes-39	Boyer	1	Y	Y
Multidimensional Diabetes Questionnaire	Talbot	2	Y	N
Newcastle Diabetes Symptoms Questionnaire	McCull	3	Y	Y
Quality of Lfe: Status and Change	Hörnquist	8	N	Y
Questionnaire on Stress in Patients with Diabetes-Revised	Herschbach	3	Y	Y
Social Psychological Health States	Given	1	Y	Y
Type 2 Diabetes Symptom Checklist	Grootenhuis	1	Y	Y
Well-being Enquiry for Diabetics	Mannucci	1	Y	N
Well-being Questionnaire	Bradley	13	Y	Y

b) Patient characteristics

The populations in which the instruments have been developed or evaluated are shown in Table IV. Just four of the twenty instruments have been developed or evaluated in UK patients: the Audit of Diabetes Dependent Quality of Life (Bradley, Todd et al., 1999) Diabetes Health Profile (Meadows, Steen et al., op. cit.), Newcastle Diabetes Symptom Questionnaire (McColl, Steen et al., op. cit.) and the Well-being Questionnaire (Bradley, 1994 op. cit.).

The majority of the remainder have been developed and evaluated in Europe and the United States. The Multidimensional Diabetes Questionnaire (Talbot, Nouwen et al., 1997) was developed in Canada and the Type 2 Diabetes Symptom Checklist (Grootenhuis, Snoek et al., 1993) underwent further evaluation in Trinidad and Tobago.

The four instruments developed within the UK have been evaluated in patients with both types of diabetes. The instruments have not been evaluated in subgroups of the populations in which they were developed, for example, older people or ethnic groups.

Table IV: Patient characteristics

<i>Name of Instrument</i>	<i>Sex</i>	<i>Age</i>	<i>Diabetes Type</i>	<i>Country</i>	<i>Setting</i>
Appraisal of Diabetes Scale	M	58	Both	USA	Outpatients
Audit of Diabetes Dependent Quality of Life	F M	52, 62	Both	UK	Outpatients, educational open day
Diabetes Care Profile	F M	54, 61	Both	USA	Community, outpatients
Diabetes Health Profile	F M	39, 40, 52, 61	Both	Holland, UK, Holland	Outpatients
Diabetes Health Status Questionnaire	F M	-	NIDDM	USA	Community
Diabetes Impact Measurement Scales	F M	45	Both	USA	Outpatients
Diabetes Quality of Life	F M	28, 34, 44, 52	Both	Canada, Fance, Germany, USA	Outpatients
Diabetes Quality of Life Clinical Trials Questionnaire	F M	34, 52	Both	USA, Canada, Germany, France	Outpatients
Diabetic Quality of Life Questionnaire	F M	52, 61	IDDM	USA	Outpatients
Diabetes-Related Knowledge and Psychosocial Functioning Questionnaire	F M	70, 68	Both	USA	Outpatients
Diabetes-Specific Quality of Life Scale	F M	36	IDDM	Germany	Community physician
Diabetes-39	F M	52, 62	Both	USA	Diabetes centre, community pharmacy, outpatients
Multidimensional Diabetes Questionnaire	F M	55	Both	Canada	Diabetes education centres
Newcastle Diabetes Symptom Questionnaire	F M	-	Both	UK	Outpatients, general practice
Quality of life: Status and Change	F M	33, 43, 45	IDDM	Sweden, Norway	Outpatients
Questionnaire on Stress in Patients with Diabetes-Revised	F M	43	Both	Germany	Inpatients, outpatients
Social Psychological Health States	F M	-	Both	USA	Diabetes care centres
Type 2 Diabetes Symptom Checklist	F M	65	Both	Holland, Tobago, Trinidad	Outpatients, health centres
Well-being Enquiry for Diabetics	F M	34, 59	Both	Italy	Outpatients
Well-being Questionnaire	F M	30, 35, 37, 48, 51, 71	Both	Albania, Croatia Finland, France, Germany, Holland, Hungary, Norway, Sweden, UK	Outpatients, health centres, private practitioners

c) Instrument content

The content of the instruments is shown in Table V; constructs that are not strictly related to quality of life (QoL) are not shown. Items relating to other conditions and items excluded from the review are not presented; the latter includes items designed for children and those that do not relate to quality of life (QoL). Where non-QoL items are present in a scale containing QoL items, the construct that these items relate to is recorded (e.g. knowledge). Items relating to whether an aspect of treatment is followed are not presented.

The first column of physical functioning covers items relating to physical ability, ability to operate machinery and drive a car as well as being able to perform basic self-care (washing, dressing). Twelve of the 20 instruments identified include items measuring physical functioning.

The second column covers diabetic symptoms and fifteen instruments include items within this category. There is a high level of agreement among the specific symptoms addressed including sleep, vitality, thirst, the need for frequent voiding, low blood sugar, and tingling of the limbs. Six instruments include items concerning vision.

The global column includes items relating to either general quality of life or the impact of diabetes on general quality of life. This definition was extended to include issues concerning freedom of people with diabetes and their ability to pursue their own plans and schedules, where this is expressed as a general issue and not an aspect of a particular type of treatment. Thirteen instruments include items addressing the global issue of quality of life or the overall impact of diabetes on quality of life.

Nineteen instruments include items relating to psychological well-being which includes self-esteem, depression, fear, aggression, moodiness, and worry. Worry was divided into two main components: disease-related worry (including worry about hypoglycaemic episodes and complications) and worry related to social or vocational issues (including family matters, the future and job-related issues). Five instruments include worry in both spheres, two address social/vocational worry and four address purely disease-related worry.

Fifteen instruments include items relating to social well-being. Items within this construct include impact on the family, sex life, ability to travel, and effect on leisure activities and social relationships.

The impact of diabetes and its treatment on cognitive functioning, including the ability to think clearly, memory and concentration problems, are rarely included in the instruments. The Type 2 Diabetes Symptom Checklist, the Well-being Questionnaire and the Diabetes-related knowledge and Quality of Life Questionnaire include items relating to this construct. The Quality of Life Status and Change instrument purports to include items relating to this construct but this has not been confirmed by questionnaire.

The role activities column refers to the impact of a disease on the ability to perform normal duties relating to employment, schooling or household activities. Twelve instruments include at least one item addressing this aspect of quality of life.

The personal constructs column refers to satisfaction with bodily appearance, life satisfaction, spirituality, and stigma and stigmatising conditions (Fitzpatrick *op. cit.*). This was extended

to include perceptions and behaviour of others, which encompasses a desire to hide the existence of diabetes or its treatment from others, the way other people react to the disease, and concern about possible embarrassing situations.

Twelve instruments include items addressing this aspect of quality of life. Four instruments include items relating to appearance, eight include items concerning stigma and the attitudes/behaviour of others, three include items relating to feeling handicapped or different, and five include items relating to life satisfaction.

Items were classified under the construct satisfaction with care when the item clearly related to satisfaction and not some aspect of quality of life. Where an item referred to satisfaction with one of the major constructs of quality of life, for example social life, this was classified under social well-being. Three instruments include items relating to satisfaction with diet and two instruments include items relating to satisfaction with the time needed for aspects of care. Three instruments included items relating to satisfaction with diet, and three included items relating to satisfaction with doctors and consultations.

The construct treatment and general quality of life was created to cover aspects of quality of life affected by diabetes care regimens not covered by the other constructs. This includes the large number of items that address the degree to which an aspect of care is a burden or bothersome, causes problems or difficulties, or is inconvenient. Twelve instruments include items relating to these issues and many include several relevant items.

The final or 'other' category includes situation-specific variables such as knowledge about the disease, attitudes towards diabetes, and perceptions about the severity of the disease.

Table V: Content of instruments included in the review

Name of Instrument	Appraisal of Diabetes Scale/ADS	Audit of Diabetes Dependent Quality of Life/ADDQoL¹	Diabetes Care Profile/DCP²	Diabetes Health Profile/DHP
Physical functioning		Physical ability 1	Active 1	
Symptom			Low blood sugar 2 High blood sugar 1 Ketones 1 Reasons for: - high blood sugar 7 - low blood sugar 8	
Global	Diabetes: - uncertainty 1 - likely to worsen 1 - developing life goals 1 - coping 1	General QoL 1 Diabetes/ general QoL 1	Diabetes/treatment: - normal activities 1 - schedule 1 - makes life difficult 1 - able to do anything 1 - effect on life 1 Things going well 2	Diabetes/ general QoL 1
Psychological well-being	Disease: upsetting 1	Future worry/ fear: personal & family 2 Motivation 1	Afraid re. disease 1 Acceptance of disease 1 Depressed 1 Inferiority 1 Handle feelings 3	Tension/edgy 2 Self-harm, no wish for life 3 Lose temper, moody 5 Depression 1 Cry 1 Worry/fear: - social 1 - re. disease 4 Wish away diabetes 1 Look forward to future 1
Social well-being		Social/ friendship 2 Family 1 Sex 1 Leisure 1 Travelling 1 Enjoyment of food 1	Going out, travelling 1 Relationships 1 Friends 1 Time alone 1	Family 1 Avoid going out 2 Stay out late 1
Cognitive functioning				
Role activities		Work 1	Financial 2 Daily activities 1	
Personal constructs		Stigma: others 1	Life satisfaction 2	
Treatment & general quality of life	Control of diabetes 2		Diet restrictive 2	Food/diet: controls life 1 Injection: planning schedule 1 Problems with diet 2
Satisfaction with care			Disease management 1	
Other			Self-care 5 Attitude 4	Attitude to food 2

¹ The ADDQoL asks patients to indicate how important each aspect of quality of life is.

² scales of diabetes control, social and personal factors, and attitudes towards diabetes, only

<i>Name of Instrument</i>	Diabetes Health Status Questionnaire/DHS³	Diabetes Impact Measurement Scales/DIMS	Diabetes Quality of Life/DQoL
<i>Physical functioning</i>	Energy Physical health	Muscular strength/ endurance	1 Using car/machinery Diabetes interferes with exercise 1
<i>Symptom</i>	Diabetes: physical symptoms 20	Vision 1 Energy 2 Sleep 1 Faint/dizzy 1 Nausea 2 Appetite 1 Bowel problems 3 Thirst/polyuria 1 Low blood sugar 2 Tingling in extremities 1	Sleep 2 Pain 1 Physically ill 1 Low blood sugar 1 Polyuria 1
<i>Global</i>	Current health ratings	Diabetes restrictive 1	
<i>Psychological well-being</i>	Emotional health	Worry: future 1 Worry: disease 1 Worry: general 1 Optimism 1 Depression 1 Feel useful 1 Feel things going well 1	Self-esteem 1 Worry: social/ Vocational 9 Worry: disease 2
<i>Social well-being</i>		Family 2 Sex 2 Social/leisure 1	Family 2 Social relations 2 Sex 2 Leisure 2
<i>Cognitive functioning</i>			
<i>Role activities</i>		Usual occupation 2 General performance 1	Daily activities 2
<i>Personal constructs</i>		Life satisfaction 1	Appearance 1 Life satisfaction 1 Stigma: others 6
<i>Treatment & general Quality of life</i>		Medicine inconvenient 1 Care worth the effort 1 Monitoring inconvenient 1 Food/diet restrictive 1	Diet restrictive 1
<i>Satisfaction with care</i>		Knowledge 2 Perception 1 In control 1 Food 1 Able to follow recommendations 1	Time: exercising, check-ups, monitoring, managing disease 4 Treatment 1 Diet 1
<i>Other</i>			Knowledge 1

³ The instrument provided by the author is different from the instrument for which published evidence was available. The content presented here is from the publication.

<i>Name of Instrument</i>	Diabetes Quality of Life Clinical Trial Questionnaire/ DQLCTQ⁴	Diabetic Quality of Life Questionnaire/ DQLQ	Diabetes-Related Knowledge & Quality of Life Questionnaire	Diabetes Specific Quality of Life Scale/DSQoLS⁵
<i>Physical functioning</i>	Moderate/vigorous activity 2 Walking 2 Movement 1 Self-care 1 Car/machinery 1 Diabetes interferes with exercise 1 Physical activity: flexibility 1			Physical strength 2 Spontaneous activity 1
<i>Symptom</i>	Vitality 3 Sleep 2 Tired 1 Physically ill 1 Vision 1 Nausea 1 Treatment pain 1 Low blood sugar 1 Polyuria 2 Thirst 1 Hunger 1 Tingling of the extremities 1		Vision 1 Hearing 1	Thirst/dry mouth 1 Blood sugar 2 Low blood sugar 2 High blood sugar 1 Complications 1 Nervous/restless 1 Tired 1 Polyuria 1 Physically unwell 1 Physical complaints 1 Infections/ itching 1 Weak/lazy 1
<i>Global</i>	General health 2 Spur of the moment activities 2	Diabetes/ general QoL 1	Diabetes & lifestyle 2	Diabetes: future plans 1 Need to think about diabetes 2 Diabetes controls life 1
<i>Psychological well-being</i>	Discouraged/despair 2 Frustrated 1 Nervous 1 Feel down 2 Happy 1 Self-esteem 1 Calm 1 Weighed down 1 Worn out 1 Full of pep 1 Afraid: health-related 1 Worry: - health related 1 - social/vocational 11 - disease-related 16 - appearance 1	Worry: disease-related 1	Depressed 1 Loneliness 1 Fear: treatment-related 1 Nervous: disease-related 1 Get upset: disease-related 1	Anxious: disease-related 1 Worry: disease-related 11 Depressed 1 Anxious/ threatened 1

⁴ This table refers to the revised version of the instrument, not the original instrument provided by the author.

⁵ Ten DSQoLS items ask how important are particular aspects of treatment/quality of life, and are not shown here.

<i>Name of Instrument</i>	Diabetes Quality of Life Clinical Trial Qu'aire/DQLCTQ	Diabetic QoL Questionnaire/ DQLQ	Diabetes-related Knowledge & Quality of Life Qu'aire	Diabetes Specific Quality of Life Scale/DSQoLS
<i>Social well-being</i>	Family 2 Relationship/friends 2 Sex 2 Leisure 2 Social 3	Social 1 Insulin: social 1	Family 1 Diet: - social 3 - family 2	Free time/ flexibility 2 Family 1 Travelling 1 Social 2 Hobbies 1
<i>Cognitive functioning</i>			Memory 1	
<i>Role activities</i>	Daily activities 5	Insulin: daily activities 1		
<i>Personal constructs</i>	Appearance 1 Life in general 1 Stigma: others 10		Stigma: treatment/ monitoring 2	Appearance 1 Stigma: others 6 Feel disabled 1
<i>Treatment & general Quality of life</i>	Willingness: insulin treatment 1 Control of diabetes 1 Flexibility: diet 5	Insulin: ease/ convenience 3 Injection: discomfort 1 Injecting: self-conscious 1	Diet: difficulty 1 Treatment: difficulty 2 Treatment/ monitoring: time 2 Monitoring: - difficulty 2 - cost 1 - bother 1	Bothersome/ burdensome: - monitoring blood sugar 1 - food plan 1 - time on treatment 1 - diet 1 - diet spontaneity 1 - diabetes supplies 1 - insulin 1
<i>Satisfaction with care</i>	Time: managing, check-ups, monitoring 3 Treatment 1 Diet: flexibility 1 Time: exercising 1 Diet: restricting 1 Insulin treatment 1		Doctor communication 2	Diet flexibility 1 Self-monitoring 1 Medical consultations: time 1
<i>Other</i>	Knowledge 1 Attitude to diabetes 3 Perception 1		Attitude 11 Disease: perception 1	Food restricted 4 Food: desire to eat without planning 1

<i>Name of Instrument</i>	Diabetes-39/D-39	Multidimensional Diabetes Questionnaire/MDQ⁶	Newcastle Diabetes Symptoms Questionnaire/NDSQ
<i>Physical functioning</i>	Walking 2 Steps 1 Self-care 1		
<i>Symptom</i>	Energy 2 Weakness 1 Vision 1 Sleep 1 Complications 1 Lose control of sugar levels 1		Thirst 1 Shaky 1 Blurred vision 1 Faint 1 Passing lots of water 1 Hunger 1 Sleepy 1 Cold hands/feet 1 Pins & needles 1
<i>Global</i>	Other health problems 2 Gen. restrictions 2 General QoL 1 Diabetes/general QoL 2	Diabetes &: - daily activities 1 - limitation of general activity 1 - interference with planning 1 - schedule restriction 1	
<i>Psychological well-being</i>	Stress 1 Depressed 1 Worry: financial 1 Worry: future 1	Worry: disease-related 2	
<i>Social well-being</i>	Family/friends 3 Sex 3	Travelling 1 Social 2 Family 1	
<i>Cognitive functioning</i>			
<i>Role activities</i>	Household duties 1	Work 1	
<i>Personal constructs</i>	Stigma: others 2		
<i>Treatment & general Quality of life</i>	Medication 1 Diet 1 Treatment plan 1 Exercise requirements 1 Self-monitoring 2 Controlling: - time 1 - diabetes 1		
<i>Satisfaction with care</i>			
<i>Other</i>	Perception 1		

⁶ scales of Interference and Severity only

<i>Name of Instrument</i>	Quality of life: status & change/ QLsc⁷	Questionnaire on Stress in Patients with Diabetes-Revised/QSD-R	Social Psychological Health States/SPHS
<i>Physical functioning</i>	Yes	Physical ability/ activity 2	Walking 1
<i>Symptom</i>	Yes	Complications 1 Wind 1 Sweating 1 Energy 1 Pain 1 Hypoglycaemia 3 Physical distress 1 Thirst 1	Sleep 1
<i>Global</i>			
<i>Psychological well-being</i>	Yes	Fear: alone 1 Irritable 2 Moody 1 Serious 1 Nervous 1 Worry: family 2 Guilt: non-compliance 2 Worry: disease 3	Nervous/restless 1 Impatience 1
<i>Social well-being</i>	Yes	Travelling 1 Sex 1 Social 2 Free time 2 Family 1	Social/friends 4 Leisure 3 Family 4 Stay home 3
<i>Cognitive functioning</i>	Yes		
<i>Role activities</i>	Yes	Daily activities 4	Household work 3 Work/job 7
<i>Personal constructs</i>	Yes	Attractive 1 Feel handicapped 2 Stigma 1	
<i>Treatment & general Quality of life</i>		Self-monitoring: problem 1 Restrictive diet: controlling 2 Equipment 1 Diet: lack of food 1	
<i>Satisfaction with care</i>		Doctors: - information 1 - time 1 - care 1	
<i>Other</i>		Knowledge 1	

⁷ as described in publication: instrument questionnaire not provided

<i>Name of Instrument</i>	Type 2 Diabetes Symptom Checklist/DSC Type 2⁸	Well-being Enquiry for Diabetics/WED	Well-being Questionnaire/WBQ⁹
<i>Physical functioning</i>		Physical Activity 1 Using car/machinery 1	
<i>Symptom</i>	Thirst 3 Voiding 1 Energy 1 Fatigue 3 Sleepiness 1 Shortness of breath 2 Palpitations 1 Aches/pain 5 Vision 5 Numbness 2 Tingling in extremities 3 Odd feeling in legs/feet 1	Sleep 1 Polyuria 1 Hypoglycaemia 1 Feel well 1 Pain 2 Sick 2	Sleep 2 Energy 3
<i>Global</i>		Diabetes limits freedom 1	
<i>Psychological well-being</i>	Irritability 2 Moodiness 1	Fear: - disease 4 - general 1 - future 1 Anxious: general 1 Depressed/sad 2 Worry: general 1 Decision making 1 Trust yourself 1 Confidence 1	Calm/nervous 2 Cope with life 2 Feel useful 1 Depressed/upset 2 Cry 1 Afraid 1 Falling apart 1 Well adjusted 1
<i>Social well-being</i>		Sex 2 Family 2 Social relations 2 Leisure 2 Travel 1	
<i>Cognitive functioning</i>	Dull head 1 Concentration 1 Attentiveness 1		Think clearly 1
<i>Role activities</i>		Normal activities 2	
<i>Personal constructs</i>		Stigma: others 6 Feel different 1	Life satisfaction 1 Life full/interesting 2 Enjoy/happy with life 2
<i>Treatment & general Quality of life</i>		Diet: limiting/burden 2 Diet: difficult 1 Treatment schedule: burden 1 Time managing diabetes 1	
<i>Satisfaction with care</i>			
<i>Other</i>		Dependence on others 1 Attitude to food 2	

⁸ final version of the instrument

⁹ 22-item version of the instrument

d) Instrument development

The methodology used in the development of the content of the 20 instruments included in the review is summarised in Table VI. Item derivation is the first stage in instrument development; this can involve several sources that are not mutually exclusive. Twelve of the twenty instruments report using a literature review as a source for items. Twelve instruments report using 'experts' in the form of focus groups, or interviews with clinicians and/or patients as a source of items. Four instruments use some theory to underpin instrument development.

Items with high levels of missing data should be considered for removal from an instrument. Levels of missing data were reported for items within four instruments. Items with large floor or ceiling effects are poor at discriminating between patients; such effects can also limit responsiveness to change. These items should likewise be considered for removal. Similarly, scales based on summated items with large end effects are not likely to be responsive to change. End effects at the item or scale level were reported for four instruments.

Dimensionality refers to the grouping of items within an instrument according to constructs, for example, physical or social functioning. Dimensionality can be assessed empirically using principal component analysis or factor analysis. Both techniques were used in the construction of eight instruments included in the review. The source of items was not reported for two of the instruments, and the remaining tests were not reported for four instruments.

Table VI: Instrument development

<i>Name of Instrument</i>	<i>Item derivation</i>	<i>Data missing</i>	<i>Floor/Ceiling effects</i>	<i>Dimensionality</i>	<i>Item-total correlation</i>
Appraisal of Diabetes Scale	Theory, literature review	-	-	Yes	Yes
Audit of Diabetes Dependent Quality of Life	Literature review, clinicians, patients	Yes	-	Yes	Yes
Diabetes Care Profile	Theory, literature review	-	-	Parent instrument only	-
Diabetes Health Profile	Literature review, clinicians, patients	Yes	Item, scale	Yes	Yes
Diabetes Health Status Questionnaire	-	-	Item	-	-
Diabetes Impact Measurement Scales	Clinicians, literature review	Yes	Scale	Yes	Yes
Diabetes Quality of Life	Literature review, clinicians, patients	Yes	-	-	-
Diabetes Quality of Life Clinical Trial Questionnaire	Literature review, clinicians, patients	-	-	-	-
Diabetic Quality of Life Questionnaire	-	-	-	-	-
Diabetes-Related Knowledge and Quality of Life Questionnaire	Literature review, clinicians, patients	-	-	Yes	-
Diabetes-Specific Quality-of-Life Scale	Literature review, clinicians, patients	-	-	Yes	-
Diabetes-39	-	-	-	-	-
Multidimensional Diabetes Questionnaire	Theory, literature review	-	-	Yes	-
Newcastle Diabetes Symptoms Questionnaire	Literature review, clinicians, patients	-	Items	Yes	Yes
Quality of Life: Status and Change	Theory	-	-	-	-
Questionnaire on Stress in Patients with Diabetes-Revised	Clinicians, patients	-	-	Yes	-
Social Psychological Health States	Literature review, patients	-	-	Yes	Yes
Type 2 Diabetes Symptom Checklist	Literature review, clinicians	Yes	Scale	Yes	Yes
Well-being Questionnaire	Clinicians, patients	-	-	Yes	-
Well-being Enquiry for Diabetics	Literature review	-	-	Yes	Yes

e) Instrument evaluation

With the exception of the Diabetes Care Profile and the original version of the Diabetes Quality of Life Clinical Trials Questionnaire, the majority of instruments have completion times of under 20 minutes. However, the completion times were not available for several instruments. Four of the instruments have more than 50 items, which may limit the scope for their inclusion in questionnaires alongside generic or situation-specific instruments, as part of a package of outcome measures.

The instruments are rated for reliability and validity: see Table VII. Very few instruments have been formally assessed for responsiveness and so a rating for this is not given. The rating method follows that of McDowell and Newell (op. cit.). The criteria for inclusion in the review included some form of testing for measurement properties. All twenty instruments have been assessed for internal consistency or test-retest reliability. Nineteen instruments have been assessed for the former, and twelve of these meet the recommendations for group comparisons. Sixteen instruments were assessed for test-retest reliability, and all the scales produced by five of these meet the recommendations for group comparisons.

These figures include any modifications made to the instrument that may have improved their reliability. For example, the reliability estimates produced by the Diabetes Quality of Life in Clinical Trials Questionnaire (Shen, Kotsanos et al., 1999) were not as acceptable for group comparisons as those produced by the revised instrument. Three of the instruments achieve ratings of reliability that can be considered as excellent: the Diabetes Quality of Life in Clinical Trials Questionnaire-Revised (Shen et al., op. cit.), the Newcastle Diabetes Symptoms Questionnaire (McColl, Steen et al., op. cit.), and the Type 2 Diabetes Symptom Checklist (Grootenhuis, Snoek et al., 1994).

Patients were reported to have been involved in the development and selection of items for ten of the instruments. However, few authors gave explicit consideration to face and content validity. Furthermore, several studies failed to describe the role of patients in the development process, making judgements of content validity difficult. Several methods have been used to assess the construct validity of instrument scores. Eleven instruments have been compared with other generic and specific instruments. Sixteen instruments have been compared with clinical and sociodemographic variables. Four instruments have been compared with single questions relating to aspects of disease impact, quality of life and life satisfaction. The Diabetes Health Status Questionnaire (Wierenga, 1994), the Diabetes-Related Knowledge and Psychosocial Functioning Questionnaire (Gilden, Hendryx et al., 1989), and the Diabetic Quality of Life Questionnaire (Henderson and Tindall, 1990) have not been assessed for construct validity.

There was little use of a priori hypotheses in instrument testing and only four authors described the expected strength of relationships in comparisons with other instruments. The Diabetes Quality of Life Measure (Jacobson, op. cit.) is the only instrument that has undergone testing for all the major forms of validity. This may be due in part to its being one of the first disease-specific instruments for diabetes. No instrument gets the maximum rating for the results of validity testing, but the majority get the second highest rating.

Table VII: Instrument ratings

Name of Instrument	No. of items	Time needed: mins	Reliability		Validity	
			a	b	a	b
Appraisal of Diabetes Scale	7	<5	+++	++	+	+
Audit of Diabetes Dependent Quality of Life	15	-	++	++	++	++
Diabetes Care Profile	234	30-40	++	++	++	++
Diabetes Health Profile	32	-	++	++	++	++
Diabetes Health Status Questionnaire	26	-	++	++	0	0
Diabetes Impact Measurement Scales	44	15-20	++	+	+	++
Diabetes Quality of Life	46	-	+++	++	+++	++
Diabetic Quality of Life Questionnaire	10	-	+	+	0	0
Diabetes Quality of Life Clinical Trials Questionnaire*	142 (57)	(10)	++	++	++	++
Diabetes-Related Knowledge and Psychosocial Functioning Questionnaire	67	-	++	++	+	+
Diabetes-Specific Quality of Life Scale	64	10-20	++	++	++	++
Diabetes-39	39	-	++	++	+	++
Multidimensional Diabetes Questionnaire	41	-	++	++	++	++
Newcastle Diabetes Symptoms Scale	9	-	+++	+++	++	++
Quality of Life: Status and Change	-	-	++	++	++	++
Questionnaire on Stress in Patients with Diabetes-Revised	45	5-15	++	++	++	++
Social Psychological Health States	28	-	++	++	++	++
Type 2 Diabetes Symptom Checklist	34	10	+++	+++	++	++
Well-being Enquiry for Diabetics	60	-	+++	++	++	++
Well-Being Questionnaire	22	-	+++	++	++	++

*numbers in brackets refer to revised version

Key

a Thoroughness rating

0 = no reported evidence of reliability or validity
 + = very basic information only
 ++ = several types of test, or several studies have reported reliability/validity
 +++ = all major forms of reliability/validity testing reported

b Results rating

0 = no numerical results reported
 + = weak reliability/validity
 ++ = adequate reliability/validity
 +++ = excellent reliability/validity

Adapted from McDowell and Newell, op.cit.

Chapter 4: INSTRUMENT REVIEWS

Appraisal of Diabetes Scale/ADS

Purpose

The ADS aims to assess how a person with diabetes evaluates the disease and its impact. It is suggested that the ADS could be used as a brief screening instrument in a clinical setting to identify those patients experiencing, or at risk for, dysphoric reactions and noncompliance problems (Carey, Jorgensen et al., 1991).

Description

The content of the ADS is based on theory and research regarding appraisal processes. Some items were adapted from a generic Attribution Questionnaire (Hammen and Mayol, 1982). The instrument comprises seven items concerning control, uncertainty, coping, effect of diabetes on life goals, predictive view of diabetes, and the degree of distress caused by diabetes. The items use a five-point adjectival scale.

Responses to the scale were evenly distributed, with mean scores close to the midpoint of all items. Principal component analysis produced a single dimension and all items had loadings above 0.40. Item-total correlations were in the range 0.28-0.59.

ADS items are summed to produce a score from 0-35: 0 represents the lowest and 25 the greatest impact of diabetes.

Patients

Two hundred adult males presenting as outpatients at a veterans medical centre were recruited. Patients had a mean age of 58.4 years and were primarily Caucasians, married and high school educated. Mean duration of diabetes was 15 years and 66% of patients were on insulin therapy.

Reliability

The ADS had a Cronbach's alpha of 0.73. Test-retest reliability was assessed by giving the ADS to one sample of patients ($n = 98$) to complete on three occasions: just before blood withdrawal, one hour after completing clinic visit and one week later. Retest questionnaires were returned by 79%. Correlations between the two sets of ADS scores were 0.89 and 0.85 for the one hour and one week retest, respectively.

Validity

It was hypothesised that the ADS would correlate with the Diabetes Regimen Adherence Questionnaire-Revised/DRAQ-R (Brownlee-Duffeck, Peterson et al., 1987), the Diabetes Health Belief Questionnaire-Revised/DHBQ-R, perceived severity and perceived susceptibility subscales only (ibid.), the Diabetic Daily Hassles Scale/DDHS (Kanner, Lee et al., 1986), the Perceived Stress Scale and the Psychiatric Symptom Index, excluding the cognitive disturbance subscale (Cohen, Kamarck et al., 1983) among one sample of patients ($n = 102$). The strength of these relationships was not hypothesised. Correlations ranged from 0.17 to 0.59 for the DRAQ-R and the DDHS, respectively. There was a low correlation of 0.18 between the ADS and HbA_{1c}. The expected level of correlation was not hypothesised.

Responsiveness

Not assessed.

Acceptability

The majority of patients found the instrument quick and easy to complete in under five minutes. The three ADS questionnaires were completed by 79% of patients.

Commentary

The content of the ADS was derived through theory and adaptation of an existing instrument. The unidimensionality of the instrument was supported through principal component analysis. The instrument has evidence for internal and test-retest reliability that makes it suitable for use in groups. The interpretation of the results of the validity testing was limited by the lack of hypotheses. The instrument has only been assessed in male diabetics and has not been tested for responsiveness.

Audit of Diabetes Dependent Quality of Life/ADDQoL

Purpose

The ADDQoL is an individualized instrument designed to measure individuals' perceptions of the impact of diabetes on their quality of life (QoL). The aim was to develop a detailed version of the instrument for research and in-depth clinical work, and a short-form for audit purposes (Bradley, Todd et al., op. cit.).

Description

The instrument comprises two summary items plus 13 items where the respondent is invited to indicate, firstly, the effect of diabetes on this particular aspect of life and, secondly, how important this aspect of life is to overall quality of life. The ADDQoL includes seven domains: physical functioning, symptoms, psychological well-being, social well-being, role activities, and personal constructs.

Two versions of the instrument were piloted, with 12 and 13 items. The latter included an item concerned with the enjoyment of food. Patients can respond "not applicable" to ten of the items. Patients respond to each item by circling a number on a seven-point scale which asks how a particular aspect of their life would be if they did not have diabetes (impact). They then rate the importance of this aspect of their life on a four-point scale (importance). Impact ratings are multiplied by importance ratings to produce a -9 to 9 score, then summed and divided by the number of domains to produce a final score from -9 to 9. Two additional summary items provide indicators of QoL and impact of diabetes on QoL.

The selection of items was influenced by a review of existing instruments, discussions with health professionals and in-depth interviews with twelve diabetes patients. The content was then reviewed by the British Diabetic Association/Royal College of Physicians Working Group and four adults with diabetes.

Six of the ADDQoL items showed responses indicating positive effects of diabetes, which demonstrates the need for bipolar scales to establish the nature of diabetes. All four importance ratings were used for the 13 domains. The authors interpret this as evidence for importance ratings which take individual impacts on quality of life into account. Mean impact scores after weighting for importance ranged from -0.350 to -0.2309. Evidence for the unidimensionality of the instrument was found through a forced one-factor solution and all items had factor loadings above 0.4. When the forced one-factor analysis was repeated, including the food enjoyment item and listwise deletion of missing variables, the factor structure changed little. Item-total correlations for the twelve items were in the range 0.37-0.67.

Patients

Two samples of patients were used to develop and evaluate the ADDQoL. The first sample consisted of 52 patients recruited from an adult outpatients diabetes clinic in Cambridge, UK (Cambridge sample). The mean age was 52 years, 46% of patients were women, and the mean age on leaving full-time education was 17.0 years. Patients not fluent in English were excluded. The mean disease duration was 12.7 years. 32 patients were on a treatment regimen of insulin and diet, 14 on a regimen of tablet and diet, and 6 on diet only. 39 patients had no complications of diabetes while ten did have complications present.

The second sample included 102 patients attending two open educational evenings for adults with diabetes and their partners in Bromley, UK (Bromley sample). The mean age was 61.6 years, 46% were women and the mean age on leaving full-time education was 16.1. The first meeting was attended by 97 patients who mainly had NIDDM. The second meeting was attended by 45 patients who mainly had IDDM. The mean disease duration was 7.3 years. 38 patients were treated with insulin and diet, 33 by tablet and diet and 30 by diet only. 74 patients were without complications and 27 had complications present.

Reliability

Cronbach's alpha for the twelve item instrument was 0.84.

Validity

The ADDQoL was correlated with two single items and, as hypothesized, the mean scores correlated better with the diabetes-specific item ($r = 0.47$) than with the generic item ($r = 0.31$); both were highly significant. The correlations are well below 1, indicating that the ADDQoL provides information additional to that provided by the single items.

ADDQoL scores were significantly correlated with perceptions of hypoglycaemia ($r = 0.32$) and the number of complications ($r = 0.21$). As hypothesised, ADDQoL scores showed a greater negative impact of diabetes on quality of life for insulin-treated patients. This difference was significant for 7 out of 13 dimensions.

Responsiveness

Not assessed.

Acceptability

Of 67 patients approached in Cambridge, 52 (78%) returned questionnaires. The response rates for the Bromley sample were 62% and 93% for the first and second meetings, respectively. Missing data for the three items presumed to relate to everyone, namely physical activity, motivation, and enjoyment of food, ranged from 3% to 8%.

Commentary

The ADDQoL represents an alternative approach to the measurement of health outcomes. The instrument incorporates an importance rating for each of the items in a similar manner to that used by individualised instruments such as the Schedule for the Evaluation of Individual Quality of Life SEIQoL (O'Boyle, McGee et al., 1992) and the Patient Generated Index (Ruta, Garratt et al., 1994). However, in contrast to these instruments, the ADDQoL uses standardised items. The authors use the wide variation in the use of importance ratings, with at least 7.5% of respondents using each response option, as a justification for importance weightings.

The instrument has been developed with input from patients and health professionals which is evidence for content validity. There is also good evidence for the unidimensional structure of the instrument. The ADDQoL has good internal consistency but has not been assessed for test-retest reliability. Evidence for construct

validity was provided by the relationship with other quality of life and clinical variables.

Although the instrument has not been assessed for responsiveness, the authors comment that diabetes had greater reported impact on diabetes-specific items than on the standard quality of life items. This suggests that it will be more sensitive to change and responsive to subgroup differences. The authors also suggest that the ADDQoL will be more responsive to the effects of changes in the management of diabetes.

The authors comment that the Bromley sample cannot be considered representative of the wider diabetic community, since this was a particularly well-motivated and informed group of patients with diabetes. However, the pattern of responses was similar for the two samples.

The authors describe how the ADDQoL could be improved, including simplifying the wording changes and increasing the breadths of items. A short version of the questionnaire is also being investigated for use in clinical audit. Additional psychometric evaluations need to be undertaken including some aspects of validity, responsiveness and test-retest reliability.

Diabetes Care Profile/DCP

Purpose

The DCP is designed to measure the social and psychological factors important in a patient's adjustment to diabetes and its treatment, and is reported to be unique in its comprehensive coverage of these aspects (Fitzgerald, Davis et al., 1996). The authors suggest that selected scales within the DCP could be used as outcome measures. The DCP could also be used in clinical settings, or as a baseline measure in intervention studies. The instrument is designed for both types of diabetes.

Description

The DCP was derived from the Diabetes Education Profile DEP (Davis et al., op. cit.) and was assessed for the measurement properties of reliability and validity (Fitzgerald, Davis et al., op. cit.). The DCP has been compared to a generic instrument (Anderson, Fitzgerald et al., 1997) and the reliability of the instrument has also been assessed in African Americans (Fitzgerald, Anderson et al., 1998).

The DCP contains 234 items divided into 14 scales measuring social and psychological factors, as well as items concerning demographic information and self-care practices. The 14 scales include control problems (18 items), social and personal factors (13 items), positive attitudes (5 items), negative attitudes (6 items), self-care ability (4 items), importance of care (4 items), self-care adherence (4 items), diet adherence (4 items), medical barriers (8 items), exercise barriers (5 items), monitoring barriers (11 items), understanding management practice (10 items), long-term care benefits (5 items), and support attitudes (6 items). The remaining items relate to demographic factors. The review covers the scales that include items relating to symptoms and quality of life, control problems, social and personal factors, positive attitudes and negative attitudes.

Factor analysis and item content analysis was reported to have guided the development of the DEP (Davis et al., op. cit.) and this was supported by confirmatory factor analysis (Hess, Davis et al., 1986; Fitzgerald, Anderson et al., op. cit.). Items were added to the DEP that assess self-care and adherence. Item-total correlation and Cronbach's alpha were used to confirm final scale content.

Items within the scales of control problems, social and personal factors, positive attitudes, and negative attitudes use five-point adjectival scales. Scale scores are equal to the mean score of the items. Scores are not computed if half of the scale items are missing.

Patients

The evaluation of the reliability and validity of the DCP was based on two patient samples (Fitzgerald, Davis et al., op. cit.). The first comprised 440 patients being cared for in a community setting in the USA, of whom 11% had IDDM and 34% had NIDDM but were using insulin. Their mean age was 61 years (sd = 13), 55% were female, 32% had 13 or more years of education, the mean disease duration was 10 years, and GHb levels averaged 10.2 mg% (range 4.6-24.7 mg%). The second comprised 576 patients receiving care at a medical centre in Michigan, USA, evenly divided between those with IDDM, with NIDDM using insulin, and with NIDDM not

using insulin. The mean age of participants was 54 years (sd = 16), 60% were female, 44% had 13 or more years of education, and the mean disease duration was 14 years.

The DCP was compared with the SF-36 (Ware and Sherbourne, op. cit.) in a sample of 255 NIDDM patients of whom 64% did not receive insulin (Anderson, Fitzgerald et al, op. cit.). The mean age was 63.4 years, 55% were female, and the mean disease duration was 8.6 years.

The DCP has also been evaluated in African Americans (Fitzgerald, Anderson et al., op. cit.). Patients with NIDDM were recruited from six sites within the metropolitan area of Detroit, USA, and included 511 African Americans and 235 Caucasians as a comparison group.

Reliability

The internal consistency of the DCP was assessed for patients recruited in the community and a medical centre (Fitzgerald, Davis et al., op. cit.). Cronbach's alphas for the two samples were: control problems (0.86, 0.86), social and personal factors (0.85, 0.90), positive attitude (0.80, 0.80), and negative attitude (0.77, 0.75).

Anderson, Fitzgerald et al. (op. cit.) reported Cronbach alphas of 0.69-0.95 for the DCP scales with the exception of the exercise barriers dimension which had an alpha of 0.60.

Fitzgerald, Anderson et al. (op. cit.) found the DCP to be reliable amongst African Americans and Caucasians. Cronbach's alphas for the two samples were: control problems (0.89, 0.91), social and personal factors (0.91, 0.91), positive attitude (0.78, 0.82), and negative attitude (0.76, 0.79).

Validity

The construct validity of the DCP has been assessed through comparisons with instruments measuring related constructs, clinical and socio-demographic variables (Fitzgerald, Davis et al., op. cit.). DCP scores were correlated with scores for the Social Provisions Scale (Cutrona, Russell et al., 1986), CES-Depression Scale (Radloff, 1977) and Happiness and Satisfaction scale (Bryant and Veroff, 1986). These scales were chosen because they measure similar constructs to the DCP scales but without a diabetes focus. Correlations of greater than 0.30 were considered indicative of concurrent validity. The control scale was expected to have a positive correlation with the CES-D.

The social and personal factors scale was expected to correlate with all three instruments, negatively with the happiness and satisfaction and social provisions scales. The positive and negative attitude scales were expected to correlate with the CES-D and the Happiness and Satisfaction scale (positively with negative attitude, and negatively with positive attitude). The resultant correlations were as hypothesised, with the exception of the social and personal factors scale and the negative attitude scale, neither of which had correlation greater than or equal to 0.30 with the Happiness and Satisfaction scale. Of the predicted correlations reaching 0.30, the highest was -0.53 for the CES-D positive attitude scale.

Construct validity was further assessed in relation to diabetes type, treatment and metabolic control. Correlations greater than or equal to 0.20 were considered supportive of validity. The four scales of control problems, social and personal factors, positive attitudes, and negative attitudes were expected to produce scores reflecting worse symptoms and quality of life for patients with IDDM compared to those with NIDDM. Patients with NIDDM using insulin were also expected to have poorer scores than those not using insulin. Finally, it was hypothesised that the dimension of control problems would correlate with metabolic control.

Scores for the DHP scales of control problems, social and personal factors and positive attitudes reflected poorer symptoms and quality of life for the IDDM patients relative to NIDDM patients. Scores for the DHP scales of control problems, social and personal factors, and negative attitudes reflected poorer symptoms and quality of life for NIDDM patients using insulin, relative to NIDDM patients not using insulin. Both comparisons produced a statistically significant difference for the control problems scale. The comparison between the IDDM and NIDDM groups not using insulin produced a statistically significant difference. The comparison between the NIDDM group using insulin and the NIDDM group not using insulin produced a statistically significant difference for the positive attitudes scale. The four scales had significant levels of correlation with GHb levels ($r = 0.16-0.21$).

Fitzgerald, Anderson et al. (op. cit.) reported that insulin use had a main effect on the social and personal factors scale, indicating that patients using insulin felt diabetes had a greater impact on their life. Their results suggest that for patients with NIDDM, insulin use has a greater social and psychological impact than ethnic background.

Anderson, Fitzgerald et al. (op. cit.) assessed the relationship between the DCP and the Short Form 36-item Health Survey/SF-36, and the relationship between instrument scores, diabetic complications and glycaemic control. Eight of the DCP scales focusing on the management of blood glucose were omitted from the study.

Anderson, Fitzgerald et al. found a greater number of significant correlations among the scales for patients with NIDDM not using insulin. Separate correlations were calculated for IDDM and NIDDM patients. As expected, the DCP scale of social and personal factors correlated significantly with the SF-36 scales of role-physical ($r = 0.37-0.38$), social functioning ($r = 0.40-0.51$), and role-emotional ($r = 0.22-0.47$). The correlation with the SF-36 scale of role-emotional ($r = 0.22$) was not significant. As expected, the DCP scales of positive and negative attitude correlated with the SF-36 scale of mental health ($r = 0.35-0.44$ and $r = 0.40-0.40$).

For NIDDM patients using insulin, the DCP scales of social and personal factors and positive attitude were significantly correlated ($r = 0.32-0.35$) with the number of complications. Among NIDDM patients not using insulin, no DCP scales correlated with the number of complications. The DCP scale of self-care ability was strongly predictive of GHb levels; data are not reported for the remainder of the DCP scales.

Responsiveness

Not assessed.

Acceptability

The DCP takes 30-40 minutes to complete (Fitzgerald, Davis et al., op. cit.). In the community study, 1017 patients were invited to participate, 517 (50.8%) agreed to take part and 440 (85.1%) took part. In the medical centre study, 1500 patients were initially approached, 576 (38.4%) agreed to participate and 428 (74.3%) of these returned questionnaires. In the evaluation among African Americans and Caucasians, Fitzgerald, Anderson et al. (op. cit.) report a response rate of 66% for 746 participants; 24 questionnaires were returned incomplete.

Commentary

The DCP was based on the Diabetes Education Profile and includes additional items relating to self-care and adherence. The instrument is lengthy to complete. However, the majority of the scales measure psychological processes which are not reviewed here. This review has been restricted to the 42 items contributing to the scales of control problems, social and personal factors, positive attitudes and negative attitudes. The scales of the DEP were constructed using factor analysis but there is no published evidence for the factor structure of the DCP. The four scales reviewed were all found to have satisfactory internal consistency, although there is no published evidence for test-retest reliability. The instrument has undergone comparison with other well-established generic measures and was found to be associated with them in the manner expected. There is no published evidence for the responsiveness of the DCP.

Diabetes Health Profile/DHP-1

Purpose

The DHP-1 is a multidimensional self-completion instrument originally designed to identify psychosocial dysfunctioning among adult insulin dependent and insulin-requiring patients in an ambulatory care setting (Meadows, Steen et al., 1996). The instrument has also been adapted for use in non-insulin dependent patients (personal communication: in press). The authors believe that the disinhibited eating subscale might be appropriate as a screening tool for potential eating problems.

Description

The DHP-1 comprises 32 items covering three dimensions: psychological distress (14 items), barriers to activity (13 items), and disinhibited eating (5 items). The items use a four-point adjectival scale.

The content of the DHP-1 was derived following a literature review, a review of available instruments, interviews with 25 IDDM and insulin-requiring patients and discussions with diabetes health care professionals (ibid.). The interviews were analysed on the basis of thematic content which generated 95 items. Four judges (health psychologist, diabetologist and two diabetic liaison psychologists) independently grouped the items into five areas. All four judges allocated 81% of the items to the same five areas; the remainder were allocated following discussion. No additional content was suggested but some items were re-worded.

Following a survey of patients, 24 items with poor levels of endorsement and low or high levels of intercorrelation were removed from the instrument. The structure of the instrument was assessed in three samples of patients using principal axis factoring (PAF). The first PAF analysis showed that there were two additional factors to those hypothesised. The 16 items loading onto these factors were removed, together with 12 items with low factor loadings. The level of correspondence between composition of the three resultant factors, and item grouping carried out by the judges was found to be moderate but satisfactory.

Following the application of a forced three factor PAF analysis on the remaining 43 items, a further 11 items were removed that had either low factor loadings or high loadings on more than one factor. The remaining 32 items contributed to three dimensions labelled psychological distress (14 items), barriers to activity (13 items) and disinhibited eating (5 items). Item-total correlations were in the range 0.47-0.75 and all items had higher item-total correlations with their own dimensions than with the other dimensions. The PAF results were confirmed across sexes and age-groups, and when the sample was randomly split in two to form two separate subsamples. One final sample of patients confirmed the factor structure of the 32-item DHP in this evaluation.

Items within the three dimensions are summed and transformed to produce a score from 0-100 where 0 represents no dysfunction. All three dimensions showed a positive skew (less dysfunctioning) and less than six percent of patients scored at the floor or ceiling on any dimension. The range of responses recorded varied between 0-85.7 for psychological distress and 0-100 for disinhibited eating.

The instrument was translated into Dutch using the forward backwards method (Goddijn et al., 1996). The structure of the DHP was assessed in a sample of Dutch NIDDM patients using a three factor forced PAF. One item within the barriers to activity dimension was removed because it was not relevant to NIDDM. Eight items loaded onto different factors to those previously found, seven of which came from the psychological distress dimension.

Three items had poor factor loadings and six items had high loadings on two factors. The original and the new factor structure correlated as follows: psychological distress ($r = 0.83$), barriers to activity ($r = 0.94$) and disinhibited eating ($r = 0.91$). Mean scores for all dimensions of the DHP were high but the data was not abnormally skewed.

Patients

The development and evaluation of the instrument used three samples of patients recruited from outpatient sites in England (Meadows et al., op. cit.).

The first sample comprised 278 IDDM patients and insulin-requiring patients aged 20-65 years registered at a hospital outpatient department. The mean age of responders was 40.9 years ($sd = 13.0$) and the mean disease duration was 13.7 years ($sd = 9.0$). There were no significant differences between responders and non-responders in terms of age, sex and disease duration.

The second sample comprised 2239 IDDM patients and insulin-requiring patients recruited at 54 hospital outpatient departments in England and Wales. The mean age of responders was 39.8 years ($sd = 10.0$), 51% were male, and mean disease duration was 13.1 years ($sd = 10.0$).

The third sample comprised 295 IDDM patients over 18 years of age requiring insulin, recruited as a convenience sample from seven hospital outpatient department in North-East England. The mean age of responders was 51.5 years ($sd = 17.3$) and 52% were male. Women were found to be significantly more likely to respond than men.

The psychometric properties of the DHP were assessed among 99 consecutive Dutch patients with NIDDM referred to a hospital outpatient department (Goddijn et al., op. cit.). The mean age was 61.2 years ($sd = 10.9$), 51.5% were female, mean disease duration was 8.5 years ($sd = 7.3$), and mean haemoglobin levels were 10.4 ($sd = 2.7$). 78.7% had at least one complication.

Reliability

Cronbach's alphas for two of the samples in which the instrument was developed were: psychological distress (0.85-0.86), barriers to activity (0.82-0.85) and disinhibited eating (0.77-0.80) (Meadows et al., op. cit.). The Dutch study reported somewhat lower values: psychological distress (0.72-0.77), barriers to activity (0.79) and disinhibited eating (0.72) (Goddijn et al., op. cit.).

Validity

The authors state that the methods of item derivation and dimension development are evidence of satisfactory face and content validity (Meadows et al., op. cit.). DHP

scores were compared with those for the Hospital Anxiety and Depression Scale/HADS (Zigmond and Snaith, 1983) and the SF-36 (Ware and Sherbourne, op. cit.). Correlations were in the range 0.17-0.68 and all were statistically significant. As hypothesised, the highest correlations were found between the psychological distress and barriers to activity dimensions, and the HADS and SF-36.

The authors also hypothesised that women would score higher than men on the psychological distress and disinhibited eating dimensions. These predictions were in part supported in one of the samples, with women under 40 scoring significantly higher than men on the psychological distress dimension, and women 65 years and under scoring significantly higher than men on the disinhibited eating dimension. In another smaller sample, the psychological distress dimension did not significantly differ between women and men, but women had a significantly higher mean score on the disinhibited eating dimension.

The Dutch study hypothesised strong correlations (>0.40) between the psychological distress and barriers to activity dimensions of the DHP, and the social functioning, mental state, vitality and general health perceptions of the SF-36 (Goddijn et al. op. cit.). No strong correlations were expected between the disinhibited eating dimension and any aspects of the SF-36. The hypotheses were supported by the data, the only exception being the low level of correlation between psychological distress and the SF-36 scale of general health perception ($r = 0.33$). Chronic diabetes complications and comorbidity correlated only moderately with the SF-36 dimensions of physical functioning and pain, and not at all with the DHP.

As hypothesised, younger patients had worse scores on the disinhibited eating subscale. Contrary to expectations, the authors did not find any significant differences between women and men on the disinhibited eating subscale, but stated that this could be due to the relatively older patient sample when compared to the study by Meadows et al. (op. cit.). The DHP psychological distress and barriers to activity dimension scores were significantly related to the hyperglycaemic complaint of fatigue. No significant association was found between the DHP scores and HbA_{1c} levels.

Responsiveness

The DHP-1 has not been formally assessed for responsiveness. However, the dimensions of psychological distress and barriers to activity within the earlier version of the DHP have been assessed for responsiveness (Whitty et al., 1997). Following a literature review and discussions with clinicians, it was hypothesised that changing NIDDM patients to insulin treatment should result in improvements in psychological distress and energy. The psychological distress and barriers to activity dimensions produced standardised response means (SRM) of 0.23 and 0.02 at six weeks follow-up, compared to an SRM of 0.85 for the Newcastle Diabetes Symptoms Questionnaire (see relevant review). Smaller SRMs were found at three months follow-up.

The Dutch version of DHP for NIDDM patients was used alongside the SF-36 as part of a longitudinal study of glycaemic control (Goddijn, Bilo et al., 1999). Decreases in mean HbA_{1c} levels from 10.4-7.8% accompanied statistically significant improvements in the DHP subscales of psychological distress and disinhibited eating, and the SF-36 scales of social functioning, role limitations due to emotional problems, and vitality.

Acceptability

Two of the samples recruited for the development of the DHP-1 produced response rates of 79.0-86.0%; anonymity meant that the response rate could not be calculated for one of the samples (Meadows et al., op. cit.). In the larger sample of 2239 patients, all 43 items were answered by 84.85% of the sample, with a significant association between lower completion rate and increasing age.

Commentary

The DHP-1 has undergone a process of development involving patients in item derivation and principal axis factoring to assess dimensionality (Meadows et al., *ibid.*). Two versions of the instrument have been developed for both types of diabetes, although the version developed for non-insulin dependent patients had not yet been published at the time of writing (Meadows, personal communication). The instrument has good evidence for internal consistency but test-retest reliability was not reported. Significant correlations between DHP dimension scores and those for the HADS and SF-36 support the validity of the instrument.

The responsiveness of the DHP-1 has not been formally assessed using an effect size statistic. An earlier version of the instrument produced a small SRM for the psychological distress scale, but there were no significant changes in scores for NIDDM patients switching to insulin, despite there being significant improvements in HbA_{1c} (Whitty et al., op. cit.). The Dutch version of the instrument produced significant score improvements for NIDDM patients with good metabolic control (Goddijn et al, 1999 op. cit.).

Diabetes Health Status Questionnaire/DHS

Purpose

The Diabetes Health Status Questionnaire (DHS) is designed to measure symptoms and general perceptions of health (Wierenga, op. cit.). The instrument was developed for a longitudinal study assessing the effectiveness of a lifestyle modification programme for weight control in NIDDM patients.

Description

The DHS comprises 26 items, 20 of which relate to the frequency of diabetes-related physical symptoms in the last six weeks. The remainder ask about physical and emotional health, current health compared with health before diabetes was diagnosed, and health compared with that of others of the same age. The items use five-point adjectival scales and sum to produce a single score where higher scores represent poorer health (ibid.).

The proportion of patients scoring at the floor and ceiling for the most frequently reported diabetes symptoms were in the range 36.8%-66.2% for the floor (“none or no symptoms”) and 5.9%-30.9% (“5 or more”) for the ceiling. Floor and ceiling effects for the health description item were 2.9% and 60.3%, respectively. Floor and ceiling effects for the physical health item were 23.5% and 4.4%, respectively. Floor and ceiling effects for the emotional health item were 22.1% and 1.5%, respectively. Floor and ceiling effects for the comparison with diagnosis before NIDDM (“how do you feel”) were 42.6% and 4.4%, respectively. Floor and ceiling effects for the comparison with diagnosis before NIDDM (“energy level”) were 22.1% and 14.7%, respectively.

Patients

Patients who had participated in an education programme in the past year were asked to participate in a joint community/university development to develop and test a community-based intervention for lifestyle modification. The study was based in Milwaukee, USA. The initial sample consisted of 66 NIDDM patients aged 30-86 years. 93.9% were white, 77.3% were married and had annual incomes between \$10,000 and \$30,000, 39.7% were high school graduates, and 36.8% had received higher education. 40% of patients had been diagnosed with diabetes for less than one year.

Reliability

Cronbach’s alpha for the DHS was 0.89.

Validity

Not assessed.

Responsiveness

The responsiveness of the instrument was not formally assessed using an effect size statistic. Instrument scores did not significantly improve following a lifestyle modification program for diabetes (ibid.).

Acceptability

The follow-up DHS questionnaire was returned by 48 (72.7%) patients. There were no differences in demographic characteristics between those patients completing and those not completing the study.

Commentary

The development of the DHS, including item derivation, is not described. The dimensionality of the instrument has not been assessed empirically and no consideration was given to validity testing. There are considerable end effects for some of the items, and the author comments that the large number of patients with high scores may have limited the instrument's ability to measure change in the longitudinal study (*ibid.*).

Diabetes Impact Measurement Scale/DIMS

Purpose

The DIMS is designed to measure longitudinal changes in health status in IDDM and NIDDM patients to quantify treatment benefit in clinical trials. The developers suggest that it may be suitable for other purposes, such as comparisons across groups of diabetes patients (Hammond and Aoki, op. cit.).

Description

The DIMS comprises 44 items covering the following domains: physical functioning, symptoms, psychological well-being, social well-being, role activities, personal constructs, and impact of treatment. Items use four- to six-point adjectival scales. The instrument is scored so that higher values represent better health. The authors suggest that each subscale be considered a separate indicator of outcome, with the average score of the major subscales being used as an overall index of diabetes impact.

The DIMS was developed following a literature review and the authors drew upon existing instruments, particularly the RAND instruments (Ware, Johnston et al., 1979), the Sickness Impact Profile (Bergner, Bobbitt et al., 1981) and the Arthritis Impact Measurement Scales (Meenan, Gertman et al., 1980). Forty four items were derived from discussions with clinicians (physicians, a diabetes nurse and a dietician). These focus on common and significant symptoms, patients' attitudes towards the disease and management, patients' abilities to fulfil social roles, and their general well-being. Items were grouped into four dimensions: symptoms, including those specific to and those less specific to diabetes, diabetes-related morale, social role fulfilment, and well-being. The language of the items was estimated to require 6th grade reading ability. Questionnaires with missing values were not excluded from the analysis; the authors suggest methods for assigning values to missing responses.

The structure of the instrument was assessed using principal component analysis, and a major component accounting for 32% of the variance was found along with nine components accounting for less than 7.5% of the variance. There was no statistical support for the unique significance of the hypothesised dimensions of the DIMS. The hypothesised dimensions were highly correlated with the first component ($r = 0.75-0.95$). Four items had poor item-total correlations which would be removed from the instrument in future.

Patients

130 patients at baseline and 52 patients at follow-up were recruited from a diabetic clinic at the University of California, USA. The mean age was 45 years ($sd = 15$) and 58% of the sample were women. 39% of patients were diagnosed as having IDDM, 59% as having NIDDM. Mean disease duration was 11 years. 77% of patients were taking insulin, 22% were taking oral hypoglycaemic agents. 39% of patients had some form of retinopathy, 28% had nephropathy, and 30% had neuropathy. Mean HbA_{1c} values were 7.6 ($sd=1.8$).

Reliability

Cronbach's alpha for the six scales were in the range 0.60-0.85; the scale of specific symptoms fell below 0.70. Test-retest questionnaires were completed by 52 patients after a minimum of one month and correlations were in the range 0.61-0.78. The

authors state that the intervals varied between test and retest for each subject. The scales of non-specific symptoms, well-being, and social role fulfilment fell below 0.70.

Validity

The instrument was validated through comparisons with global rating scales and clinical variables. It was hypothesised that higher DIMS scores would generally be negatively correlated with clinical variables indicating the presence of disease. However, no a priori hypotheses were made regarding the level of correlation.

The questionnaire included two global visual analogue scales relating to diabetic control and general health respectively. Clinicians completed the same scales on behalf of the patient. Correlations between the DIMS scales and the diabetes control scale were in the range 0.22-0.55 and 0.24-0.38 for the patient and clinician ratings, respectively. Correlations between the DIMS scales and the general health scale were in the range 0.27-0.47 and 0.29-0.45 for the patient and clinician ratings, respectively.

The DIMS scales of non-specific symptoms, combined symptoms, well-being, diabetes related morale and total DIMS scores were correlated with HbA_{1c} levels. There were no significant correlations between the DIMS scores and the main diabetes-related complications, although correlations were observed for the symptom dimension. DIMS scores were not significantly correlated with the diabetes complications index (sum of all complications). Age was positively correlated with two scales and sex was correlated with three scales, as well as the overall DIMS score. Significant correlations in the direction hypothesised were found between the DIMS scales and a number of diagnoses.

Responsiveness

Not assessed.

Acceptability

The questionnaire takes an average of 15-20 minutes to complete. Items were easily answered by most patients, though some items were more frequently left unanswered than others; the authors provide some explanation for this. Every item was answered by 90 (69.2%) patients from the sample of 130 patients. Four items had relatively high rates of missing values, particularly items relating to sexual functioning and libido.

Commentary

The DIMS is designed for use in patients with IDDM and NIDDM, and was developed following a review of the literature and consultation with clinicians. The DIMS consists of six scales, the existence of which was not supported empirically. There was some evidence for the existence of a single dimension and the authors included a total DIMS score in their analyses. The instrument has evidence for internal consistency, but the scale of specific symptoms fell below the level of reliability necessary for groups. There was limited evidence for test-retest reliability but there was a large gap between test and retest, and no attempt was made to identify whether health had changed. The interpretation of the results of validity testing is limited by the lack of a priori hypotheses regarding the size of correlations. The instrument has not been assessed for responsiveness.

Diabetes Quality-of-Life Measure/DQoL

Purpose

The DQoL was originally developed for use in a clinical trial comparing the efficacy of two different treatment regimens on the appearance and progression of chronic complications of patients with IDDM. However its structure allows for application to other patients with IDDM and NIDDM (Jacobson, 1997). The developers state that the DQoL could be used in clinical settings as a screening measure to identify patients with concerns about diabetes (Jacobson, Barofsky et al., 1988).

Patients are asked to rate the common issues of having diabetes in terms of their current functioning. The developers suggest that since the questions are posed from three perspectives (the impact generated by diabetes, the patient's satisfaction with him/herself, and worry about anticipated effects of diabetes) the DQoL can be considered a battery instrument.

With the exception of the worry: social/vocational subscale among NIDDM subjects and the satisfaction/worry interconnection among IDDM patients in the first study, all correlations between the subscales were significant at the $p \leq 0.01$ level. Correlations ranged from 0.37 (satisfaction and worry: social/vocational) to 0.68 (worry: diabetes related and worry: social/vocational). Item-total correlations for NIDDM patients were of the order 0.52-0.93 and for IDDM patients, 0.66-0.89.

Description

The development and initial validation of the DQoL was carried out by the Diabetes Control and Complications/DCCT Research Group (Jacobson, Barofsky et al., op. cit.). It was subsequently evaluated in two studies comparing it with generic instruments (Parkerson, Connis et al., 1993; Jacobson, De Groot et al., 1994). There have been two reviews of the instrument (Jacobson and DCCT Research Group, 1994; Jacobson, op. cit.). The Diabetes Quality of Life and Clinical Trial Questionnaire/DQLCTQ includes the DQoL, and evidence on the DQoL's reliability and validity is also available from this study (Shen, Kotsanos et al., 1999).

The content of the DQoL was derived from the following three sources: a literature review identifying the concerns of diabetic patients and problems that impact on their lives, consultation with clinicians knowledgeable about diabetes, and patients with IDDM. The meaning, relevance and readability of the instrument was assessed during its development by giving draft versions to IDDM patients; drafts were also reviewed by health professionals. The initial item pool comprised items considered to be of greatest relevance to patients with IDDM undergoing treatments of differing intensity (Jacobson, Barofsky et al., op. cit.).

The instrument has 46 core items forming four scales: satisfaction with treatment (15 items), impact of treatment (20 items), worries about future effects of diabetes (4 items), and worries about social and vocational issues (7 items). The instrument also includes a generic health item that does not contribute to the scales. Adolescent and youth versions of the DQoL have been developed (Jacobson, Hauser et al., 1997).

The dimensions and DQoL total scores (average score across the four dimensions) are scored 0-100 where 0 represents the lowest possible quality of life and 100 the

highest. It is recommended that the diabetes and social worry scales be deleted if two or more of the item scores are missing. For the satisfaction scale, this limit rises to three missing items and for the impact scale, the limit rises to four. No information is given about how DQoL dimensions were constructed.

Patients

The original development and evaluating of the DQoL included 210 IDDM patients with similar characteristics to those in the DCCT trial, as judged by the individual participating centres (Jacobson, Barofsky et al., op.cit.). 192 patients agreed to participate; the mean age was 28 years (sd = 7), 136 were adults (114 men and 76 women) who were primarily white. The patients were from Hollingshead social classes I-IV and 40% of the sample were married. The mean disease duration was 8 years (sd = 4.3) and all patients were receiving insulin injections. Patients were chosen who had no clinically evident complications.

The DQoL was compared with generic instruments in a sample of 170 adult insulin-dependent diabetic patients attending various medical clinics in the USA (Parkerson et al., op. cit.). The mean age was 33.7 years (sd = 8.7) and approximately half were women. The study population was white and the largest group of patients (60.3%) had received 13-16 years of schooling. Most were married and living with their spouses (63.9%) and the mean disease duration was 15.6 years (sd = 9.8). Just over half the sample had at least one diabetic complication and around one-third had comorbid conditions. Varying insulin treatment regimens (conventional insulin, intensified insulin, insulin pump) were fairly evenly represented. Those completing and not completing the DQoL were similar, except in one respect: those not participating were more likely to be receiving intensive treatment.

The DQoL was also compared to a generic instrument in a more heterogeneous sample of 240 consecutive patients attending an outpatients clinic in the USA, of whom 111 had IDDM and 129 had NIDDM (Jacobson, de Groot et al., 1994 op. cit.). The mean age was 44 (se = 16) for the IDDM subjects and 60 (se = 12) for the NIDDM subjects. 53% of IDDM patients and 49% of NIDDM patients were women; 63% of IDDM patients and 69% of NIDDM patients were married. Mean disease duration was 18.8 years (se = 11.5) for the IDDM patients and 12 years (se = 8) for the NIDDM patients. Just over half the NIDDM patients were taking insulin, 38% of the NIDDM patients were on oral-hypoglycaemic agents and 9% were treated by diet alone. Diabetic complications were present, the most common being neuropathy which was found in 35.1% of IDDM patients and 48.8% of NIDDM patients.

Details of the sample used in the development of the Diabetes Quality of Life Clinical Trials Questionnaires (Shen et al., op. cit.), which includes the DQoL as part of a battery of instruments, can be found in the relevant review of this instrument.

Reliability

The DCCT Research Group (Jacobson et al., 1988 op. cit.) reported Cronbach's alpha coefficients ranging from 0.69 to 0.92 for the scales of diabetes-related worry and total scores respectively. Only the former dimension fell below 0.70. Parkerson et al. (op. cit.) reported alpha values in the range 0.52-0.88 for the diabetes worry and total DQoL scores, respectively. Jacobson, de Groot et al. (1994, op.cit.) reported alpha values in the range 0.47-0.87 for patients with IDDM and NIDDM. With the

exception of the diabetes worry scale ($r = 0.47-0.49$), the reliability estimates were regarded as similar to those reported in previous studies. Shen et al. (op. cit.) reported alpha coefficients ranging from 0.53 to 0.86 for the social worry and satisfaction scales, respectively.

Test-retest reliability was assessed by asking patients to complete a second questionnaire at a mean of nine days after it was first administered. Pearson correlations were in the range 0.78-0.92 for the social/vocational worry and total scores respectively (Jacobson, Barofsky et al., op. cit.). Shen et al. report intraclass correlation coefficients ranging from 0.56 to 0.80 for the diabetes worry and impact scales, respectively.

Validity

The DQoL items were derived from IDDM patients and clinicians, together with the literature on psychosocial aspects of diabetes. Selected patients, as well as clinicians, then reviewed the items for content relevance. On the basis of patient input, the instrument was expanded to include worries about the future (Jacobson, Barofsky et al., op. cit.). The content of the DQoL was subsequently reviewed but methodological details were not given (Parkerson et al., op. cit.).

In the original evaluation, the DQoL was compared with three instruments: the Symptom Checklist-90-R/SCL, the Bradburn Affect Balance Scale/BABS, and the Psychosocial Adjustment to Illness Scale/PAIS (Jacobson, Barofsky et al., op. cit.). Several hypotheses were constructed. First, the DQoL worry scales would have larger correlations with the SCL total score than the PAIS and BABS. Second, the DQoL worry scales would have similar levels of correlation with the BABS and PAIS. Third, the DQoL satisfaction scale would have the largest correlation with the BABS. Fourth, the DQoL impact scale would have the largest correlation with the PAIS scales, with the exception of the PAIS distress scale. Finally, the DQoL total scores would have significant correlations with all instrument scores and the DQoL scales would have positive correlations with all instrument scores. Correlations were expected to fall within the range 0.3-0.7, indicating that constructs were similar but not identical.

The two DQoL worry scales were significantly correlated with the SCL total score ($r = 0.40-0.50$) and these were stronger than all correlations with the PAIS and BABS except for the PAIS scale of psychological distress ($r = 0.46$). The DQoL worry scales had similar low levels of correlation with the ABS and the PAIS except for the aforementioned psychological distress scale of the PAIS. The DQoL satisfaction scale had a significant correlation ($r=-0.55$) with the ABS but had a slightly larger correlation with the PAIS scale of health-care orientation. The DQoL impact scale did produce the largest correlations with the PAIS scales although the PAIS psychological distress scale correlated more highly with the DQoL impact scale than expected. Finally, the DQoL total scores did have significant correlations with all the instrument scores and all correlations were positive.

The DCCT study found two small but significant associations with sex: women reported DQoL scores reflecting a greater impact of diabetes and greater diabetes-related worries (Jacobson, Barofsky et al., op. cit.).

Two studies have compared the DQoL with generic instruments. The first compared the DQoL with the Duke-UNC Health Profile/DUHP, the General Health Perceptions Profile/GHP, and the Health and Daily Living Form/HDL (Parkerson et al., op. cit.). There were no formal hypotheses but the authors expected DQoL scores to explain greater variance in disease indicators than scores for the generic instruments. Of the disease indicators (duration of diabetes, complications and intensity of treatment), only the complications variable was a statistically significant predictor.

The DQoL total scores had 28% of their variation explained by four comorbidity and psychosocial variables. The DQoL social/vocational worry dimension had the most variance explained (41%) by these variables. The impact dimension had the least variance explained (12%). Similar analyses of a modified DQoL that separated the instrument into generic and disease-specific components found that more variance was explained by the generic component. Neither of the modified scales had a statistically significant relationship with the diabetes-related variables.

In a stepwise regression analysis, sex and age did not enter the equation when DQoL total scores and satisfaction, impact and diabetes worry scales were the dependent variables. However, age did enter the equation when social/vocational worry was the dependent variable. Age was predictive of less social worry. Marriage entered the equation when the two DQoL worry dimensions were dependent variables: being married was predictive of less worry and better mental health.

The second study compared the DQoL with the SF-36 scales of physical functioning, social functioning, role limitations due to physical problems, pain, and general health perception (Jacobson, de Groot et al., 1994 op. cit.). The total DQoL had small to moderate levels of correlation with the SF-36 scales ($r = 0.33-0.60$). The DQoL scales of satisfaction and impact had the largest correlations with the SF-36 scales, ranging 0.28-0.50 and 0.30-0.59 respectively.

This study also assessed the relationship between the DQoL and complications using regression analysis, after adjusting for sociodemographic factors. The DQoL impact and satisfaction scales, and total scores had a significant relationship with the number of complications among patients with IDDM. The DQoL total scores, and impact, satisfaction and diabetes worry scales had a significant relationship with the severity of diabetes among patients with IDDM. The DQoL satisfaction scale had a significant relationship with the number of complications in patients with NIDDM. The DQoL impact and satisfaction scales, and total scores had a significant relationship with the severity of diabetes among patients with NIDDM. DQoL total scores were significantly correlated with age ($r = 0.34$). Separated or divorced patients were found to experience worse quality of life than their counterparts, but data were not presented.

In the evaluation of the DQLCTQ, it was hypothesised that, relative to their counterparts, IDDM patients, men, patients with tight metabolic control, and patients who had good self-perceived control would have better scores on DQLCTQ, of which the DQoL is a component (Shen et al., op. cit.). Patients with IDDM were found to have significantly better scores for all but the scale of social worry. Patients with tight metabolic control were found to have significantly better scores for all but the social worry scale. Male patients were found to have a significantly better scores for the

scale of diabetes worry. Patients with good self-perceived control had significantly better scores for all four scales.

Responsiveness

The responsiveness of the DQoL has not been formally assessed but the instrument's authors cite two studies as evidence for the responsiveness of the instruments (Jacobson, *op. cit.*). In the first, patients with end-stage renal disease were given either a kidney transplant or a combined pancreas/kidney transplant. There was a significant improvement in the DQoL total scores and all subscales in patients who received the combined transplant, while there was no improvement for those receiving the kidney transplant alone (Nathan, Fogel et al., 1991). The second study compared the quality of life of patients who received an implantable pump with those receiving normal insulin treatment (Selam, Micossi et al., 1992). The DQoL scale of satisfaction showed an improvement but there were no other changes.

The DQoL was also assessed for responsiveness as a part of the Diabetes Quality of Life in Clinical Trials Questionnaire which is reviewed below (Shen et al., *op. cit.*). The responsiveness of the instrument was assessed by mean changes in scores for patients whose metabolic control had improved or worsened over six months. For the improved group, the scales of DQoL satisfaction and treatment satisfaction scale produced significantly better scores compared to baseline. For the worsened group, the DQoL did not produce any significant changes.

Acceptability

Information relating to the acceptability of the DQoL is available only for the two studies reporting comparisons with generic instruments. In the first, 131 out of 179 IDDM patients completed the DQoL and there were no missing items (Parkerson et al., *op. cit.*). The analysis was limited to those patients completing the DQoL. There were no significant differences between responders and non-responders to the DQoL for any of the demographic, psychosocial or comorbidity variables collected. There were also no significant differences for disease duration and complications.

However, a significant difference was found for intensity of treatment and 79.5% of non-responders were insulin-pump patients. The second study reported that 88% of patients agreed to participate (Jacobson, De Groot et al., 1994 *op. cit.*). There were differences in responses to DQoL subscales, reflecting the fact that the social/vocational worry subscale is less suitable for older NIDDM patients. The responses were different for the subscales of satisfaction (n = 228), diabetes worries (n = 219), impact (n = 217), and social/vocational worries (n = 61).

Commentary

The content of the DQoL was derived from a literature review with input from patients and clinicians, which lends support to the content validity of the instrument. The DQoL is designed to be relevant to both IDDM and NIDDM patients, and the instrument has been tested with both groups of patients. However, the patients contributing to the instrument content were IDDM patients.

The developers do not adequately describe the methodology and results of item selection. Guyatt's methodology (Guyatt, Veldhuyzen Van Zanten et al., 1989) is referred to, but the results of this approach are not presented. For example, the

developers do not provide information relating to the initial item pool nor the means by which this was reduced. There is no evidence that the structure of the instrument has been assessed empirically, for example using principal component analysis.

The DQoL scale of diabetes worry has consistently poor levels of internal consistency, making it unsuitable for use with groups of patients. The results of the test-retest reliability indicate that the scales are suitable for group comparisons in IDDM patients. Two of the dimensions, namely satisfaction and impact, are close to being suitable for use in individual patients.

The DQoL has been assessed for validity in patients with IDDM and NIDDM. Significant small to moderate correlations were found between the DQoL dimensions and a number of generic instruments, including the SF-36 (Parkerson et al., op. cit., Jacobson, de Groot et al., 1994 op. cit.). However, in assessing the relationship between instrument scores and clinical, psychosocial and demographic variables, Parkerson et al. concluded that the generic instruments outperformed the DQoL. Jacobson and the DCCT group (op. cit.) provide some evidence that complications are significantly correlated with DQoL dimensions. The DQoL has not been formally assessed for responsiveness though the results of two studies provide some anecdotal evidence (Jacobson, op. cit.).

Diabetes Quality of Life Clinical Trial Questionnaire/DQLCTQ

Purpose

The DQLCTQ was developed to be a valid and reliable measure of health-related quality of life, for use in multinational clinical trials of patients with IDDM and NIDDM (Shen et al., op. cit.).

Description

The DQLCTQ was developed and published alongside a revised version of the instrument referred to as the DQLCTQ-R. The DQLCTQ comprises 142 items across 20 domains, 3 self-efficacy questions and 4 demographic questions. For the most part, items use five-point adjectival scales. The DQLCTQ-R, comprises 57 items across 8 domains.

Patient focus groups (30 patients) and expert clinician panels (11 clinicians) in the USA and France identified domains of importance. The literature was then reviewed to find generic and disease-specific instruments containing these domains. The major components of the draft instrument were based on these findings, and data were extracted from validated generic and disease-specific instruments. The former included the SF-20 (Stewart, Greenfield et al., 1989) and SF-36 (Ware and Sherbourne, op. cit.), while the latter included the Diabetes Quality of Life Measure/DQoL (Jacobson, Barofsky et al., 1988 op. cit.) and the Hypoglycaemia Fear Survey/HFS (Cox, Irvine et al., 1987). Instruments dealing with social stigma, treatment satisfaction and symptoms were not available and items were therefore developed. This process produced 293 items that were assessed for face and content validity by a group of researchers expert in the measurement of health-related quality of life.

The draft instrument was evaluated in patients attending five internal medicine practices and diabetes care centres in the USA. Following this study, and using the results of the focus groups and clinician panels, the instrument was reduced to make it acceptable for multinational clinical trials. Redundant items and domains were removed and domains with poor psychometric properties were modified or removed. Two domains were created for insulin-specific comparisons, treatment satisfaction and treatment flexibility.

Each domain is scored on a scale of 0-100 where higher scores represent better quality of life. Summary scores were calculated for each domain. The instrument was translated into German and French following Guyatt's recommendations (Guyatt, 1993).

The instrument was revised following tests for reliability and validity. Single item dimensions were excluded because of high variability of responses and low reliability. The social worry and diabetes worry dimensions of the DQoL were excluded because of poor reliability. The worry scale of the HFS was excluded because it failed to discriminate between IDDM and NIDDM patients, and between good and poor metabolic control.

The bothersomeness of symptoms domain was excluded because it correlated strongly with the frequency of symptoms dimension and was not as responsive. The impact

scale of the DQoL was excluded because it correlated well with the satisfaction scale of the DQoL and was not as responsive. This revised instrument, the DQLCTQ-R, comprises 57 items across eight domains: physical function, energy/fatigue, health/distress, mental health, satisfaction, treatment satisfaction, treatment flexibility, and frequency of symptoms.

Patients

The initial pilot sample included 123 patients, 59% of whom had NIDDM. The instrument was further evaluated with 942 patients of whom 63.5% were from the USA, 20% from Germany, 8.9% from France and 7.6% from Canada (Shen et al., op. cit.). Patients with IDDM and NIDDM were enrolled in separate clinical trials and a pooled analysis was performed as the psychometric properties were similar between the two groups. The mean age of patients was 33.8 and 58.2 years for IDDM and NIDDM patients, respectively, and 409 (43.4%) patients were women. Of the 942 patients, 468 (49.7%) had IDDM, mean disease duration was 12.6 years, and mean baseline HbA_{1c} levels were 8.6 (sd = 1.6).

Reliability

Cronbach's alpha coefficients for the DQLCTQ domains were greater than 0.70, with the exception of the DQoL dimensions of social worry (0.62) and the DQoL diabetes worry (0.53). The newly developed domains produced alpha coefficients ranging from 0.77 to 0.89 for the frequency of symptoms and treatment flexibility domains, respectively.

Test-retest reliability was assessed 7-10 days after baseline among the initial pilot sample of patients. Intraclass correlation coefficients were in the range 0.49-0.90 for the social stigma and health distress dimensions, respectively. The diabetes worry and the social stigma dimensions produced coefficients below 0.70. Revisions to the instrument meant that the test-retest reliability was not reported for the newly created domains of treatment satisfaction, treatment flexibility, frequency of symptoms or bothersomeness of symptoms. The revised version of the instrument, the DQLCTQ-R, has good levels of reliability with alpha and test-retest coefficients all above 0.70 (Shen et al., op. cit.).

Validity

The draft instrument was assessed for face and content validity by a group of researchers with expertise in the measurement of health-related quality of life. The DQLCTQ was further assessed for validity through comparisons with clinical and sociodemographic variables. On the whole, the hypotheses were supported by the data. Patients with good metabolic control had significantly higher mean DQLCTQ scores than those with poor metabolic control. Patients who considered themselves to be in good control of their diabetes had significantly higher mean DQLCTQ scores than those who felt they were in poor control. With the exception of the domains of social worry, worry of the HFS, treatment satisfaction, and treatment flexibility, patients with IDDM had higher mean DQLCTQ scores than those with NIDDM. With the exception of the dimensions of satisfaction, impact, and social worry and social stigma, women patients had poorer mean DQLCTQ scores than men.

Responsiveness

The responsiveness of the DQLCTQ was assessed by mean changes in DQLCTQ scores for patients whose metabolic control had improved or worsened over six months. For the improved group, the DQoL satisfaction and treatment satisfaction scales produced significantly better scores compared to baseline. For the worsened group, the mental health scale produced a significantly worse score compared to baseline. The DQLCTQ-R scales of treatment satisfaction, health/distress, mental health, and DQoL satisfaction produced significantly better scores compared to baseline for the improved group.

Acceptability

The DQLCTQ-R can be completed within ten minutes. Less than 10% of items were missing for 83% of questionnaires administered (Kotsanos, Vignati et al., 1997).

Commentary

The development of the DQLCTQ tried to ensure that all domains of importance to patients and clinicians were included, and that the draft instrument had content validity, as assessed by a group of quality of life experts. The internal consistency and test-retest reliability of some of the DQLCTQ scales failed to achieve the levels of reliability required for group comparisons. The revised instrument produced satisfactory estimates but data were not presented. Evidence for the validity of the DQLCTQ was provided through comparisons with clinical variables, but the instrument was not compared to other generic and specific instruments measuring related constructs. The DQLCTQ and DQLCTQ-R have not been formally assessed for responsiveness, but some of the scales showed higher levels of quality of life at follow-up for patients whose metabolic control had improved.

The DQLCTQ and the revised shorter-form represent a battery approach to measurement, and incorporate domains from existing instruments and newly developed scales. The reader is referred to the review of the DQoL for further evidence of the measurement properties of this component. There are existing short-form disease-specific and generic instruments that, used in conjunction for assessing outcomes, perform a similar task to the DQLCTQ. The inclusion of a generic instrument in its entirety, as compared to selected domains, improves comparability across studies and is likely to be of greater usefulness for purposes of economic evaluation.

Diabetic Quality of Life Questionnaire/DQLQ

Purpose

The Diabetic Quality of Life Questionnaire was developed to assess the extent to which taking insulin and being diabetic affected patient quality of life (Henderson and Tindall, op. cit.).

Description

The questionnaire comprises ten items. Eight items assess how much being diabetic and taking insulin interferes with day-to-day life, including discomfort and inconvenience. Two items assess confidence regarding insulin doses and concern about evidence of hypoglycaemia. All items use ten-point scales ranging from “not at all” to “a great deal”, “not at all” to “extremely so”, or from “very easy” to “extremely difficult”.

Patients

The instrument was administered to two groups of patients. Twelve IDDM patients volunteered to test the Novopen II, and 20 IDDM patients attending the same outpatient clinic were selected at random as a control group. The mean age of the Novopen patients was 61.2 years (sd = 7.8), 25% were female, the duration of diabetes was 17.1 years (sd = 12.3), and the mean HbA_{1c} was 11.4 (sd = 2.0). The mean age of the control patients was 52.1 years (sd = 11.1), 40% were female, the duration of diabetes was 17.9 years (sd = 11.3), and the mean HbA_{1c} was 10.4 (sd = 1.1).

Reliability

Test-retest reliability was assessed by administering the instrument a second time to patients in the control group at two weeks. The correlation between the two sets of scores were 0.87, 0.85 and 0.68 for the interference with day-to-day life dimension, and the two single items referring to insulin doses and hypoglycaemia concern, respectively. The correlations were statistically significant.

Validity

Not assessed.

Responsiveness

Not assessed.

Commentary

The DQLQ was developed concurrently for use as an outcome measure in a study comparing different methods of insulin administration (Henderson and Tindall, op. cit.). The method of item derivation is not described, and content and construct validity are not considered. Two of the three scales appear to have adequate test-retest reliability but the dimensionality and internal consistency of these scales was not assessed. The DQLQ was not tested for responsiveness to change.

Diabetes-Related Knowledge and Psychosocial Functioning Questionnaire

Purpose

The Diabetes-Related Knowledge and Psychosocial Functioning Questionnaire was developed to assess the effectiveness of a diabetes education programme aimed at improving knowledge, self-care and stress management (Gilden, Hendryx et al., 1989 op. cit.). The instrument assesses knowledge of diabetes and its care and includes a dimension relating to quality of life and diabetes. The instrument has been modified and used in a study examining the effects of self-monitoring of blood glucose on the quality of life of the elderly (Gilden, Hendryx et al., 1992).

Description

Twenty of the 67 items within the instrument assess the impact of diabetes on quality of life, in relation to diet, exercise, medication and self-monitoring of blood glucose (Gilden, Hendryx et al., 1989 op. cit.). The domain of social activities comprises three items that relate to involvement in leisure activities. One further question classified with two knowledge questions seems to refer to cognitive impairment. The remaining 34 items relate to knowledge of diabetes and family involvement and support. The development of these items is not described. An additional nine items which were adapted from the ATT-39 (Dunn, Smartt et al., 1986) form the stress dimension relating to attitudes towards diabetes.

Responses to the 20 items relating to the impact of treatment on quality of life are based on five-point Likert scales. The responses to items are summed to produce a 0-100 scale, where higher scores indicate less impact on quality of life. This component of the questionnaire is further divided into QL_a and QL_b. The former relates to more demanding and intensive lifestyle changes: general ones, and those due to diet and exercise. The latter relates to medication and self-testing. The face validity of the quality of life component was assessed by other diabetes educators. The stress items are scored on a three-point scale with a higher score indicating fewer stress-related problems. The items sum to produce a 0-18 score where higher scores indicate greater diabetes-associated stress. The three social items are scored on a five-point scale ranging from “very frequently” to “very rarely”.

The instrument was subsequently modified with 53 items covering the domains of general quality of life (14 items), interference with quality of life from the diabetic diet (10 items), exercise (3 items), medication (9 items), and urine/blood glucose monitoring (17 items) (Gilden, Casia et al., 1990). Most items were derived from the original instrument and use a six-point adjectival scale; two items are scored on a 10-point scale, and several items which are not used in scoring use a yes/no format. The instrument was administered by interview with trained medical students and diabetes nurses not known to the patient.

The instrument was further modified for use in a longitudinal evaluation (Gilden, Hendryx et al., 1992 op. cit.). The psychosocial component of the instrument comprised 20 items covering the domains of diet (4 items), exercise (2 items), medical administration (5 items), monitoring blood tests (9 items), and three general items.

Patients

The instrument was administered to patients attending a Diabetes Clinic in Chicago, USA (Gilden, Hendryx et al., 1989 op. cit.). There were two groups of patients. The first consisted of 45 males with a mean age of 70 years (sd = 4.3); 29 (64.4%) were married. The mean disease duration was 10.1 years (sd = 9.1) and 83% had NIDDM. The second consisted of 22 males with a mean age of 57 years (sd = 8.1 years). All patients were ambulatory and without any significant visual impairment.

The modified versions of the instrument were administered to two groups of patients. The first was a selected sample of 20 older diabetic patients with a mean age of 68.3 years (se = 1.2) (Gilden, Casia et al., op. cit.). All patients were performing blood- or urine-testing for at least one year. The second included 32 older male diabetic patients recruited from a diabetes clinic in Chicago, USA. The mean age of patients was 68 years (se = 1.3) and the mean disease duration was 10 years (se = 2).

Reliability

The quality of life dimension of the original instrument produced a Cronbach's alpha of 0.91, the stress dimension produced an alpha of 0.94 and the social involvement dimensions produced an alpha of 0.91 (Gilden et al., 1989 op. cit.). Cronbach's alpha for the entire Diabetes-Related Knowledge and Quality of Life Questionnaire was 0.93 (Gilden, Casia et al., op. cit.).

Validity

Face validity was confirmed by obtaining the opinions of ten diabetes educators (Gilden, Hendryx et al., 1989 op. cit.; Gilden, Casia et al., op. cit.). The Diabetes-Related Knowledge and Psychosocial Functioning Questionnaire was also reported to have content validity (Gilden, Casia et al., op. cit.).

Responsiveness

The responsiveness of both versions of the instrument has not been formally assessed but different versions of the instrument have been used in longitudinal evaluations of education programmes (Gilden, Hendryx et al., 1989 op. cit.; Gilden, Hendryx et al., 1992 op. cit.). Following an education programme, the total quality of life and QL_a scores showed significant improvements for older diabetic patients. The QL_b scores showed a significant improvement for younger diabetic patients (Gilden, Hendryx et al., 1989 op. cit.). Following a support group intervention for older diabetic patients, the total quality of life, QL_a and QL_b scores showed significant improvements (Gilden, Hendryx et al., 1992 op. cit.).

Acceptability

Not reported.

Commentary

The items relating to quality of life, stress, and involvement in social activities within the Diabetes Knowledge and Psychosocial Functioning Questionnaire were developed to evaluate the specific concerns of diabetic patients. However, patients were not involved in item derivation. The instruments were reported to have face validity but this was assessed by health care professionals. Moreover, with the exception of the stress scale, the basis of these items (including any literature review) is not described.

The performance of individual items, including endorsement frequencies and item-total correlation, is not reported and the dimensionality of the instrument was not assessed empirically. The instruments were not assessed for construct validity, either through comparisons with other instruments or with clinical and sociodemographic variables. Cronbach's alpha was reported for the entire quality of life component but not for the QLa and QLb subscales (Gilden, Hendryx et al., 1989 op. cit.). The rationale for modifications made to the instrument is not provided which hinders decisions relating to application, for example in clinical trials.

Diabetes-Specific Quality-of-Life Scale/DSQoLS

Purpose

The DSQoLS is designed to measure diabetes-specific quality of life among patients with IDDM and to assess patient preference-weighted treatment satisfaction (Bott, Mühlhauser et al., 1998). It may be used to inform individual treatment strategies and to identify motivational deficits. The authors suggest that the DSQoLS is not only helpful as an outcome measure but may also facilitate long-term care as preferences, treatment satisfaction and impact on quality of life can be assessed simultaneously.

Description

The DSQoLS has been extended to cover fear of hypoglycaemia as an additional factor, resulting in a 77-item questionnaire (Bott, personal communication).

The original instrument comprised 64 items covering individual treatment goals (10 items), satisfaction with treatment success (10 items), and diabetes-related distress (44 items). The individual items are rated using six-point adjectival scales. The preference-weighted treatment satisfaction score (PWTSS) is calculated by multiplying the patient's importance rating of each treatment goal (1 = totally unimportant, 6 = very important) by the degree of satisfaction with the achievement of the corresponding goals.

The items were selected from existing instruments developed in Germany. The DSQoLS assesses physical, emotional and social burdens together with daily functioning. Additional items were added to reflect the differences between IDDM and NIDDM, comprising patients' constraints, treatment goals, and treatment satisfaction. These items were derived from structured group discussions with IDDM patients. Two physicians, a diabetes educator, a dietician, and two educators reviewed instrument content, and helped to improve wording and item selection.

Principal component analysis was used to assess the dimensionality of the 44 diabetes-related distress items. Five items with poor component loadings or low levels of item-total correlation were removed from the instrument. The remaining 39 items formed six scales: physical complaints (8 items), worries about the future (5 items), social relations (11 items), daily hassles (4 items), and diet restrictions (5 items). Items are summed and scaled to produce scores from 0-100 where 0 represents the worst possible and 100 the best possible quality of life or treatment satisfaction.

Patients

A representative sample of 684 patients with IDDM were recruited from 630 randomly selected community-based physicians in Nordrhein, Germany. The DSQoLS was completed by 657 patients. The mean age was 36 years (sd = 11) and 42% were women. The mean disease duration was 18 years (sd = 11), 80% were receiving three or more injections per day, and mean haemoglobin levels were 8.0 (sd = 1.5).

Reliability

Cronbach's alpha for the seven dimensions of the DSQoLS ranged from 0.70 to 0.88 for the daily hassles and social relations dimensions, respectively.

Validity

The content of the DSQoLS was reviewed by four health professionals and two educators.

The DSQoLS scales were assessed for intercorrelations and were compared to the six-item Positive Well-being Scale (Bradley and Lewis, 1990) but there were no a priori hypotheses relating to specific scales or size of expected correlations. Correlations between the DSQoLS and Positive Well-being Scale ranged from 0.35 to 0.53 for the diet restrictions and social relations scales, respectively. It was hypothesised that DSQoLS dimension scores would relate to glycaemic control, treatment regimen, diabetic complications, and socioeconomic variables, and that these associations would be stronger than those for the Positive Well-being Scale.

Correlations with HbA_{1c} ranged from 0.00 to 0.24 for the daily hassles and physical complaints scales, respectively, and were significant for three of the scales. DSQoLS scores were related to the type of insulin treatment and these results were significant for the scales of leisure time flexibility, worries about the future, and diet restrictions. Patients taking a flexible approach to insulin dosage according to carbohydrate intake had better scores across the DSQoL scales, the majority being statistically significant.

DSQoLS scores were related to the presence of late complications, including retinopathy and nephropathy. The DSQoLS scales of social relations, leisure time flexibility, physical complaints and worries about the future had a statistically significant relationship with the degree of retinopathy. DSQoLS scales of leisure time flexibility, physical complaints, and worries about the future were significantly lower for patients with nephropathy. Correlations with the frequency of mild hypoglycaemia ranged from -0.03 to -0.23 for diet restrictions and treatment satisfaction dimensions, respectively, and the latter was significant.

The evidence supported the hypothesis that the DQoLS has stronger associations with diabetes-specific variables than the generic measure. The Positive Well-being Scale did not discriminate between the type of insulin treatment, the presence of a flexible adaptation strategy, nor the degree of retinopathy (except when it caused blindness).

Correlations with age ranged from 0.00 to -0.23 for the treatment satisfaction and leisure time flexibility dimensions, respectively, and were significant for two of the scales. Correlations with social status ranged from -0.02 to 0.24 for the daily hassles and physical complaints dimensions, respectively, and were significant for two of the scales. The DSQoLS scale of treatment satisfaction produced better scores for patients living with a partner (data not shown).

Responsiveness

The responsiveness of the DSQoLS has not been formally assessed. However, the instrument has been used in an evaluation of a structured inpatient treatment and teaching programme for IDDM patients (Haak et al., 1997). The subscales of treatment satisfaction, physical complaints, diet restrictions, worries about the future, social relations, and positive well-being showed significant score improvements for the 77 patients participating in the programme (Bott et al., op. cit., Haak et al., op. cit.).

Acceptability

Of the 684 patients recruited into the study, 657 (96.1%) patients attempted the DSQoLS. 27 patients (3.9%) declined to complete the questionnaire and another 18 patients (2.6%) did not answer more than 3 questions. The instrument took 10-20 minutes to complete.

Commentary

The DSQoLS was developed following a review of existing instruments, group discussions with patients and input from clinicians. Although patients were involved in item generation, their exact role is not described.

The dimensionality of the instrument was assessed using principal component analysis and the resulting scales had levels of internal consistency reliability that make the DSQoLS suitable for use in groups. The instrument has not been assessed for test-retest reliability. Validity was assessed through correlations with the generic Six-item Well-being Scale, and the instrument's association with clinical and sociodemographic variables. The developers did not, however, provide a priori hypotheses regarding the degree of statistical association they expected for particular scales, making it difficult to conclude whether the hypotheses had been met.

The DSQoLS has not been formally assessed for responsiveness but the developers refer to a study showing that several DSQoLS scales produced significant score changes following a structured inpatient treatment programme for patients with IDDM. The instrument appears to be acceptable to patients. An extended version of the instrument has been developed and is available from the developers but there are no published evaluations.

Diabetes 39/D-39

Purpose

The D-39 is an evaluative instrument designed to assess the quality of life of patients with diabetes; a slightly modified version has also been developed for use in clinical trials. The authors state that the instrument could be used to indicate areas of difficulty which could stimulate conversation between the caregiver and patient. Its purpose is to shed light on issues of importance to all diabetic patients and to determine their unmet needs (Boyer and Earp, 1997).

Description

The D-39 comprises 39 items, plus two global ratings covering the domains of physical functioning, symptoms, psychological well-being, social well-being, role activities, personal constructs, and treatment effects. Items use seven-point visual analogue scales ranging from “not affected at all” to “extremely affected”.

The D-39 was developed to be of greatest relevance to diabetes patients, and to be representative of the widest possible range of patients, irrespective of age, sex, education level, health level, and ethnicity. No attempt was made to define “quality of life” for the diabetes patient, rather the instrument specifically asks patients to indicate the impact of each item on their quality of life according to their own definition. The rationale for this approach was that it elicits responses that more accurately reflect the individual burden of diabetes and its impact on the patient.

Instrument development was in two phases. In the first phase, information derived from a literature review, existing quality of life instruments, and unstructured interviews with diabetes patients and health professionals (physicians, diabetes educators, pharmacists) was used to develop 93 items considered as addressing important aspects of patients’ lives. Each item asked the respondent to assess the extent to which their quality of life was affected during the past month by the action or activity within the item.

Following the application of factor analysis and item analysis, the instrument was reduced to 42 items in six domains. Item standard deviations were found to be approximately equal within each scale. With the exception of two items, larger correlations were found between items and scale scores than with the remaining scales. Item-total correlations were in the range 0.50-0.84.

In the second phase of the study, confirmatory factor analysis was used to confirm the presence of the six domains previously identified. Items were assessed for equivalent variances and item-total correlation. Item-total correlations were in the range 0.45-0.84. The instrument was reduced to 39 items and five domains, namely energy and mobility (15 items), diabetes control (12 items), anxiety and worry (4 items), social and peer burden (5 items), and sexual functioning (3 items). Items are summed and transformed to produce a 0-100 scale where 0 represent the best and 100 the worst possible quality of life.

Patients

The D-39 was developed and tested in three groups of patients.

The first comprised 1000 patients randomly selected from the list of a regional diabetes centre in Carey, North Carolina, USA. They were given the first questionnaire with 93 items. The mean age of patients was 52.4 years (sd = 15.8) and 55% were female; 11.8% of patients were Hispanic or Mexican, the rest were white. Thirty years of age was determined as the cut-off point for deciding whether patients had IDDM or NIDDM. On this basis, 32.5% were identified as having probable IDDM, 67.5% probable NIDDM. The mean disease duration was 14.2 years (sd = 10.9). The majority of patients (71%) were insulin users, although patients using oral medication, non-drug therapy or a combination were also represented. The proportion of patients with complications and comorbidity is not given, although comorbidity was present in some.

The D-39 was then evaluated in a group of patients who had been treated for diabetes in the two years prior to January 1989. These patients were identified from a community pharmacy list in Iowa, USA. 165 patients produced 'usable' questionnaires. The mean age of patients was 61.7 years (sd = 17.8), 55% of the sample were women and all of the patients were white. Those classed as having probable IDDM formed 20% of the total and the mean disease duration was 11.5 years (sd = 10.9). Almost half the patients (48%) were insulin users, slightly fewer were on oral medication, and a small proportion were treated by a combination of methods, or diet alone. Comorbidity was present but details are not given.

The third sample comprised 262 patients recruited from a diabetic outpatient clinic in North Carolina, USA. The mean age was 55.3 years (sd = 12.5) and 65% were women. 54.4% were black and there was a small proportion of patients from other ethnic minorities. Those classed as having probable IDDM formed 25% of the sample and the mean disease duration was 10.1 years (sd = 7.8). The majority of patients (61%) were insulin users. Comorbidity was present but details are not given.

Reliability

The six domains from the first phase (Carey sample) produced Cronbach's alpha coefficients in the range 0.81-0.92. The final five-domain instrument produced alpha coefficients in the range 0.82-0.93 and 0.81-0.93 for the patients recruited from the community (Iowa) and from the hospital outpatient department (North Carolina), respectively. Estimates of internal consistency were all above the criterion of 0.7 for subgroups of older patients and patients who had no high school education.

Validity

In the first phase of instrument development, D-39 scores were correlated with global ratings of quality of life. There were no a priori hypotheses. Not all the results were significant but they were all in the anticipated direction. Four of the six dimension scores were significantly related to self-ratings of diabetes severity. Patients with seven or more concomitant conditions had the poorest scale scores (data not shown). Patients with no concomitant conditions had the best scores on five of the six dimensions (data not shown). Patients reporting depression as a concomitant condition had poorer scores on each of the six scales (data not shown).

Compared to younger patients, those aged more than 75 years had significantly poorer scores on the scales of energy and mobility. Younger patients had poorer scores, although not always significant, on the scales of diabetes control, anxiety and worry,

social and peer burden, and diabetes medication (data not shown). Women had significantly poorer scores for the scales of energy and mobility, diabetes control, and anxiety and worry (data not shown). Patients who were not married had significantly poorer scores for the scales of energy and mobility, and anxiety and worry, and significantly better scores on the sexual functioning scale.

In the second phase, the D-39 scores were compared with those for the eight scales of the SF-36 (Ware and Sherbourne, op. cit.). The instrument was assessed in both the community and the outpatient groups. As hypothesised, the largest correlations were found between the D-39 dimension of energy and mobility, and the SF-36 scale of physical functioning ($r = 0.71$), between the D-39 dimension of anxiety and worry, and the SF-36 scale of mental health ($r = 0.64$), and between the D-39 dimension of social burden and the SF-36 scale of social functioning ($r = 0.48$). The great majority of correlations were statistically significant. All five dimensions of the D-39 had significant correlations with the self-reported global quality of life ($r = 0.21-0.44$) and self-reported diabetes severity ($r = 0.15-0.56$).

Relative to their counterparts, patients with comorbid conditions had significantly poorer scores on the D-39 energy and mobility dimensions. Compared to younger patients, those aged over 75 had poorer scores on the scales of energy and mobility. Although not always statistically significant, younger patients had poorer scores on the scales of diabetes control, anxiety and worry, social and peer burden, and diabetes medication. With the exception of sexual functioning, in which men had significantly poorer scores, there were no significant score differences between men and women. Finally, compared to patients with no employment-limiting disabilities, those with employment-limiting disabilities had poorer scores across the five dimensions.

Compared to their counterparts, patients with insulin-dependent diabetes had significantly poorer scores for the D-39 dimensions of diabetes control, and anxiety and worry. These results had the greatest levels of statistical significance in the sample of patients recruited from the community. Patients using a combination of insulin and oral therapies had poorer scores across the five dimensions.

Responsiveness

Not assessed.

Acceptability

Of the 1000 questionnaires mailed to the Carey sample, 542 were returned (54.2%). There was a 73.3% response rate from the community pharmacy sample and a 45.8% response rate from the outpatient sample. Of the questionnaires returned, 70.8% and 41.4% were deemed usable from the community and outpatient samples, respectively.

Commentary

The D-39 is designed to be of relevance to the widest range of patients, but the methods of item derivation are not described in sufficient detail for the relevance of the items to be ascertained. It is not clear what contribution patients made to the process of item derivation nor how many patients were involved. The instrument was assessed for dimensionality using factor analysis and the resulting dimensions have good evidence for internal consistency reliability. However, the instrument has not been assessed for test-retest reliability.

For purposes of validity testing, instrument scores were compared to a number of variables including SF-36 scores, and clinical and sociodemographic variables. Overall, the results of these analyses lend support to the validity of the instrument but their interpretation was limited by the lack of clearly defined hypotheses. The instrument was not assessed for responsiveness.

There was a large number of non-responders and a large number of questionnaires were deemed unusable, which has implications for the acceptability of the instrument. After looking at background data, the authors state that there is nothing to suggest that non-responders were different from responders. The scoring of the instrument is time-consuming.

The Multidimensional Diabetes Questionnaire/MDQ

Purpose

The Multidimensional Diabetes Questionnaire is designed to be a multidimensional but brief instrument assessing psychosocial adjustment to diabetes (Talbot, Nouwen et al., 1997). It is theoretically linked to a social learning perspective of diabetes designed to provide information about cognitive and social factors. The authors suggest that clinical and research applications of the MDQ are likely to include assessment, prevention and treatment. The assessment element may relate to using the instrument as a means of detecting individuals with diabetes at risk of developing depressive symptoms, or experiencing problems with adhering to their treatment regime.

Description

The MDQ was developed from a pilot 28-item scale based on the authors' experience of individuals with diabetes, comments of diabetologists, patients with diabetes attending a diabetes education programme, and their partners. Items were judged for content validity by three of the authors; following examination of the items, 7 were removed and 20 new items were created. The final version consists of three sections. The first includes the extent to which diabetes interferes with daily activities, work and leisure activities (9 items), perceptions about severity of diabetes (3 items), and social support (4 items). The second includes social incentives related to self-care activities. The final section includes self-efficacy and outcome expectancies.

The first section which measures aspects of health-related quality of life, is the focus of this review. The items within the first section use seven-point adjectival scales. Higher scores indicate higher levels of perceived interference, social support and severity. Confirmatory factor analyses were conducted to confirm the existence of a three-factor model corresponding to the three dimensions of the first section of the instrument. The comparative fit index showed an acceptable fit of the data, and the authors concluded that the postulated model was adequately supported.

Patients

The MDQ was assessed for reliability and validity in patients with NIDDM attending two diabetic education centres in Canada. Eligible patients were aged 35-65 years, had been diagnosed for at least a year, had undergone no major treatment changes in the previous six months, and had not participated in an education programme within the previous two months. The mean age of the 249 responders was 54.9 years (sd = 7.3), 45% were female and the mean disease duration was 8.3 years (sd = 7.2).

The internal consistency of the MDQ was assessed in a group of 101 patients with both types of diabetes recruited from four hospital-based diabetes education centres in Quebec City and Montreal, Canada. Eligible patients were aged 18-65 years, had been diagnosed for at least a year, and had undergone no major treatment modification in the previous year. The mean age of patients was 47.8 years (sd = 13.6); 37% were women. The mean disease duration was 11.4 years (sd = 9.3); 32% were IDDM patients. 54% were using oral hypoglycaemics and 44% were using insulin.

Reliability

Cronbach's alphas for the dimensions of section 1 of the MDQ ranged from 0.77 to 0.91 for social support and interference, respectively (ibid.). The instrument has been further assessed for internal consistency producing alpha coefficients of 0.89 for Interference, 0.72 for Severity and 0.68 for Social Support (Nouwen, Gingras et al., 1997).

Validity

The three dimensions of section 1 of the MDQ were correlated with the Beck Depression Inventory/BDI (Beck, Steer et al., 1988). There were no a priori hypotheses, but the relationship between the MDQ scales and depression was of particular interest. Correlations between section 1 of the MDQ and the BDI ranged from 0.32 to 0.56 for the social support and interference dimensions, respectively.

There was a statistically significant correlation between the MDQ interference scores and complications ($r = 0.26$), but no association was found between any of the dimension scores and HbA_{1c}. Sex was the only demographic variable included in the analysis that was found to be significantly associated with MDQ scores; women reported significantly lower levels of social support. There was no association between MDQ scores and age, education or income.

Responsiveness

Not assessed.

Acceptability

Of the 367 patient asked to participate, 329 (89.6%) agreed and 249 (75.7%) returned a questionnaire.

Commentary

The first section of the MDQ addresses health-related quality of life constructs. There is empirical evidence for the existence of the three dimensions of interference, severity of diabetes symptoms, and social support within this section. The internal consistency of three dimensions is acceptable for group comparisons but test-retest reliability has not been assessed. However, a second study found considerably lower levels of internal consistency and the social support dimension was below 0.70 (Nouwen et al., op. cit.). The MDQ has limited evidence for validity and has not been assessed for responsiveness.

The MDQ was further assessed to see whether it could be used to distinguish patient profiles according to psychological process measures: in this case the sub-groups were adaptive copers, low support-low involvement, and spousal over-involvement (ibid.). Such a classification system could be used to increase the fit between patient characteristics and specific psychosocial interventions.

Newcastle Diabetes Symptoms Questionnaire/NDSQ

Purpose

The NDSQ was developed as a component of a battery of instruments, designed to be used routinely in judging the effectiveness of care for patients with both types of diabetes being cared for in ambulatory care settings (McColl et al., op. cit.). This battery included components of generic instruments and diabetes-specific instruments such as the Diabetes Health Profile (Meadows et al., op. cit.).

Description

The development of the NDSQ has been described in two papers (Steen and McColl, 1995; McColl et al. op. cit.) and responsiveness was assessed in a separate paper (Whitty et al., op. cit.). An initial pool of 20 items was developed following a literature review of diabetes-specific instruments and using the results of a patient focus group. The items were reviewed for completeness and appropriateness by consultant physicians and general practitioners, who felt that asking patients about their symptoms over the preceding month was an acceptable period of time.

Only 3% of patients with non-insulin treated diabetes experienced a hypoglycaemic attack and many patients reported not understanding the question, so the item was removed. Several items were removed on the basis of a comparison between the responses of the diabetes sample and a matched control group without diabetes, which showed that several symptoms were not diabetes-specific. The final instrument comprises nine items which ask how often the patient has experienced symptoms over the previous month and use a five-point adjectival scale from “never” to “every day”. The final score is transformed into a 0-100 scale where 0 represents no symptoms at all and 100 represents all symptoms occurring every day.

The dimensionality of the NDSQ was assessed using exploratory factor analysis. Two and three factor solutions resulted but since the factors had little clinical significance, the factors were ignored and a single score was calculated.

Patients

The development process involved collecting data from 715 patients with diabetes, 33% of whom were treated with insulin. All patients were over 18 years of age (Steen and McColl, op. cit.).

The psychometric testing of the instrument has been described in two studies. In the first, the NDSQ was administered as part of a battery of instruments to 800 patients with diabetes aged 18 years or over, recruited from hospital outpatients and GP clinics in the UK (McColl et al., op. cit.). In the second, 48 NIDDM patients were recruited from six diabetes clinics in the North of England. Almost half the sample were over 61 years of age (range 31-83), 58% were female, the mean duration of diabetes was 7.2 years (sd = 4.2), and they had been treated with oral hypoglycaemic agents for a mean of 6.8 (sd = 5.4) years.

Reliability

The NDSQ produced a Cronbach's alpha of 0.78 (Whitty et al., op. cit.). Test-retest reliability was assessed over a three-week interval and the 95% confidence interval for the intra-class correlation coefficient was 0.90-0.96.

Validity

The NDSQ scores were correlated with the general health perception single item of the SF-36. Correlations were 0.55 and 0.63 for NIDDM and IDDM patients, respectively (McColl et al., op. cit.). Changes in NDSQ scores had a correlation of 0.32 with HbA_{1c} over a three-month period (Whitty et al., op. cit.). All correlations were statistically significant.

Responsiveness

At six weeks follow-up, the NDSQ produced a large standardised response mean (SRM) of 0.85 compared with small to moderate SRMs for the remaining instruments, which included physical function and vitality from the SF-36, depression from the Hospital Anxiety and Depression Scales, and psychological distress and barriers to activity from the Diabetes Health Profile (ibid.). Changes in scores for the instruments were much smaller at three months but the NDSQ still produced the largest SRM of 0.11; SRMs for the other instruments were in the range 0.00-0.08.

Acceptability

The first postal survey recruited 800 patients with diabetes (McColl et al., op. cit.), of whom 715 (89.4%) returned a questionnaire (Steen and McColl, op. cit.). In the second study, which assessed the responsiveness of the instrument, 48 patients were recruited and all completed the baseline assessment. Of these 43 (89.6%) and 47 (97.9%) completed the six weeks and three months follow-up questionnaires.

Commentary

The Newcastle Diabetes Symptom Scale is one of the few symptom-based instruments that has undergone a rigorous process of development and psychometric evaluation. The instrument has evidence for reliability and test-retest coefficients which indicate that the instrument may be suitable for use in individual patients. The instrument has only had limited testing for validity but demonstrates good responsiveness to change. The instrument is brief and may be useful for clinical practice.

Quality of Life: status and change/QLsc

Purpose

The QLsc is a generic instrument, partly tailored for patients with IDDM. Quality of life is defined as perceived well-being and life satisfaction, globally as well as within key domains and functions. The instrument measures satisfaction with several life domains (including social, psychological, physical), with special emphasis on well-being. The authors suggest that the QLsc could be used in analyses of coping strategies and in evaluative contexts (Hörnquist, Wikby et al., 1993).

The theory underpinning the instrument is that, first, the inner state determined by subjective experience is of more importance to quality of life than external factors. Secondly, quality of life should be considered in its entirety. Thirdly, quality of life perceptions are of a more permanent nature. Fourthly, quality of life studies are primarily descriptive.

Description

Different versions of the QLsc have been evaluated in a number of studies but the instrument generally consists of two components: a life domain rating and a well-being rating (Hörnquist et al., 1993 op. cit.; Hanestad and Albrektsen, 1991 op. cit.; Hanestad and Albrektsen, 1992; Wikby, Stenstrom et al., 1998).

The life domain rating is subdivided into two components: one measuring point status and the other measuring changes that have occurred over a period of time, or changes due to an intervention in each of the domains. The instrument is mostly reported to consist of six life domains: somatic (bodily health and specific sickness impact), psychological (well-being and intellectual function), social (social contact, family and sexual life), behavioural/activity (mobility, working capacity and basic habits), material (economic situation), and structural (religion).

The developers proposed twelve well-being domains (each containing ten items, both positively and negatively worded) including resignation, anxiety, endurance, loneliness, future-orientation, security, basic mood, inferiority, guilt, tension, sociability, and attitude of the environment.

The QLsc has been described as a flexible instrument that can be lengthened with items relating to disease-specific symptoms or side-effects, or shortened, depending on the time available for administration, is one of its positive features (Hörnquist et al., 1993 op. cit.). The maxi version of the QLsc comprises 12 well-being and 37 life domain scales. The midi comprises six well-being and 18 life domain scales. The mini comprises nine specific life domain ratings and the well-being scale “resignation” (Hörnquist, 1989; Hörnquist et al., 1993 op. cit.). As well as a version existing for carers, one has also been developed for use among children and adolescents (ibid.).

Among the life domains, items mostly use adjectival scales ranging from -3 to +3. Higher scores indicate greater satisfaction with life domains. For individual items in the well-being domains, respondents indicate whether the statements match “fully, partially, or not at all”. Scale scores range from 0 to 10. For the negatively defined scales such as inferiority, an optimal score is zero. For the scales measuring positive

aspects such as sociability, an optimal score is 10. A final composite index was also calculated from the domain indices and item scores.

The original evaluation of the QLsc used 26 life domain items: 17 generic items (comprising the psychological index, social index, behavioural index, habits index, and two broader single item scores on bodily health and global life satisfaction) and nine IDDM-specific items, of which eight form the illness index (*ibid.*). The IDDM-specific questions relate to stability, level of blood glucose, hypoglycaemia, injections, self-monitoring, compliance, and contact with health care personnel. These items use seven point scales except the life intrusion scale, which has four.

The material life sphere (personal economic status) and structural life sphere (principally religion) usually incorporated were considered less relevant for the treatment evaluation aim and were therefore excluded. In the initial and follow-up assessments, the well-being scales were anxiety, security, resignation, loneliness, endurance, and future-orientation. However, at the initial assessment a further seven scales were included. A shortened version of the life-domain rating was also given to a spouse/close friend.

There have been further evaluations of the QLsc but instrument composition was different from that of the original QLsc (Hanestad and Albrektsen, 1993; Hanestad and Albrektsen, 1991 *op. cit.*). The life domains included in these studies were physical, psychological, social and behavioural/activity, and included a total of 28 items. The diabetes-specific illness index was not included, though diabetes-specific items were integrated into the behavioural life domain and a physical domain not used in the original evaluation. The well-being ratings also differed with the exclusion of endurance, attitude of the environment, and resignation, and the inclusion of obsessive traits and indolence.

Patients

The QLsc was originally evaluated in 74 adult male and female Swedish patients with IDDM who had previously changed from syringe therapy to insulin pen treatment, attending an outpatient clinic in Sweden (Hörnquist et al., 1993 *op. cit.*). 66 patients underwent re-examination at the clinic in 1990. The mean age was 43, mean disease duration was 19 years, and complications were present. 56 patients were married or cohabiting.

The retrospective life domain change rating component asked patients to rate those changes that had occurred up to 1988 as a result of previously converting from syringe to pen delivery. At the 1990 follow-up, they were asked to rate changes in health status since the 1988 evaluation. In the corresponding parallel life domain status rating and well-being ratings, point status in 1988 and 1990 was evaluated. This group of patients was also the focus of a five-year follow-up study (Wikby et al., *op. cit.*). Of the original 74 patients, 56 were available for follow-up on all three occasions, i.e. in 1988, 1990 and 1993. Based on the 1993 data, there were 25 women, the mean age was 45.4 (sd = 14.3), the mean disease duration was 21.5 years (sd = 10.5), and 10 patients had complications.

The QLsc was further evaluated in 247 patients with IDDM at a University hospital in Norway (Hanestad and Albrektsen 1991, *op.cit.*; Hanestad and Albrektsen 1992,

op.cit., Hanestad and Albrektsen 1993 op.cit.). The mean age was 33 years (sd = 13) and 44% were women. The mean disease duration was 13 years (sd = 9); mean haemoglobin levels were 9.7% (sd = 2.4).

Reliability

Two studies have assessed the internal consistency of the QLsc (Hörnquist et al., 1993 op. cit.; Hanestad and Albrektsen, 1992 op. cit.). In the first, Cronbach's alpha for the six well-being scales were in the range 0.72-0.82. It was not reported for the seven life domain scales, but the composite quality of life score (mean score across the life domain scales) produced an alpha of 0.93 (Hörnquist et al., 1993 op. cit.). In the second, Cronbach's alphas for the life domain scales were in the range 0.49-0.86; the physical scale produced alpha values well below 0.70. Cronbach's alphas for the well-being domain scales were all above 0.70 at both assessments, the only exception being the obsessive traits scale with values of 0.46 and 0.26 (Hanestad and Albrektsen, 1992 op. cit.).

Validity

The QLsc was assessed for convergent and discriminant validity but no specific hypotheses were given (Hörnquist et al., 1993 op. cit.). The well-being scale of resignation and the composite quality of life score had the strongest level of correlation with the remaining scales, ranging from -0.23 to -0.62 and -0.25 to 0.90, respectively. Larger correlations were found within rather than between the well-being and life domain scales. Correlations between changes in scores for the life domains status scales and the parallel life domain change scales ranged from 0.34 to 0.60 for the illness index and composite quality of life scale, respectively.

Changes according to the differences in the repeated life domain status ratings between 1988 and 1990 were less consistently rated, and ranged between 0.06 and 0.55 for the ability to detect hypoglycaemia item and meaningful life item, respectively. The diabetes-specific illness domain produced relatively weak correlations with other domains and the well-being scores in particular. The authors suggest this could be due to the heterogeneity of this domain. Sex, age, cohabitation, metabolic control, and onset and duration of diabetes had little or no effect on quality of life status or change (Hörnquist et al., 1993 op. cit.).

In a further evaluation of the QLsc, a sample of 247 patients were separated on the basis of their score on one item in the physical life domain measuring the impact of IDDM on daily life (Hanestad and Albrektsen, 1992 op. cit.). The group reporting a greater impact of diabetes on daily life had significantly lower scores for the physical, psychological, social and behavioural life domains. Those reporting a greater impact of IDDM on daily life also scored significantly higher on the loneliness, tension, basic mood, anxiety, security, indolence, and future orientation well-being scales. Scores for the impact of diabetes on daily life were not related to HbA_{1c} levels; however, scores were significantly related to the complication of neuropathy though not other complications. Scores were not related to sex, age and education level.

For the same sample of 247 patients, the physical, social and psychological life domains were negatively affected by living alone (Hanestad, 1993 op. cit.). The well-being scales of indolence, safety and loneliness were significantly affected by living alone. The well-being scales of future orientation, safety, obsessive traits, loneliness,

and sociability were significantly associated with education. The well-being scales of tension, guilt, and sociability were significantly related to sex, with women reporting more guilt, tension, and sociability.

The psychological life domain scores were found to be significantly related to age, with older patients reporting less satisfaction. The well-being scale of future orientation had a significant negative association with age. The well-being scales of obsessive traits and guilt had significant positive associations with age. The satisfaction with physical life domain had a significant negative association with the presence of nephropathy, but the satisfaction with social life domain had a significant positive association with neuropathy (ibid.).

Mean HbA_{1c} levels were significantly lower at the one year follow-up assessment. Analysis of change in mean scores over the 12-month period showed a corresponding significant improvement in the activity/behavioural life domain, and the well-being scales measuring indolence, sociability, loneliness, safety, and guilt (ibid.). Well-regulated patients had significantly higher average scores on the somatic and activity/behavioural life domain ratings, although the latter was statistically significant only for well-educated patients. Well-regulated patients had better scores on the well-being scales of sociability and loneliness (Hanestad, Hörnquist et al., 1991).

In the five-year follow-up evaluation there were no significant relationships found between the QLsc, and age and complications (Wikby et al., op. cit.). Furthermore, the QLsc did not reflect significant deteriorations in metabolic control over time for the poorly regulated patients. It is suggested that this was due to differences within the sub-group of poorly regulated patients regarding their quality of life assessments, recent life events and metabolic control patterns.

Responsiveness

Not assessed.

Acceptability

The study of 247 patients with IDDM excluded 4.1% who did not wish to participate and 4.4% who did not return the questionnaire (Hanestad et al., 1991 op. cit.). The full version of the QLsc which includes life domain status and change, and 14 well-being scales can be completed within one hour (Hörnquist, 1989 op. cit.). The reduction of the well-being scales at follow-up was due to the extra burden of including an additional life event questionnaire in the battery (Hörnquist et al., 1993 op. cit.).

Commentary

The QLsc is a generic instrument adaptable for use with specific patient groups, through the inclusion of additional items within the life domains component of the instrument (Hörnquist, 1989 op. cit.). The questionnaire is lengthy and the full version takes up to one hour to complete. The dimensionality of the QLsc has not been tested empirically, and it is possible the psychometric properties of the instruments could be improved if further consideration is given to the composition of the instrument.

Two studies have compared scores for stability (Hörnquist et al., 1993 op. cit.; Hanestad and Albrektsen, 1992 op. cit.) but, given the large time interval between the

assessments and the lack of a health transition item to assess whether patients had remained stable, these data should not be considered estimates of test-retest reliability. Many correlations were produced in support of validity testing; however, without specific hypotheses it is difficult to assess the evidence for the validity of the instrument. The QLsc was not assessed for responsiveness.

Questionnaire on Stress in Diabetic Patients - Revised/QSD-R

Purpose

The QSD-R is designed to assess psychosocial stress associated with problems in daily living with both types of diabetes. The authors suggest it could be used to identify patients with psychosocial problems related to diabetes (Duran, Herschbach et al., 1995). The instrument was designed to allow patients to report any diabetes-related problems from their perspective, and to be easy to administer and interpret. Herschbach comments that one of the applications of the QSD-R is in the evaluation of treatment, although evidence of responsiveness has yet to be established (Herschbach, Duran et al., 1997).

Description

The original QSD comprised 90 items designed to screen for diabetes-related stress in a wide range of everyday situations. The items were selected following a literature review, and interviews of 5 diabetologists and 76 patients with diabetes (Duran et al., op. cit.). Patients are asked to indicate whether a given item applies to them and, if so, rate the extent to which daily stresses in that area causes them problems. Items use five-point adjectival scales ranging from “only a slight problem” to “a very big problem”.

Principal component analysis (PCA) of QSD produced 10 components. Three components cover social problems, namely problems with relationship/family, strained doctor-patient relationship, and problems with work; three cover emotional problems, namely acceptance of diabetes, fear of complications, and feeling patronised. Two components cover barriers to adherence, namely dietary restrictions and difficulties with treatment regimen; two cover physical limitations due to diabetes, namely problems with hypoglycaemia and reduction of performance.

The revised QSD-R comprises 45 items describing situations that are a source of stress for many people with diabetes, including leisure time, depression/fear of the future, hypoglycaemia, treatment regimen/diet, physical complaints, work, partner, and the doctor-patient relationship. The instrument was revised on the basis of psychometric evidence (Duran et al., op. cit.) and clinical experience with patients. The 45 items were subjected to PCA producing eight components. Clinical considerations led to the re-grouping of 14 items.

Patients

The QSD was developed in a sample of 617 diabetes patients (392 with IDDM and 225 with NIDDM), half of whom were being treated as inpatients and half as outpatients at 10 different treatment facilities in Germany. The mean age was 42.6 years (sd = 17.4), 55% were women, and 368 (61%) were married. The mean disease duration was 11.7 years (sd = 10.6); 51% of patients had at least one long-term complication.

The QSD-R was administered to 1930 patients recruited from seven inpatient and five outpatient departments in different parts of Germany (Herschbach et al., op. cit.). Of these patients, 915 (47.3%) had IDDM and of the 1015 (52.7%) with NIDDM, 253 (25%) were being treated with insulin. Patients were aged 18-85 years and there were 394 (46.2%) women and 344 (37.1%) men in the IDDM and NIDDM groups,

respectively. The mean disease duration was 15.5 years for the IDDM group and 9.1 years for the NIDDM group. 45% of patients had long-term complications. Construct validity and test-retest reliability was performed on a subgroup of 82 patients which did not differ markedly from the main group.

Reliability

Cronbach's alpha for the ten scales identified through PCA for the QSD was in the range from 0.63 to 0.88 (Duran et al., op. cit.). The scales of dietary restriction and feeling patronised produced alpha values below 0.7. Cronbach's alpha for the eight dimensions and total QSD-R ranged from 0.69 to 0.93 for the partner and total scores, respectively (Herschbach et al., op. cit.). Test-retest reliability was assessed for the QSD-R at five weeks among a subgroup of patients who had not undergone any treatment changes. Test-retest correlation was 0.63 for the QSD-R total scores, ranging from 0.45 to 0.73 for the individual scales.

Validity

The Complaint List and Questions on Satisfaction with Health were administered alongside the QSD. The former includes 24 items assessing somatic and general complaints (von Zerssen, 1976), while the latter measures satisfaction with one's state of health and was developed by the authors of the QSD. The QSD produced significant correlations of 0.48 and -0.42 with the Complaint List and the Questions on Satisfaction with Health, respectively. The ratings on the QSD were significantly higher for those with long-term or acute complications.

The QSD-R was compared with the Beck Depression Inventory/BDI (Beck et al., op. cit.) and the State-Trait Anxiety Inventory/STAI (Spielberger, Gorsuch et al., 1970). The depression/fear of future scale of the STAI was examined for particular correlation with the QSD-R. Correlations were in the ranges 0.39-0.67 and 0.33-0.71 for the BDI and STAI, respectively. As expected, the QSD-R dimension of depression/fear of future had the largest level of correlation with the STAI. All but one of the QSD-R dimension scores were significantly higher for patients with long-term complications. The dimension scores and total scores were all significantly higher for patients with poor metabolic control (HbA_{1c}).

Responsiveness

Not assessed.

Acceptability

Of the 838 patients given a questionnaire, 617 (73.6%) returned completed questionnaires. Patients took 5-15 minutes to complete the QSD-R (Herschbach et al., op. cit.).

Commentary

The QSD and QSD-R have been evaluated in patients with both types of diabetes receiving a range of treatment modes. Patients were involved in the derivation of items which lends the instrument content validity. The dimensionality of both versions of the instrument has been assessed but the performance of individual items in terms of endorsement frequencies and item-total correlation was not reported.

The revised instrument has greater levels of internal consistency than the QSD but two scales fell below 0.70. The low estimates of test-retest reliability are cause for concern but real changes in health may have occurred over the five-week interval between test and retest. Evidence for construct validity is good, although the generation of more hypotheses would allow a more thorough evaluation.

The shorter-form QSD-R is likely to be more acceptable to patients and is more practical for use in clinical settings. However, little information is provided on the rationale for the modification of the QSD. Neither version of the instrument has been assessed for responsiveness.

Social Psychological Health States/SPHS

Purpose

The SPHS attempts to measure the psychosocial effects, and impact on job-related performances and social role performances of maturity onset diabetes for ambulatory patients taking part in an experimental nursing intervention (Given, 1984). The instrument was designed to detect limitations imposed by the disease and measure improvements in health as a result of an intervention aimed at increasing patients' involvement in the management of their disease. It is not clear whether the instrument is self- or interviewer-administered.

Description

The content of the SPHS was developed following a review of measures of social-perceived and psychological health states, and from in-depth interviews with a sample of 25 diabetic patients and 23 hypertensive patients. These patients were asked to describe how their disease had affected four different aspects of their life: social interactions with family and friends, recreational activities, cognitive and affective reactions, and job-related activities and interactions. 45 items were developed covering these four areas. Responses were scaled using five-point Likert scales.

Items were divided into three hypothesised dimensions of usual daily activities, affective and cognitive reactions, and impact on job performance. The items were then assessed for their scale contribution using factor analysis and item-total correlation. These analyses produced two dimensions: psychosocial effects (21 items) and impact on job performance (7 items). The scales were internally consistent and the results were replicated in a further sample of patients.

With the exception of one item which fell below 0.40, the item-total correlations for the psychosocial effects scale were in the range 0.43-0.75. With the exception of one item which fell below 0.40, item-total correlations for the impact on job performance dimensions were in the range 0.46-0.84.

Patients

The instrument was administered to two samples of 156 and 292 diabetic patient recruited from 11 ambulatory care centres. Patients who were pregnant/lactating, who had evidence of stroke, cancer, end-stage renal failure, and psychoses were excluded. Patients in both samples were aged 18-65 years, the majority of patients was female (61% and 73%) and the majority was white (93% and 77%).

Reliability

Cronbach's alpha for the psychosocial effects and impact on job performance dimensions were 0.92 and 0.86-0.87 respectively.

Validity

The instrument was compared to a symptom severity scale, the Health Perception Questionnaire/HPQ (Ware, 1976), the anxiety and depression scales from the Hopkins Symptom Checklist (Derogatis et al., 1973), and the Functional Limitations and Physical Abilities scale developed by the Rand Corporation (Ware, Johnston et al., 1979). The authors hypothesised that the psychosocial health-state scale would be negatively correlated with the current and prior health subscales of the HPQ, and

positively related to the anxiety and depression scales and the symptom-severity scale. Furthermore, the authors expected limitations in function and physical abilities to correlate more highly with the job performance than with the psychosocial health-state scale.

Both scales were negatively and similarly associated with the current and prior health HPQ subscales. The psychosocial scale was positively associated with depression ($r = 0.56$) and symptom severity ($r = 0.58$) but not with anxiety ($r = -0.00$). Contrary to expectations, the functional and physical limitations scales correlated more strongly with psychosocial health ($r = 0.73$ and 0.60 respectively) than with impact on job performance ($r = 0.65$ and 0.26 respectively). The authors comment that as one third of the sample suffered no limitation in functional or physical abilities, the two dimensions may not be sensitive to variations in health status.

Responsiveness

Not assessed.

Acceptability

Information relating to response rates and completion rates is not provided.

Commentary

The SPHS was developed to measure the psychosocial effects, job-related performances and social role performances of diabetes and hypertension. The instrument was not developed solely for diabetes patients, which may limit its measurement properties and its relevance in terms of content validity with diabetes patients.

There is empirical evidence for the three dimensions and, on the whole, individual items have a good level of correlation with dimension scores. The internal reliability of the SPHS is acceptable for use in groups but the instrument has not been assessed for test-retest reliability. Evidence for the validity of the instrument was demonstrated through correlations with measures of functional limitations and physical abilities scales. The instrument has not been assessed for responsiveness.

Type 2 Diabetes Symptom Checklist/DSC-Type 2

Purpose

The DSC-Type-2 was developed for use in clinical and epidemiological research, and is designed to measure differences in symptom severity between patients and detect changes over time within patients (Grootenhuis, Snoek et al., 1994).

Description

The DSC-Type 2 is a 34-item instrument covering six symptom dimensions. The patient is asked about the frequency of a symptom during the past four weeks on a four-point scale (not at all, one or more times a month, one or more times a week, daily) and if the symptom occurs they are asked to indicate the level of discomfort using a four-point scale (not at all, a little, quite a bit, very much). The dimension scores are calculated by summing relevant frequency scores, and the scores are weighted by multiplying the frequency of an item by the corresponding discomfort score. Final scores are converted to a 0-10 scale, where 0 signifies the absence of symptoms and 10 the worst possible symptoms.

Originally 78 physical and psychological symptoms and complications associated with NIDDM were derived from a literature review and discussions with diabetologists. These items were sent to 20 experienced clinicians (diabetologists, general practitioners and diabetes educators) who were asked to consider whether the symptoms occur more frequently in diabetic patients than in the general population. The clinicians were asked to contribute additional items of importance; 42 were added. The clinicians were then asked to assess the resulting 120 items. Items were removed if less than 50% of clinicians stated they were associated with diabetes, or more than 25% stated they were not associated. The resulting instrument comprised 51 items in six dimensions, namely hyperglycaemia (7 items), hypoglycaemia (6 items), psychological symptoms (10 items), cardiovascular symptoms (8 items), neuropathy (10 items), ophthalmological symptoms (5 items), and rest (5 items).

Three items relating to signs rather than symptoms were removed from the instrument and three were removed because they were considered too abstract. The definition of these concepts and related criteria are not provided. 11 items were removed because they had levels of item-total correlation below 0.35. Principal component analysis was undertaken for the remaining 34 items which supported the existence of six dimensions: hyperglycaemia (4 items), hypoglycaemia (3 items), cardiovascular symptoms (4 items), neuropathy (4 polyneuropathic pain items and 6 polyneuropathic sensory items), psychological symptoms (4 psychological fatigue items and 4 psychological distress items), and ophthalmological symptoms (5 items).

Because of a high proportion of zero scores, the distribution of the final scores for all ten scales and total scores was positively skewed (Grootenhuis et al., op. cit.). This result was found in a further study (Gulliford and Mahabir, 1999) although aggregate DSC-Type 2 scores have been found to be less skewed than other instruments (Van der Does, de Neeling et al., 1996).

Patients

The original evaluation of the instrument comprised 116 patients visiting a diabetes outpatient clinic in Amsterdam, Holland, 29 patients referred for suspected

polyneuropathy and 46 patients from a general practice. The mean age across all the samples was 65 years (sd = 11) and 88 (47.6%) were women. The mean disease duration was 11.8 years (sd = 7.1), most patients (61%) were using insulin and the prevalence of comorbidity was 61%.

The instrument was further evaluated in patients with diabetes attending government health centres in Trinidad and Tobago (Gulliford and Mahabir, *op. cit.*). There were 2117 patients in the final sample although only 1880 providing completed questionnaires were included. Of these, 33% of patients were aged 65 or older, 70% were women, 32% were Afro-Trinidadian, and 50% were Indo-Trinidadian. Three proxies for socio-economic status were used: no school education (9%), no piped drinking water supply to the home (30%), and employment (192 subjects had full-time jobs). It was assumed that the majority of patients would have NIDDM. Patients had suffered from diabetes for 10 years or more, and 71% were taking oral hypoglycaemic agents.

Reliability

Cronbach's alpha coefficients for the frequency and weighted scores ranged from 0.76 to 0.93 for the hypoglycaemia and psychological fatigue scales, respectively. Cronbach's alpha coefficients for the total scores ranged from 0.94 to 0.95 for the frequency and weighted scores, respectively. Test-retest reliability was assessed among 116 patients from the outpatient clinic and general practice after an interval of 3 weeks. The test-retest coefficients for the frequency and weighted scores ranged from 0.79 to 0.94 for the hyperglycaemia and neuropathy scales, respectively. Test-retest coefficients for the total scores ranged from 0.94 to 0.92 for the frequency and weighted scores, respectively (Grootenhuis *et al*, *op. cit.*).

For the Trinidad and Tobago study, alpha coefficients ranged from 0.63 to 0.76 for the hypoglycaemic and cardiovascular scales, respectively. The total score had a Cronbach's alpha of 0.94. Test-retest reliability was measured in a convenience sample of 53 of the original sample, by the mean difference between score results. This difference ranged from 0.12 (-0.39 to 0.63) to 0.56 (-0.02 to 1.13) for hyperglycaemia and hypoglycaemia, respectively.

Validity

It was hypothesised that relative to their counterparts, patients with diabetes-related comorbidity would score higher on scales related to diabetic complications, and that patients with insulin-treated diabetes would score higher still. Higher scores were found for patients with comorbidity compared to those without comorbidity for all ten scale and total scores; the only insignificant finding was for the polyneuropathic scale. When insulin-treated patients with suspected polyneuropathy were compared to other insulin-treated patients, the former had significantly higher scores on the polyneuropathic sensory and polyneuropathic scales. Compared to patients treated with oral hypoglycaemic agents and after stratifying by comorbidity, patients with insulin-treated diabetes had higher scores for the majority of scales, except for hyperglycaemia.

The DSC-Type 2 scores have been compared to HbA_{1c} levels, with correlations ranging from 0.02 to 0.19 for the hypoglycaemia and hyperglycaemia scales, respectively. The majority of the correlations were statistically significant. The total

scores produced a statistically significant correlation with HbA_{1c} of 0.23 (Van der Does et al., op. cit.).

The instrument has also been compared with the SF-36 for purposes of assessing construct validity (Gulliford and Mahabir, op. cit.). For each of the scales, a worsening score on the DSC-Type 2 was associated with lower SF-36 scores. After adjusting for comorbidity, the association between the DSC-Type 2 and the mental health component summary scores of the SF-36 remained the same, although the association with the physical component summary scores was reduced. After adjusting for social and demographic factors, the SF-36 physical component summary scores were found to be associated with the psychological and neuropathic symptoms of the DSC, whilst the SF-36 mental component summary scores were associated with psychological and hypoglycaemic symptoms.

The relationship between the DSC-Type 2 and various disease variables was also examined. The DSC-Type 2 was not found to be associated with age. Higher scores were associated with female sex, Indo-Trinidadian ethnic origin, lower levels of education, absence of a piped water supply to the house, and not having a full-time job. The DSC-Type 2 scores were not significantly related to HbA_{1c} levels.

Responsiveness

The minimal detectable mean change in score over time was calculated by dividing patients into two groups: those who were stable over time but had a mean change in score less than the median, and those who were stable but had a mean change in score greater than the median (Grootenhuis et al., op. cit.). Responsiveness was calculated on the basis of standard deviations of within-subject differences between the first and second administrations of the instrument. The minimal detectable mean change in score in a group of 100 patients was in the range 0.07-0.25 for the group with a score change less than the median, and 0.14-0.58 for the group with a score change greater than the median.

Acceptability

Sixty six percent of the diabetic outpatient clinic sample returned the questionnaire. Among those attending a neurology clinic, 100% returned the questionnaire. 77% of the general practice patients returned the questionnaire. 3% of the returned questionnaires were unsuitable for analysis. The follow-up questionnaire was completed by 74% of patients. Administration of the questionnaire took about ten minutes and items were formulated so as to be understandable to people with a low reading level. There were up to 7.6% missing values per item, although 77% completed the questionnaire without missing any items. Six patients were excluded due to incorrect completion of the questionnaire (ibid.).

Gulliford et al. (1999, op. cit.) reported that 96% provided complete scores for the DSC, whilst there was a 91% response rate among the sample chosen for the test-retest evaluation.

Commentary

The DSC-Type 2 is a symptom-based instrument with good evidence for the reliability and validity of its scales. The instrument was based on a literature review and expert judgements. The internal consistency reliability estimates all exceed the level necessary for group comparisons. The instrument scores were significantly

related to those for the SF-36. The responsiveness of the instrument was not assessed using an effect size statistic, but minimally clinical important differences have been calculated (Grootenhuis et al., op. cit.). The authors conclude that although strong support for two separate polyneuropathy dimensions was not found, this is clinically relevant.

Well-Being Enquiry for Diabetics/WED

Purpose

The Well-Being Enquiry for Diabetics aims to measure disease-specific quality of life among patients with both types of diabetes and is intended for application in different clinical settings. The dimension serenity is proposed as a screening tool for mental disorders (Mannucci, Ricca et al., 1996).

Description

The instrument was developed “using the contribution of several diabetologists, psychiatrists, nurses and diabetic patients” (Manucci et al., op. cit.). Some of the items were derived from other instruments, including the Diabetes Quality of Life Questionnaire (Jacobson et al., 1988 op cit.). The items contribute to four subscales: symptoms (20 items), discomfort (10 items), serenity (10 items), and impact (20 items). The symptoms subscale covers diabetes-related somatic symptoms and physical functioning. The discomfort subscale covers diabetes-related worries and emotional status. The serenity subscale covers mental health. The impact subscale covers family relationships, role functioning, and social network. Some items relate directly to the impact of diabetes while others, particularly in the serenity dimension, are not specifically restricted to the impact of diabetes. The items use five-point adjectival scales.

Factor analysis produced four alternative subscales of symptoms and impact, depression and shame, diet and eating attitudes, anxiety and worries. These new subscales were examined for reliability and validity. However, the authors conclude that the original WED serenity dimension retains its validity as an overall measure of mental health, whilst the subdivision of impact suggested by factor analysis does not produce additional relevant clinical information.

Subscale scores are calculated by summing responses to individual items. The total WED score is the sum of the subscale scores. Higher WED scores indicate a better quality of life.

Patients

The Italian translation of the instrument was evaluated in a consecutive sample of 267 patients (70 with IDDM and 197 with NIDDM) aged 15-67 attending an outpatient clinic at the University of Florence, Italy. The sample consisted of patients with a range of treatment modalities. In addition the existence of diabetic complications and comorbidity was documented. The IDDM patients had a mean age of 33.6 years (sd = 11.7), 60% were female, mean disease duration was 12.7 years (sd = 10.6) and mean HbA_{1c} levels were 7.2 (sd = 1.5). The NIDDM patients had a mean age of 58.9 years (sd = 9.8), 52.3% were female, mean disease duration was 9.1 years (sd = 7.8) and mean HbA_{1c} levels were 7.1 (sd = 1.6).

Reliability

Cronbach's alpha were in the range 0.81-0.84 for the four subscales and total scores. Patients were asked to complete the questionnaire twice within an interval of 15-60 days. 29 patients declined. There was no external criterion to assess whether patients had remained stable. For the WED total score, a correlation of 0.89 was observed.

Test-retest correlations ranged from 0.68 to 0.89 for the symptoms and total scores, respectively.

Validity

The WED was compared to the Diabetes Quality of Life Questionnaire/DQoL (Jacobson et al., 1988 op. cit.), Hamilton's Depression Rating Scale/Ham-D (Hamilton, 1960), State and Trait Anxiety Inventory/STAI (Spielberg et al., op. cit.) and Bulimic Investigation Test Edinburgh/BITE (Henderson and Freeman, 1987).

WED scores were also related to a number of clinical variables. There were no a priori hypotheses. The correlations between WED and DQoL scores ranged from -0.05 to -0.68 . The correlations between WED and STAI scores ranged from -0.13 to -0.63 . The correlations between WED and Ham-D scores ranged from -0.29 to -0.49 . The correlations between the WED and BITE scores ranged from -0.26 to -0.35 . The vast majority of these correlations were statistically significant. WED scores for the subscales of symptoms and discomfort were significantly lower for patients with complications compared to those without complications. There was a significant correlation ($r = -0.35$) between WED scores and HbA_{1c} for the IDDM patients but not for the NIDDM patients.

There were no differences in WED scores between males and female IDDM patients, but females had significantly worse scores on the dimensions of symptoms and serenity in the NIDDM group. WED scores were significantly correlated with age in patients with IDDM ($r = 0.25$) and NIDDM ($r = 0.32$).

Acceptability

Of the 294 patients approached to take part in the study, 27 (9.2%) refused to participate.

Commentary

Health professionals and patients contributed to the development of the WED but the process of development is not described. Some of the items were derived from other instruments including the DQoL. The WED subscales were not supported by factor analysis but have satisfactory levels of internal consistency. The test-retest coefficients are lower but this may be due to weaknesses in the design of the test-retest study, including the absence of criteria to assess whether patients had remained stable between administrations of the instrument. For purposes of validation, the WED was compared to several other instruments and clinical variables, but the lack of a priori hypotheses limits the interpretation of results. The WED has not been assessed for responsiveness.

The Well-Being Questionnaire/WBQ

Purpose

The WBQ was developed to provide a measure of mood, anxiety, and aspects of positive well-being for use in a World Health Organisation study evaluating new treatments for diabetes (Bradley, 1994 op. cit.). The instrument is designed to assess the efficacy of new interventions. The authors state that the instrument may be useful in the routine audit of established treatment. The inclusion of a positive well-being dimension is designed to increase the sensitivity of the instrument. The instrument was initially developed with IDDM patients but has also been developed with NIDDM tablet-treated patients).

The positive well-being dimension is designed to assess psychological aspects of well-being, both negative and positive, independently of glycaemic control, and has been validated independently (Bardsley, Astell et al., op. cit.). The WBQ has been recommended for identifying individuals who require referral to psychological services, as well an audit tool via the DIABCARE routine audit dataset, to enable the monitoring of care throughout Europe (Vaughan, Bradshaw et al., 1994; Williams, 1992; Wilson, Home et al., 1993).

The WBQ has been translated into over 20 languages. There are three short-form versions of the instrument: the twelve item (Depressed) Well-being Questionnaire, the W-BQ12, and the WHO (Ten) Well-being Index. The twelve item (Depressed) Well-being Questionnaire was developed for routine use in a study of audit of diabetes care (Wilson and Home, op. cit.). The WHO (Ten) Well-being Index was developed to identify items from the WHO questionnaire belonging to an overall index of positive and negative well-being (Bech, Gudex et al., 1996). The instrument was not intended to replace the WBQ but to act as an alternative when a unidimensional scale of well-being is required. The W-BQ12 was developed as a short-form version of the WBQ for use by clinicians and researchers (Pouwer, van der Ploeg et al., 1999).

Description

The original instrument comprised 28 items. The depression and anxiety components were taken from an existing instrument (Warr, Banks et al., 1985). The positive well-being items were devised in consultation with diabetologists and psychologists. The final instrument includes four subscales: depression (6 items), anxiety (6 items), positive well-being (6 items), and energy (4 items).

The WHO (Ten) Well-Being Index consists of one item from the three scales of depression, anxiety, and energy (Bech et al., op. cit.). An item on sleep was retained, as were all six items from the positive well-being dimension. The W-BQ12 questionnaire includes three dimensions, each with four items: negative well-being, positive well-being, energy, and a summary score labeled general well-being (Pouwer, van der Ploeg et al., op. cit.). The former consists of two anxiety and two depression items, all negatively worded. Four items were selected from the original positive well-being dimensions (all positively worded). The energy dimension remained unchanged.

The twelve item (Depressed) Well-being Questionnaire (Wilson and Home, op. cit.) includes the first six items from the depression dimensions of the WBQ and six diabetes-specific items from the Diabetes Health Profile (Meadows et al., op. cit.).

The WBQ was assessed for dimensionality using principal component analysis (PCA) in two studies with tablet-treated NIDDM and IDDM patients respectively (Bradley and Lewis, *op. cit.*; Bradley, 1994 *op. cit.*). The first study used PCA to select six positive well-being items from the original sixteen. The six positive well-being items loaded onto the first component of a forced three component PCA. The anxiety items loaded primarily onto the second component and the depression items largely accounted for component three. Item-total correlations for the positive well-being dimension were in the range 0.62-0.76.

The second study sought to confirm the PCA results in IDDM patients and used forced and unforced PCA on the positive well-being items to confirm the existence of a six-item dimension and to explore the construction of an additional dimension of energy. The presence of the dimensions was confirmed in three European centres. PCA of the six anxiety items led to single component solutions for the English and French centres, though the German data produced two components. PCA of the six depression items resulted in a single component for each centre; however, up to two items performed poorly for each centre.

The WBQ was subjected to a PCA that produced three components (Wredling, Stalhammar et al., 1995). The six positive well-being items loaded onto the first component, four of the anxiety items loaded onto the second and the two remaining anxiety items loaded onto the third. Four of the depression items loaded onto the third component; the other two loaded onto the second component. There was no distinctive component for the energy dimension.

Bech et al. (1996, *op. cit.*) reported a PCA of the WBQ and found one general component which included the 28 items. Promax rotation produced four components. The results were similar to those reported by Bradley except for a fourth component with the highest component loadings on love and sex life. In the development of the 10-item well-being scale, one item was chosen to represent three of the dimensions. An item on sleep was retained along with the six items of positive well-being.

Pouwer, van der Ploeg et al. (*op. cit.*) used factor analysis to assess the dimensionality of the W-BQ12 and found evidence for a three-factor solution supporting the dimensions of positive well-being, negative well-being and energy. The item-total correlations for the W-BQ12 were in the range 0.38-0.75.

Each item is scored on a 0 to 3 Likert scale where 0 represents “not at all” and 3 “all the time”. Ratings for the items are summed, after reversing where necessary. A higher score indicates more of the specific mood state.

Patients

Two studies are reported by the original developers of the WBQ (Bradley et al., 1990 *op. cit.*; Bradley, 1994 *op. cit.*). The first recruited 219 patients with NIDDM attending an outpatient clinic in Sheffield, UK. Patients were aged 45-65 and 59.4% were male. They were given the original 28-item WBQ and 184 completed questionnaires were available for analysis. In the second study, the patients included IDDM patients recruited for a multi-centre European study. Psychometric analyses were performed on English (n = 41), French (n = 69) and German patients (n = 40) of

both sexes. Mean ages were: English 35 years (sd = 10.0), French 37 years (sd = 10.0), German 30 (sd = 9.0). Mean disease duration was: English 12 years (sd = 8.0), French 12 years (sd = 8.0), German 11.5 (sd = 7.0).

The WBQ subscale of positive well-being was compared with the Nottingham Health Profile and an anglicized version of the Sickness Impact Profile in 284 patients with both types of diabetes, recruited from a hospital database in the UK (Bardsley et al., op. cit.). The sample was stratified by age and type of diabetes to be representative of the patient population. Ages were fairly evenly distributed in the range from 40 to just over 65 years.

The WBQ underwent further evaluation in a sample of 423 patients (83.6% of those invited), recruited from hospital outpatient clinics, primary health care centres and private practitioners in two Swedish counties (Wredling et al., op. cit.). The sample consisted of middle-aged to elderly patients; 152 were insulin-treated, 270 were treated with diet only, or diet and oral hypoglycaemic agents.

The WHO (Ten) Well-being Index was evaluated in patients attending ten European centres in eight countries (Bech et al., op. cit.). Participants were NIDDM patients aged 16-60 years who had been on insulin for at least two years and who needed more than 0.3 units/kg, but not more than four injections of insulin per day. The analysis was limited to the 358 patients who completed the WBQ and four general assessment questions at the time of randomisation.

Paper and pen administration of the WBQ was compared with computerised administration in a randomised cross-over design (Pouwer, Snoek et al., 1998). In the Netherlands, 105 patients over the age of 16 with both types of diabetes were asked to participate. The mean age was 48.1 years (sd = 17.9), 39% were female, mean level of HbA_{1c} was 8.1% (sd = 1.3). The patient groups were found to be comparable in terms of age, type of diabetes and HbA_{1c} levels.

The Well-Being Questionnaire was evaluated in a postal survey of 1000 diabetic patients aged 60 or over (mean age 71 years), randomly selected from the Salford Collaborative Diabetes Care Programme register in the UK (Petterson, Lee et al., 1998). Patients received a range of treatment modalities but the proportions are not given. The mean age was 71 years (sd = 7), 51% were male, and the mean disease duration was 8 years. Of the responders, 17% were on diet alone, 55% were on tablets, and 28% were on insulin.

The W-BQ12, HADS, and the trait scale of the STAI and other instruments were mailed to 3000 members of a diabetes self-help group in the Netherlands (Pouwer, van der Ploeg et al., op. cit.). A second set of questionnaires was sent out containing the Centre for Epidemiological Studies Depression/CES-D scale and the SF-36. The sample consisted of patients with both types of diabetes, a mean age of 51 years, and a mean disease duration of 16 years. Almost half the sample reported suffering from diabetic complications.

Reliability

The original evaluation of the WBQ produced Cronbach's alpha coefficients in the ranges 0.70-0.88 and 0.64-0.80 for the English sample with NIDDM and IDDM

patients, respectively. For the French and German samples, the ranges were 0.46-0.89 and 0.66-0.75, respectively. The depression subscale had a very low Cronbach's alpha in the French sample (Bradley, 1994 op. cit.). Subsequent studies have produced alpha coefficients in the range 0.64-0.89; the depression subscale fell below 0.7 (Wredling et al., op. cit.). Cronbach's alpha for the WBQ total scores was reported as 0.90 for the 22-item WBQ and 0.85 for the 10-item WBQ (Bech et al., op. cit.). Cronbach's alpha for the W-BQ12 dimensions ranged from 0.73 to 0.91 for negative well-being and general well-being, respectively (Pouwer, van der Ploeg et al., op. cit.).

Test-retest reliability was assessed in one study, where patients completed both paper-and-pencil and computerised versions of the questionnaire within seven days (Pouwer, Snoek et al., op. cit.). Test-retest correlations ranged from 0.72 to 0.85 for the depression and general well-being dimension scores, respectively.

Validity

In the original evaluation of the WBQ, it was hypothesised that, compared to their counterparts, patients with complications of diabetes or other physical disorders would be expected to report lower levels of well-being (Bradley and Lewis, op. cit.; Bradley, 1994 op. cit.). There were significant differences between the groups on all three WBQ subscales and the total well-being scores (Bradley and Lewis, op. cit.). Compared to men, women had significantly higher depression and anxiety scores and significantly lower general well-being scores. The results follow previous findings in general population studies (Bradley, 1994 op. cit.).

No relationship was found between treatment modalities for the WBQ positive well-being dimension among NIDDM patients. The relationship between the WBQ and several other variables was also assessed but there were no a priori hypotheses. As desired, there was no correlation between WBQ dimensions and HbA_{1c} which implies that the instrument measures well-being without confounding it with the effects of blood glucose (Bradley and Lewis, op. cit.).

The WBQ positive well-being dimension has been compared with the Nottingham Health Profile/NHP and the Sickness Impact Profile/SIP (Bardsley et al., op. cit.). The strongest correlations were with the NHP dimensions of energy ($r = 0.45$) and emotional reactions ($r = 0.45$), and SIP dimensions of social interactions ($r = 0.33$) and home management ($r = 0.32$). All correlations were statistically significant.

The positive well-being subscale was assessed by comparison with responses to the questions: "How do you cope with your life?" and "How does diabetes affect your life?". Patients were also asked to provide a summary rating of their health on a five-point ordinal scale ("very poor" to "very good") and this was compared to the scale score. Patient responses to these questions tended to correlate with the positive well-being dimension but were not statistically significant. There was no relationship between the positive well-being scores and the presence of complications, and there was no association with HbA_{1c} (Bardsley et al., op. cit.).

WBQ scores have also been correlated with responses to four summary questions relating to perceived health situation, disease impact and extent of thinking about diabetes (Wredling et al., op. cit.). The depression and anxiety dimensions were negatively correlated with perceived health situation (both $r = 0.43$). The energy and

general well-being dimensions were correlated with general health situation (both $r = -0.48$). The depression dimension correlated with the impact of disease on daily living ($r = 0.42$) and thinking about diabetes ($r = 0.59$). All correlations were significant. There was no significant correlation between the WBQ scores and HbA_{1c}. Women had significantly higher scores for depression than men. Compared to men, insulin-treated women had significantly higher anxiety levels, and lower energy and general well-being scores.

In a study of a large community sample of older people with diabetes, no association was found between between the WBQ and HbA_{1c} (Pettersen et al., op. cit.). Women were also found to have significantly poorer scores than men on all but the energy dimension.

Evidence of criterion validity of the 22- and 10-item versions of the WBQ was demonstrated by the large significant correlations with the parent 28-item WBQ, which were in the range 0.93-0.98 (Bech et al., op. cit.). The 22- and 10-item WBQ scores were also compared with four general assessment questions. For both instruments, highly significant relationships were found with higher levels of well-being associated with less disruption of lifestyle, better present state of health, more satisfaction with total quality of life, and more satisfaction with life in general.

It was hypothesised that the WBQ would discriminate between patients receiving insulin by injection and pump; it was found that pump patients experienced significantly better scores on all three versions of the WBQ. Age was not significantly associated with WBQ scores. Men scored higher on all WBQ scales than women, reaching significance when the data from individual centres were combined (Bech et al., op. cit.).

The W-BQ12 has been compared with the Hospital Anxiety and Depression Scale/HADS (Zigmond and Snaith, op. cit.), State Trait Anxiety Inventory/STAI (Spielberger et al. op. cit.), Center for Epidemiological Studies Depression Scale/CES-D (Radloff, op. cit.), and the SF-36 (Ware and Sherbourne, op. cit.; Pouwer, Snoek et al., op. cit.). Correlations between the W-BQ12 dimensions and the HADS and STAI were in the range 0.52-0.82. Correlations between the W-BQ12 dimensions and the CES-D and SF-36 were in the range 0.22-0.85.

The study also assessed which dimensions were most useful in discriminating between patients who were under the care of a psychologist or psychiatrist and their counterparts, patients who were suffering from depression and their counterparts, and patients suffering from chronic fatigue and their counterparts. The four W-BQ12 dimensions produced significant differences for each of the comparisons and produced higher odds ratios than the HADS and STAI for the second comparison. Women had significantly higher levels of negative well-being and anxiety, and significantly lower levels of energy and general well-being.

Responsiveness

The WBQ has not been formally assessed for responsiveness, but a number of studies have been cited where the instrument has been used in longitudinal evaluations as evidence for responsiveness (Bradley, 1994 op. cit.). The first study included NIDDM patients invited to take part in a trial of insulin therapy using either CSII pumps or

injections. Patients not meeting the selection criteria for this study were invited to attend an education session. WBQ scores showed insignificant and small levels of deterioration following the education session. WBQ scores showed no change for patients with poorly-controlled diabetes who were switched to two forms of insulin therapy. It was concluded that the introduction of insulin therapy led to significant improvements in metabolic control without any deterioration in psychological well-being.

The second study compared CSII with injection treatment for IDDM patients as part of a multi-centre study and results are presented for French and German centres. In the French centre, where patients had been using the pumps for at least two months following conventional insulin treatment, the WBQ scales of depression, positive well-being and energy showed significant improvements. Similar patterns of change were found during the crossover phase of the study, and the energy subscale showed a significant difference. In the German centre, CSII produced significant improvements in energy, positive well-being and general well-being, compared with conventional insulin treatment.

Acceptability

In the original evaluation of the WBQ, 187 patients (85.4%) returned a completed questionnaire; of these 184 completed the WBQ (98.4%) (Bradley and Lewis, *op. cit.*). There were no differences between responders and non-responders in age, sex, duration of diabetes, HbA_{1c} or percent body weight. In the second study, the response rate was 84% (*ibid.*; Bradley, 1994 *op. cit.*).

Other studies have reported response rates of 81% (after a single reminder) among older patients (Pettersen *et al.*, *op. cit.*) and 75% with no differences between responders and non-responders (Bardsley *et al.*, *op. cit.*). Wredling *et al.* (*op. cit.*) reported that 83.6% of those invited agreed to participate.

The W-BQ12 study produced a 49% response rate for the first questionnaire; of these, 81% agreed to participate in the test-retest study (Pouwer, van der Ploeg *et al.*, *op. cit.*).

Commentary

The WBQ and its shorter-forms have been tested on a range of patients with varying ages, type of diabetes and treatment modalities. There is good evidence for the dimensionality of the instrument. The internal consistency reliability of the WBQ is generally good but one or two scales have been shown to perform badly in different populations (Bradley, 1994 *op. cit.*; Wredling *et al.*, *op. cit.*). Initial results suggest the W-BQ12 may perform better (Pouwer, van der Ploeg *et al.*, *op. cit.*). Test-retest reliability correlations were acceptable for use in groups (Pouwer, Snoek *et al.*, *op. cit.*). The lack of a priori hypotheses limits the interpretation of the results of validity testing in all studies.

WBQ scores have been shown to be related to the presence of complications (Bradley and Lewis, *op. cit.*). Moderate levels of correlation were found between WBQ scores and responses to a number of global questions relating to perceived health, disease impact, and life satisfaction (Wredling *et al.*, *op. cit.*; Bech *et al.*, *op. cit.*). Positive well-being scores were moderately correlated with those for the Nottingham Health

Profile (Bardsley et al, op cit.). On the whole, WBQ scores have been shown to be unrelated to HbA_{1c} levels which is considered a desirable property of the instrument (Bradley and Lewis, op. cit.). The moderate to large levels of correlation with a number of instruments measuring related constructs are evidence for the validity of the W-BQ12 (Pouwer, Snoek et al., op. cit.).

Bradley (1996) has argued that while a short, unidimensional form of the WBQ would be useful for routine clinical monitoring, the strategy used by Bech and colleagues (op. cit.) to reduce the WBQ is conceptually and methodologically unsound. The combining of data from centres using different translations of the instrument is criticised, particularly in the light of translation inadequacy. Little attention was given to the content of items selected and the wording of one of the items was changed without justification.

There has been no formal assessment of the responsiveness of the WBQ and its variants, but the instrument has been used in a number of clinical trials and the results have been commented upon (Bradley, 1994 op. cit.).

Chapter 5: CONCLUSIONS

There are two broad approaches to measuring health outcomes from the perspective of the patient: generic instruments that aim to include aspects of health and quality of life of relevance to the general population, and disease-specific instruments of relevance to patients with a particular disease. For the purposes of measuring health outcomes within diabetes, a third approach has been suggested that is situation-specific (Beaser et al., op. cit.). Situation-specific instruments are designed to measure the effects of specific interventions on knowledge, attitudes and behaviour. Interventions designed to influence these factors are only desirable if they improve the quality of life of patients. Therefore evaluation should include some measure of self-perceived health related quality of life.

This review has focused on disease-specific instruments measuring components of quality of life and diabetic symptoms. The review was based on a comprehensive search strategy applied to the relevant electronic databases, and the hand-searching of relevant diabetes and health outcomes literature. The references were downloaded into a database and assessed against inclusion criteria. The authors of instruments meeting these criteria were contacted and asked for further information relating to development, testing and instrument scoring.

The search strategy produced 20 instruments that were included in the review. Two instruments did not strictly meet the inclusion criteria. The Well-being Questionnaire is generic in focus but has had considerable application in diabetes. The Quality of Life: Status and Change instrument is a generic instrument but has a disease-specific component for patients with IDDM (Hörnquist et al., 1993 op. cit.). The majority of instruments has been developed and evaluated in the United States and Europe for use in patients with both types of diabetes. Four instruments have undergone development and evaluation in a UK patient population, namely the Audit of Diabetes Dependent Quality of Life/ADDQoL (Bradley, Todd et al., op. cit.), the Diabetes Health Profile/DHP-1 (Meadows et al., op. cit.), the Newcastle Diabetes Symptoms Questionnaire/NDSQ (McCull et al., op. cit.), and the Well-being Questionnaire (Bradley, 1994 op. cit.).

The content of the twenty instruments covered ten domains relating to aspects of health-related quality of life and diabetic symptoms. Most of the instruments reviewed include some diabetic symptoms in their content, and two instruments have symptoms as their focus. The greatest proportion of items related to the domains of symptoms, general quality of life and diabetes, psychological well-being, quality of life and treatment, and social well-being. There was considerable variation in the spread of items across the remainder of the domains.

With the exception of the two instruments which include a substantial generic component (the Well-being Questionnaire and the Quality of Life: Status and Change instrument) the majority of the instruments reviewed have just one published evaluation relating to their measurement properties. Of the strictly disease-specific instruments, the Diabetes Quality of Life Measure/DQoL (Jacobson, op. cit.) has the largest number of published evaluations, and has evidence for its reliability and validity across four nationalities of diabetic patients.

The three disease-specific instruments developed in the UK, namely the ADDQoL, the DHP-1, and the NDSQ, have not been evaluated as extensively as the DQoL but have undergone a more rigorous approach to development. For example, the DQoL has not been assessed for dimensionality using principal component analysis or factor analysis. The evidence for the measurement properties of the three instruments compares favourably with that for the other instruments reviewed. Just three of the remaining instruments with comparable evidence for reliability and validity have undergone a similar process of development, namely the Diabetes-specific Quality of Life Scale (Bott et al., *op. cit.*), Diabetes-39 (Boyer and Earp, *op. cit.*) and the Well-being Enquiry for Diabetics (Mannucci, *op. cit.*).

Very few instruments have been evaluated for responsiveness to change. Only two instruments have been formally assessed for responsiveness using an effect size statistic, namely an earlier version of the Diabetes Health Profile (DHP-1) and the Newcastle Diabetes Symptoms Questionnaire (NDSQ). The instruments were evaluated as part of the same longitudinal study using the standardised response mean to quantify responsiveness (Whitty et al., *op. cit.*). The DHP and NDSQ produced SRMs representing small and large levels of change respectively.

The three UK instruments are based on different approaches to patient-reported health outcomes. The DHP-1 is a health profile based on the traditional psychometric approach to instrument development. The ADDQoL incorporates an individualised importance weighting for each of 13 items. The NDSQ is a more narrowly focused measure of diabetes symptoms. The three instruments are relatively brief, making them suitable for inclusion alongside generic and situation-specific instruments as part of a package of outcome measures. The DHP-1 and the ADDQoL have greater content validity from the perspective of the patient and are therefore recommended as potential primary outcome measures for clinical and health services research, including clinical trials. Where a clinically relevant responsive measure of symptoms is required, it is recommended that the nine-item NDSQ be used alongside these instruments.

For comparability across conditions and for purposes of economic evaluation, it is recommended that, where practical, these instruments are used alongside a generic instrument with evidence for its measurement properties within a diabetic population. Furthermore, such an instrument has greater potential to measure unexpected effects, including the side-effects of treatment. The content of the SF-36 has been described as being mostly relevant to diabetes patients and there is evidence to support this suggestion (Bradley, 1996 *op. cit.*; McColl, Steen et al., *op. cit.*). Furthermore, the SF-36 includes scales that measure physical functioning and role limitations, domains covered by several instruments included in the review but not by the DHP-1. Where a detailed evaluation of psychological well-being is required, it is recommended that the WBQ (Bradley, 1994 *op. cit.*) or short-form version of the instrument is used.

There is still limited evidence for the usefulness of patient-reported measures of health outcome within clinical practice (Greenhalgh and Meadows, 2000). None of the instruments included in the review has the recommended levels of reliability for use in individual patients. Where an instrument is required for clinical practice at the group level, for example within clinical audit, it is recommended that the content of the ADDQoL, DHP-1 and NDSQ are assessed for relevance.

Further research is needed to evaluate the measurement properties of these instruments. The DHP and ADDQoL have evidence of internal consistency but neither instrument has been assessed for test-retest reliability. This should be conducted within a short time-frame, with a supporting health transition question to assess whether patients' diabetes and general health has remained stable between administrations. For purposes of assessing validity the instruments should be compared to generic instruments such as the SF-36, and specific instruments measuring similar constructs. The latter could include an anglicised version of the DQoL. Such testing should be supported by a priori hypotheses.

Further comparison should be made between instrument scores, and clinical and sociodemographic variables with a priori hypotheses informed by previous research (Rubin and Peyrot, *op. cit.*). Responsiveness to change has been the most neglected area of instrument evaluation within diabetes. Instruments should be administered longitudinally before and after changes in treatment known to improve health-related quality of life. Again, health transition questions should be included as an external criterion for assessing whether patient health has really changed. Finally, instruments should be assessed concurrently so that relative responsiveness can be assessed. This research will inform decisions regarding the selection of instruments for applications, including clinical trials.

REFERENCES

- Anderson, J. S., Sullivan, F., and Usherwood, T. P. (1990). The Medical Outcomes Study Instrument (MOSI) - use of a new health status measure in Britain. *Fam.Pract* **7**, 205-218.
- Anderson, R. M., Fitzgerald, J. T., Wisdom, K., Davis, W. K., and Hiss, R. G. (1997). A comparison of global versus disease-specific quality-of-life measures in patients with NIDDM. *Diabetes-Care* **20**, 299-305.
- Avis, N. E. and Smith, K. W. (1994). Conceptual and methodological issues in selecting and developing quality of life measures. In "Advances in medical sociology" (R. Fitzpatrick, Ed.), pp. 255-280. London: JAI Press Inc.
- Baker, G. A. and Meadows, K. (1993). A review of quality of life assessment: Diabetes mellitus. *Practical-Diabetes* **10**, 18-20.
- Bardsley, M. J., Astell, S., McCallum, A., and Home, P. D. (1993). The performance of three measures of health status in an outpatient diabetes population. *Diabetic-Medicine* **10**, 619-626.
- Beaser, R. S., Garbus, S. B., and Jacobson, A. M. (1996). Diabetes Mellitus. In "Quality of life in clinical trials" (B. Spiler, Ed.), pp. 983-991. Philadelphia: Lippincott-Raven Publishers.
- Bech, P., Gudex, C., and Johansen, K. S. (1996). The WHO (Ten) Well-being index: Validation in diabetes. *Psychotherapy and Psychosomatics* **65**, 183-190.
- Beck, A. T., Steer, R. A., and Garbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: twenty-five years of evaluation. *Clin. Psychol. Rev.* **8**, 77-100.
- Bergner, M., Bobbitt, R. A., Carter, W. B., and Gilson, B. S. (1981). The Sickness Impact Profile: development and final revision of a health status measure. *Med Care* **19**, 787-805.
- Bott, U., Mühlhauser, I., Overmann, H., and Berger, M. (1998). Validation of a diabetes-specific quality-of-life scale for patients with type 1 diabetes. *Diabetes-Care* **21**, 757-769.
- Bowling, A. (1995). Other disease and condition specific scales. In "Measuring disease" (A. Bowling, Ed.), pp. 259-267. Buckingham: Open University Press.
- Boyer, J. G. and Earp, J. A. (1997). The development of an instrument for assessing the quality of life of people with diabetes. Diabetes-39. *Med.Care* **35**, 440-453.
- Bradley, C. (1996). The WHO (Ten) Well-being Index: a critique letter. *Psychother.Psychosom* **65**, 331-333.
- Bradley, C. and Lewis, K. S. (1990). Measures of psychological well-being and treatment satisfaction developed from the responses of people with tablet-treated Type 2 diabetes. *Diabetic-Medicine* **7**, 445-451.

Bradley, C., Pierce, M., Hendrieckx, C., Riazi, A., and Barendse, S. (1998). Diabetes Mellitus. In "Comprehensive Clinical Psychology" (A. S. Bellack and M. Hersen, Eds.), pp. 277-304.

Bradley, C., Todd, C., Gorton, T., Symonds, E., Martin, A., and Plowright, R. (1999). The development of an individualized questionnaire measure of perceived impact of diabetes on quality of life: The ADDQoL. *Quality-of-Life-Research* **8**, 79-91.

Bradley, C. E., Ed. (1994). "Handbook of psychology and diabetes: A guide to psychological measurement in diabetes research and practice." Chur, Switzerland: Harwood Academic Publishers.

Brazier, J. E., Harper, R., Jones, N. M., O'Cathain, A., Thomas, K. J., Usherwood, T., and Westlake, L. (1992). Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *BMJ* **305**, 160-164.

Brownlee-Duffeck, M., Peterson, L., Simonds, J. F., Goldstein, D., Kilo, C., and Hoette, S. (1987). The role of health beliefs in the regimen adherence and metabolic control of adolescents and adults with diabetes mellitus. *J Consult Clin.Psychol* **55**, 139-144.

Bryant, F. B. and Veroff, J. (1984). Dimensions of subjective mental health in American men and women. *Journal of Health and Social Behavior* **25**, 116-135.

Carey, M. P., Jorgensen, R. S., Weinstock, R. S., Sprafkin, R. P., Lantinga, L. J., Carnrike-CL, J., Baker, M. T., and Meisler, A. W. (1991). Reliability and validity of the appraisal of diabetes scale. *J Behav Med.* **14**, 43-51.

Charlton, J. R., Patrick, D. L., and Peach, H. (1983). Use of multivariate measures of disability in health surveys. *J Epidemiol.Community Health* **37**, 296-304.

Cohen, S., Kamarck, T., and Mermelstein, R. (1983). A global measure of perceived stress. *J Health Soc.Behav* **24**, 385-396.

Cox, D. J. and Gonder-Frederick, L. (1992). Major developments in behavioral diabetes research. *J Consult Clin.Psychol* **60**, 628-638.

Cox, D. J., Irvine, A., Gonder-Frederick, L., Nowacek, G., and Butterfield, J. (1987). Fear of hypoglycemia: quantification, validation, and utilization. *Diabetes Care* **10**, 617-621.

Cox, W. M., Blount, J. P., Crowe, P. A., and Singh, S. P. (1996). Diabetic patients' alcohol use and quality of life: Relationships with prescribed treatment compliance among older males. *Alcoholism: Clinical and Experimental Research* **20**, 327-331.

Cutrona, C., Russell, D., and Rose, J. (1986). Social support and adaptation to stress by the elderly. *Psychol Aging* **1**, 47-54.

Davis, W. K., Hess, G. E., Van Harrison, R., and Hiss, R. G. (1987). Psychosocial adjustment to and control of diabetes mellitus: differences by disease type and treatment. *Health Psychol.* **6**, 1-14.

- Derogatis, L., et al. (1973). The SCL-90: An outpatient psychiatric rating scale - a preliminary report. *Psychopharmacology Bulletin* **9**, 13-18.
- Dunn, S. M., Smartt, H. H., Beeney, L. J., Turtle, J. R. (1986). Measurement of emotional adjustment in diabetic patients: validity and reliability of the ATT39. *Diabetes Care* **9**, 480-489.
- Duran, G., Herschbach, P., Waadt, S., Strian, F., and Zettler, A. (1995). Assessing daily problems with diabetes: a subject-oriented approach to compliance. *Psychol.Rep.* **76**, 515-521.
- Eiser, C. and Tooke, J. E. (1995). Quality of life in type II diabetes: evaluation and applications. *PharmacoEconomics* **8 Suppl 1**, 17-22.
- Fitzgerald, J. T., Anderson, R. M., Gruppen, L. D., Davis, W. K., Aman, L. C., Jacober, S. J., and Grunberger, G. (1998). The reliability of the Diabetes Care Profile for African Americans. *Evaluation-and-the-Health-Professions* **21**, 52-65.
- Fitzgerald, J. T., Davis, W. K., Connell, C. M., Hess, G. E., Funnell, M. M., and Hiss, R. G. (1996). Development and validation of the diabetes care profile. *Evaluation-and-the-Health-Professions* **19**, 208-230.
- Fitzpatrick, R., Davey, C., Buxton, M. J., and Jones, D. R. (1998). Evaluating patient-based outcome measures for use in clinical trials. *Health Technol.Assess.* **2**, i-74.
- Gilden, J. L., Casia, C., Hendryx, M., and Singh, S. P. (1990). Effects of self-monitoring of blood glucose on quality of life in elderly diabetic patients. *Journal-of-the-American-Geriatrics-Society* **38**, 511-515.
- Gilden, J. L., Hendryx, M., Casia, C., and Singh, S. P. (1989). The effectiveness of diabetes education programs for older patients and their spouses. *Journal of the American Geriatrics Society* **37**, 1023-1030.
- Gilden, J. L., Hendryx, M. S., Clar, S., Casia, C., and Singh, S. P. (1992). Diabetes support groups improve health care of older diabetic patients. *Journal of the American Geriatrics Society* **40**, 147-150.
- Given, C. W. (1984). Measuring the social psychological health states of ambulatory chronically ill patients: hypertension and diabetes as tracer conditions. *J Community Health* **9**, 179-195.
- Goddijn, P., Bilo, H., Meadows, K., Groenier, K., Feskens, E., and Meyboom, d. J. (1996). The validity and reliability of the Diabetes Health Profile (DHP) in NIDDM patients referred for insulin therapy. *Quality-of-Life-Research* **5**, 433-442.
- Goddijn, P. P., Bilo, H. J., Feskens, E. J., Groeniert, K. H., Van der Zee, K. I., and Meyboom, d. J. (1999). Longitudinal study on glycaemic control and quality of life in patients with Type 2 diabetes mellitus referred for intensified control. *Diabet.Med.* **16**, 23-30.

- Greenhalgh, J., and Meadows K. (1999). The effectiveness of the use of patient-based measures of health in routine practice in improving the process and outcomes of patient care: a literature review. *Journal of Evaluation in Clinical Practice* **5**, 401-416.
- Grootenhuis, P. A., Snoek, F. J., Heine, R. J., and Bouter, L. M. (1994). Development of a type 2 diabetes symptom checklist: A measure of symptom severity. *Diabetic-Medicine* **11**, 253-261.
- Gulliford, M. C. and Mahabir, D. (1999). Relationship of health-related quality of life to symptom severity in diabetes mellitus: a study in Trinidad and Tobago. *Journal of Clinical Epidemiology* **52**, 773-780.
- Guyatt, G. H. (1993). The philosophy of health-related quality of life translation. *Qual.Life Res* **2**, 461-465.
- Guyatt, G. H., Veldhuyzen Van Zanten, S. J., Feeny, D. H., and Patrick, D. L. (1989). Measuring quality of life in clinical trials: a taxonomy and review. *CMAJ*. **140**, 1441-1448.
- Haak, T., Bott, U., Freitlag, D., Farac, J., Siegmund, T., Usadel, K. (1997). Verbessert eine strukturierte stationäre Schulung die Lebensqualität und Therapiezufriedenheit von Patienten mit insulinpflichtigem Diabetes mellitus? *Diabetes Stoffwechsel* **16**, 70-71.
- Hamilton, M. (1960). A rating scale for depression. *J. Neurol. Neurosurg. Psychiatry* **23**, 56-65.
- Hammen, C. and Mayol, A. (1982). Depression and cognitive characteristics of stressful life-event types. *J Abnorm.Psychol* **91**, 165-174.
- Hammond, G. S. and Aoki, T. T. (1992). Measurement of health status in diabetic patients: Diabetes impact measurement scales. *Diabetes-Care* **15**, 469-477.
- Hanestad, B. R. and Albrektsen, G. (1991). Quality of life, perceived difficulties in adherence to a diabetes regimen, and blood glucose control. *Diabetic-Medicine* **8**, 759-764.
- Hanestad, B. R. and Albrektsen, G. (1992). The stability of quality of life experience in people with Type 1 diabetes over a period of a year. *Journal-of-Advanced-Nursing* **17**, 777-784.
- Hanestad, B. R., Hörnquist, J. O., and Albrektsen, G. (1991). Self-assessed quality of life and metabolic control in persons with insulin-dependent diabetes mellitus (IDDM). *Scandinavian Journal of Social Medicine* **19**, 57-65.
- Hanestad, B. R. and Albrektsen, G. (1993). The effects of participation in a support group on self-assessed quality of life in people with insulin-dependent diabetes mellitus. *Diabetes-Research-and-Clinical-Practice* **19**, 163-173.
- Henderson, M. and Freeman, C. P. (1987). A self-rating scale for bulimia. The 'BITE'. *Br.J Psychiatry* **150**, 18-24.

- Henderson, M. J. and Tindall, H. (1990). Evaluation of consumer satisfaction and quality of life in patients changing to Novopen II. *Practical-Diabetes* **7**, 206-208.
- Herschbach, P., Duran, G., Waadt, S., Zettler, A., Amm, C., and Marten-Mittag, B. (1997). Psychometric properties of the questionnaire on stress in patients with diabetes – Revised (QSD-R). *Health Psychology* **16**, 171-174.
- Hess, G. E., Davis, W. K., and Harrison, R. V. (1986). A diabetes psychosocial profile. *Diabetes Educator* **12**, 135-140.
- Home, P., Coles, J., Goldacre, M., Mason, A., Wilkinson, E. (eds). Health Outcome Indicators: Diabetes. Report of a working group to the Department of Health. Oxford: National Centre for Health Outcomes Development, 1999.
- Hörnquist, J. O. (1989). Quality of life: Concept and assessment. *Scandinavian-Journal-of-Social-Medicine* **18**, 69-79.
- Hörnquist, J. O., Wikby, A., Hansson, B., and Andersson, P. O. (1993). Quality of life: Status and change (QLsc) reliability, validity and sensitivity of a generic assessment approach tailored for diabetes. *Quality-of-Life-Research* **2**, 263-279.
- Hörnquist, J. O., Wikby, A., Stenstrom, U., Andersson, P. O., and Akerlind, I. (1995). Type II diabetes and quality of life: a review of the literature. *PharmacoEconomics* **8 Suppl 1**, 12-16.
- Hunt, S. M., McKenna, S. P., McEwen, J., Williams, J., and Papp, E. (1981). The Nottingham Health Profile: subjective health status and medical consultations. *Soc.Sci Med [A]* **15**, 221-229.
- Jacobson, A., Barofsky, I., Cleary, P., and Rand, L. (1988). Reliability and validity of a diabetes quality-of-life measure of the Diabetes Control and Complications Trial (DCCT). *Diabetes-Care* **11**, 725-732.
- Jacobson, A. M. (1997). Quality of life in patients with diabetes mellitus. *Semin.Clin.Neuropsychiatry* **2**, 82-93.
- Jacobson, A. M., De Groot, M., and Samson, J. A. (1994). The evaluation of two measures of quality of life in patients with type I and type II diabetes. *Diabetes-Care* **17**, 267-274.
- Jacobson, A. M., De Groot, M., and Samson, J. A. (1997). The effects of psychiatric disorders and symptoms on quality of life in patients with type I and type II diabetes mellitus. *Qual.Life Res* **6**, 11-20.
- Jacobson, A. M., Hauser, S. T., Willett, J. B., Wolfsdorf, J. I., Dvorak, R., Herman, L., and De Groot, M. (1997). Psychological adjustment to IDDM: 10-year follow-up of an onset cohort of child and adolescent patients. *Diabetes-Care* **20**, 811-818.
- Jacobson, A. M. and Diabetes Control and Complications Trial Research Group (1994). The Diabetes Quality of Life Measure. In "Handbook of psychology and diabetes: A guide to psychological measurement in diabetes research and practice" (C. Bradley, Ed.), pp. 65-87. Chur,Switzerland: Harwood Academic Publishers.

- Jenkinson, C. (1995) Measuring the outcomes of medical care: possibilities and limitations. *Soc. Sci. Med.* **41**, 1395-1401.
- Kanner, A. D., Lee, D. K., Rountree, L., and Wilson, A. M. (1986). Relationship of psychological stress and glycaemic control in diabetic children. *Diabetes* **35**, 443.
- Kotsanos, J. G., Vignati, L., Huster, W., Andrejasich, C., Boggs, M. B., Jacobson, A. M., Marrero, D., Mathias, S. D., Patrick, D., Zalani, S., and Anderson, J. (1997). Health-related quality-of-life results from multinational clinical trials of insulin lispro. Assessing benefits of a new diabetes therapy. *Diabetes Care* **20**, 948-958.
- Mannucci, E., Ricca, V., Bardini, G., and Rotella, C. M. (1996). Well-being enquiry for diabetics: A new measure of diabetes-related quality of life. *Diabetes, -Nutrition-and-Metabolism-Clinical-and-Experimental* **9**, 89-102.
- McCull, E., Steen, I. N., Meadows, K. A., Hutchinson, A., Eccles, M. P., Hewison, J., Fowler, P., and Blades, S. M. (1995). Developing outcome measures for ambulatory care - An application to asthma and diabetes. *Social-Science-and-Medicine* **41**, 1339-1348.
- McDowell, I. and Jenkinson, C. (1996). Development standards for health measures. *J Health Serv Res Policy* **1**, 238-246.
- McDowell, I. and Newell, C. (1996). "Measuring Health: A Guide to Rating Scales and Questionnaires". 2nd Ed. Oxford University Press, New York.
- Meadows, K., Steen, N., McCull, E., Eccles, M., Shiels, C., Hewison, J., and Hutchinson, A. (1996). The diabetes health profile (DHP): A new instrument for assessing the psychosocial profile of insulin requiring patients - Development and psychometric evaluation. *Quality-of-Life-Research* **5**, 242-254.
- Meenan, R. F., Gertman, P. M., and Mason, J. H. (1980). Measuring health status in arthritis. The arthritis impact measurement scales. *Arthritis Rheum.* **23**, 146-152.
- Nathan, D. M., Fogel, H., Norman, D., Russell, P. S., Tolkoff-Rubin, N., Delmonico, F. L., Auchincloss, H., Jr., Camuso, J., and Cosimi, A. B. (1991). Long-term metabolic and quality of life results with pancreatic/renal transplantation in insulin-dependent diabetes mellitus. *Transplantation* **52**, 85-91.
- Nouwen, A., Gingras, J., Talbot, F., and Bouchard, S. (1997). The development of an empirical psychosocial taxonomy for patients with diabetes. *Health Psychol.* **16**, 263-271.
- O'Boyle, C. A., McGee, H., Hickey, A., O'Malley, K., and Joyce, C. R. (1992). Individual quality of life in patients undergoing hip replacement. *Lancet* **339**, 1088-1091.
- Parkerson, H., Jr., G. R., Connis, R. T., Broadhead, W. E., Patrick, D. L., Taylor, T. R., and Tse, C. K. (1993). Disease-specific versus generic measurement of health-related quality of life in insulin-dependent diabetic patients. *Medical Care* **31**, 629-639.

- Petterson, T., Lee, P., Hollis, S., Young, B., Newton, P., and Dornan, T. (1998). Well-being and treatment satisfaction in older people with diabetes. *Diabetes-Care* **21**, 930-935.
- Pouwer, F., Snoek, F. J., van der Ploeg, H. M., Heine, R. J., and Brand, A. N. (1998). A comparison of the standard and the computerized versions of the Well-being Questionnaire (WBQ) and the Diabetes Treatment Satisfaction Questionnaire (DTSQ). *Qual.Life Res* **7**, 33-38.
- Pouwer, F., van der Ploeg, H. M., Ader, H. J., Heine, R. J., and Snoek, F. J. (1999). The 12-item Well-Being Questionnaire: an evaluation of its validity and reliability in Dutch people with diabetes. *Diabetes-Care* **22**, 2004-2010.
- Radloff, L. S. (1977). The CES-D Scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement* **1**, 385-401.
- Rodin, G. (1990). Quality of life in adults with insulin-dependent diabetes mellitus. *Psychotherapy and Psychosomatics* **54**, 132-139.
- Rubin, R. R. and Peyrot, M. (1999). Quality of life and diabetes. *Diabetes/Metabolism-Research-and-Reviews* **15**, 205-218.
- Ruige, J. B., de Neeling, J. N., Kostense, P. J., Bouter, L. M., and Heine, R. J. (1997). Performance of an NIDDM screening questionnaire based on symptoms and risk factors (SR-Q). *Diabetes Care* **20**, 491-496.
- Ruta, D. A., Garratt, A. M., Wardlaw, D., and Russell, I. T. (1994). Developing a valid and reliable measure of health outcome for patients with low back pain. *Spine* **19**, 1887-1896.
- St. Vincent Declaration (1989). *Diabetic Medicine* **7**, 360.
- Selam, J. L., Micossi, P., Dunn, F. L., and Nathan, D. M. (1992). Clinical trial of programmable implantable insulin pump for type I diabetes. *Diabetes Care* **15**, 877-885.
- Shen, W., Kotsanos, J. G., Huster, W. J., Mathias, S. D., Andrejasich, C. M., and Patrick, D. L. (1999). Development and validation of the Diabetes Quality of Life Clinical Trial Questionnaire. *Med.Care* **37**, AS45-AS66.
- Slevin, M. L., Plant, H. J., Lynch, D., Drinkwater, J., and Gregory, W. M. (1988). Who should measure quality of life, the doctor or the patient? *Br J Cancer* **57**, 109-112.
- Spielberger, C. D., Gorsuch, R. L., and Lushene, R. E. (1970). STAI manual. Consulting Psychologist Press: Palo Alto.
- Steen, N. and McColl, E. (1995). Developing and testing symptom-based outcome measures. In "Health Outcome Measures for Primary and Outpatient Care" pp. 23-44.
- Stewart, A. L., Greenfield, S., Hays, R. D., Wells, K., Rogers, W. H., Berry, S. D., McGlynn, E. A., and Ware, J. E., Jr. (1989). Functional status and well-being of

patients with chronic conditions. Results from the Medical Outcomes Study. *JAMA* **262**, 907-913.

Streiner, D. L. and Norman, G. R. (1995). "Health Measurement Scales: A Practical Guide to their Development and Use". 2nd Ed. Oxford University Press: New York.

Talbot, F., Nouwen, A., Gingras, J., Gosselin, M., and Audet, J. (1997). The assessment of diabetes-related cognitive and social factors: the Multidimensional Diabetes Questionnaire. *J Behav Med.* **20**, 291-312.

Van der Does, F. E. E., de Neeling, J. N. D., Snoek, F. J., Kostense, P. J., Grootenhuys, P. A., Bouter, L. M., and Heine, R. J. (1996). Symptoms and well-being in relation to glycemic control in type II diabetes. *Diabetes-Care* **19**, 204-210.

Vaughan, N. J. A., Bradshaw, C., Bradley, C., Brown, F., Cleary, R., Clements, D., Dixon, P., Gadsby, R., Home, D., Hopkins, A. P., Hutchinson, A., Jolley, S., Keen, H., Kopelman, P., Kyne-Grebalski, D., Meadows, K., Murphy, M., Pierce, M., Redmond, S., and et al. (1994). Measuring the outcomes of diabetes care. *Diabetic Medicine* **11**, 418-423.

Von Zerssen, D. (1976). Die Beschwerden-Liste. In "Klinische Selbstbeurteilungsskalen (KSb-S) aus dem Münchner Psychiatrischen Informationssystem (PSYCHIS-München)". Beltz: Weinheim.

Ware, J. E., Jr. (1976). Scale for measuring general health perceptions. *Health Services Research* **11**, 396-415.

Ware, J. E., Jr., Johnston, S. A., Davies-Avery, A. and Brook, R. H. (1979). Conceptualisation and Measurement of Health for Adults in the Health Insurance Study: Mental Health. Vol. 3. Santa Monica: Rand Corporation.

Ware, J. E., Jr. and Sherbourne, C. D. (1992). The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* **30**, 473-483.

Whitty, P., Steen, N., Eccles, M., McColl, E., Hewison, J., Meadows, K., Clapp, Z., and Hutchinson, A. (1997). A new self-completion outcome measure for diabetes: Is it responsive to change? *Quality-of-Life-Research* **6**, 407-413.

Wierenga, M. E. (1994). Life-style modification for weight control to improve diabetes health status. *Patient Educ.Couns.* **23**, 33-40.

Wikby, A., Stenstrom, U., Andersson, P., and Hörnquist, J. (1998). Metabolic control, quality of life, and negative life events: a longitudinal study of well-controlled and poorly regulated patients with Type 1 diabetes after changeover to insulin pen treatment. *Diabetes-Educator* **24**, 61-66.

Wilkin, D., Hallam, L., and Doggett, M.A. (1992). "Measures of need and outcome for primary health care". Oxford: Oxford University Press.

Williams, D. R. (1992). A proposal for continuing audit of diabetes services. Home and Members of a Working Group of the Research Unit of the Royal College of

Physicians and British Diabetic Association see comments. *Diabetic Medicine* **9**, 759-764.

Wilson, A. E., Home, P. D., and for the Diabetes Audit Working Group of the Research Unit of the Royal College of Physicians and the British Diabetic Association (1993). A dataset to allow exchange of information for monitoring continuing diabetes care. *Diabetic Medicine* **10**, 390.

Woodend, A. K., Nair, R. C., and Tang, A. S. (1997). Definition of life quality from a patient versus health care professional perspective. *International Journal of Rehabilitation Research* **20**, 71-80.

Wredling, R., Stalhammar, J., Adamson, U., Berne, C., Larsson, Y., and Ostman, J. (1995). Well-being and treatment satisfaction in adults with diabetes: A Swedish population-based study. *Quality-of-Life-Research* **4**, 515-522.

Zigmond, A. S. and Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica* **67**, 361-70.

Appendix: measures excluded from the review

Barriers

Barriers to Adherence Questionnaire

Glasgow, R. E., Hampson, S. E., Strycker, L. A., and Ruggiero, L. (1997). Personal-model beliefs and social-environmental barriers related to diabetes self-management. *Diabetes Care* **20**, 556-561.

Barriers to Self Care Scale

Glasgow, R. E. (1994). Social-environmental factors in diabetes: Barriers to diabetes self-care. In Bradley, C. E. "Handbook of psychology and diabetes: A guide to psychological measurement in diabetes research and practice." Chur, Switzerland: Harwood Academic Publishers.

Diabetes Fear of Injecting and Self-testing/D-FISQ

Snoek, F. J., Mollema, E. D., Heine, R. J., Bouter, L. M., and van der Ploeg, H. M. (1997). Development and validation of the diabetes fear of injecting and self-testing questionnaire (D-FISQ): First findings. *Diabetic Medicine* **14**, 871-876.

Environmental Barriers to Adherence Scale/EBAS

Irvine, A. A., Saunders, J. T., Blank, M. B., and Carter, W. R. (1990). Validation of scale measuring environmental barriers to diabetes-regimen adherence. *Diabetes Care* **13**, 705-711.

Experience of Treatment Benefits and Barriers/ETBB

Lewis, K. S. and Bradley, C. E. (1994) Measures of diabetes-specific health beliefs. In Bradley, C. E. "Handbook of psychology and diabetes: A guide to psychological measurement in diabetes research and practice." Chur, Switzerland: Harwood Academic Publishers.

Ipswich Diabetes Self-Management Questionnaire

Day, J. L., Bodmer, C. W., and Dunn, O. M. (1996). Development of a questionnaire identifying factors responsible for successful self-management of insulin-treated diabetes. *Diabetic Medicine* **13**, 564-573.

Summary of Diabetes Self-care Activities Questionnaire

Toobert, D. J. and Glasgow, R. E. (1994) Assessing diabetes self-management. In Bradley, C. E. "Handbook of psychology and diabetes: A guide to psychological measurement in diabetes research and practice." Chur, Switzerland: Harwood Academic Publishers.

Beliefs

Diabetes Health Belief Scale

Harris, R., Linn, M. W., and Pollack, L. (1984). Relationship between health beliefs and psychological variables in diabetic patients. *Br.J.Med.Psychol* 253-259.

Health Belief Model Scale

Given, C. W. and Given, B. A. (1984). Development of scales to measure beliefs of diabetic patients. *Research in Nursing and Health* **6**, 127-141.

Hurley, A. C. (1990). The Health Belief Model: Evaluation of a Diabetes Scale. *The Diabetes Educator* **16**, 44-48.

Control

Diabetes Empowerment Scale/DES

Anderson, R. M., Funnell, M. M., Fitzgerald, J. T., and Marrero, D. G. (2000). The Diabetes Empowerment Scale: a measure of psychosocial self-efficacy. *Diabetes Care* **23**, 739-743.

Diabetes Locus of Control Scale

Ferraro, L. A., Price, J. H., Desmond, S. M., and Roberts, S. M. (1987). Development of a diabetes locus of control scale. *Psychol-Rep* **61**, 763-770.

Diabetes Onset Locus of Control Scales/DOLoC

Bradley, C. E. "Handbook of psychology and diabetes: A guide to psychological measurement in diabetes research and practice." Chur, Switzerland: Harwood Academic Publishers (1994).

Perceived Control of Diabetes Scale/PCDS

Bradley, C., Brewin, C. R., Gamsu, D. S., and Moses, J. L. (1984). Development of scales to measure perceived control of diabetes mellitus and diabetes-related health beliefs. *Diabetic Medicine* **1**, 213-218.

Coping

Adaptation to Diabetes Scale/ADS

Lernmark, B., Persson, B., Fisher, L., and Rydelius, P.-A. (1999). Symptoms of depression are important to psychological adaptation and metabolic control in children with diabetes mellitus. *Diabetic Medicine* **16**, 14-22.

Diabetes Coping Measure/DCM

Welch, G. W. In Bradley, C. E. "Handbook of psychology and diabetes: A guide to psychological measurement in diabetes research and practice." Chur, Switzerland: Harwood Academic Publishers (1994).

Issues for Coping with IDDM scale

Grey, M., Boland, E. A., Davidson, M., Yu, C., Sullivan, B. S., and Tamborlane, W. V. (1998). Short-term effects of coping skills training as adjunct to intensive therapy in adolescents. *Diabetes Care* **21**, 902-908.

Knowledge

Diabetes Knowledge Questionnaire/CCQ

Meadows, K. A., Fromson, B., Gillespie, C., Brewer, A., Carter, C., Lockington, T., Clark, G., and Wise, P. H. (1988). Development, validation and application of computer-linked knowledge questionnaires in diabetes education. *Diabet.Med* **5**, 61-67.

Diabetes Knowledge Scales/DKN

Beeney et al.(1994) In Bradley, C. E. "Handbook of psychology and diabetes: A guide to psychological measurement in diabetes research and practice." Chur, Switzerland: Harwood Academic Publishers.

Patient Knowledge Test

Garrard, J., Joynes, J. O., Mullen, L., McNeil, L., Mensing, C., Feste, C., and Etwiler, D. D. (1987). Psychometric study of patient knowledge test. *Diabetes Care* **10**, 500-509.

Non-English

Elderly Diabetes Impact Scales/EDIS) (*Japanese*)

Araki, A., Izumo, Y., Inoue, J., Takahashi, R., Takanashi, K., Teshima, T., Yatomi, N., Shimizu, Y., and Ito, H. (1995). Development of Elderly Diabetes Impact Scales (EDIS) in elderly patients with diabetes mellitus. *Nippon Ronen Igakkai Zasshi* **32**, 786-796.

Diabetes Emotional Adjustment Scale (*Spanish*)

De Leon, O. A. (1995). Development, reliability and validation of the Diabetes Emotional Adjustment Scale in Spanish: preliminary findings. *Int J Psychiatry Med* **25**, 81-92.

Psychological/attitudinal

The ATT-39

Welch, G., Dunn, S. M., and Beeney, L. J. (1994). The ATT39: A measure of psychological adjustment to diabetes. In Bradley, C. E. "Handbook of psychology and diabetes: A guide to psychological measurement in diabetes research and practice." Chur, Switzerland: Harwood Academic Publishers (1994).

Attitudes towards diabetes and self-care

Wikblad, K. F., Wibell, L. B., and Montin, K. R. (1990). The patient's experience of diabetes and its treatment: construction of an attitude scale by a semantic differential technique. *Journal of Advanced Nursing* **15**, 1083-1091.

Diabetes Attitude Scale/DAS-2

Anderson, R. M., Donnelly, M. B., and Dedrick, R. F. (1990). Measuring the attitudes of patients towards diabetes and its treatment. *Patient Educ. Couns.* **16**, 231-245.

Fear of Hypoglycaemia/HFS-II

Cox, D. J., Irvine, A., Gonder-Frederick, L., Nowacek, G., and Butterfield, J. (1987). Fear of hypoglycemia: quantification, validation, and utilization. *Diabetes Care* **10**, 617-621.

Ideas about Diabetes – Revised/IAD-R

Dion, G. A. (1990). Testing of an instrument to measure acceptance of diabetes: "Ideas about Diabetes-Revised" (IAD-R). Thesis/Dissertation.

Problem Areas in Diabetes Survey/PAIDS

Polonsky, W. H., Anderson, B. J., Lohrer, P. A., Welch, G., Jacobson, A. M., Aponte, J. E., and Schwartz, C. E. (1995). Assessment of diabetes-related distress. *Diabetes Care* **18**, 754-760.

Support

Diabetes-specific Family Behaviour Scale/DFBS

Waller, D.A., Chipman, J. J., Hardy, B. W., Hightower, M. S., et al. (1986). Measuring diabetes-specific family support and its relation to metabolic control: A preliminary report. *Journal of the American Academy of Child Psychiatry* **25**, 415-418.

Diabetes Family Behaviour Checklist/DFBC

Schafer, L. C., McCaul, K. D., and Glasgow, R. E. (1986). Supportive and nonsupportive family behaviors: relationships to adherence and metabolic control in persons with type I diabetes. *Diabetes Care* **9**, 179-185.

Symptom measures

Bagne et al Questionnaire

Bagne, C. A., Luscombe, F. A., Damiano, A. (1995). Relationships between glycemic control, diabetes-related symptoms, and SF-36 scales scores in patients with non-insulin dependent diabetes mellitus. *Quality-of-Life-Research* **4**, 392-1. *Abstract*.

Diabetic Clinic Pain Survey

Chan, A. W., MacFarlane, I. A., Bowsher, D., Wells, J. C., Bessex, C., and Griffiths, K. (1990). Chronic pain in patients with diabetes mellitus: comparison with a non-diabetic population. *The Pain Clinic* **3**, 147-159.

Diabetes and Lower Extremity Ulcers Questionnaire

Brod, M. (1998). Quality of life issues in patients with diabetes and lower extremity ulcers: Patients and care givers. *Quality-of-Life-Research* **7**, 365-372.

Diabetic Neuropathy Questionnaire/DNQ (German)

Schandry, R. and Leopold, C. (1998). On the validity of patients' symptom reporting in diabetics neuropathy. *Medizinische-Welt* **49**, 169-175.

Edinburgh Hypoglycaemia Scale

Deary, I. J., Hepburn, D. A., MacLeod, K. M., and Frier, B. M. (1993). Partitioning the symptoms of hypoglycaemia using multi-sample confirmatory factor analysis. *DIABETOLOGIA* **36**, 771-777.

Neuroqol

Vileikyte, L., Bundy, C., Tomenson, B., Walsh, T., and Boulton, A. J. (1998). Neuropathy-Specific Quality of Life Measure: Construction of Scales and Preliminary Tests of Reliability and Validity (Abstract #24). American Diabetes Association. (*Conference Proceedings*)

NIDDM Screening Questionnaire Based on Symptoms and Risk Factors/SR-Q

Ruige, J. B., de Neeling, J. N., Kostense, P. J., Bouter, L. M., and Heine, R. J. (1997). Performance of an NIDDM screening questionnaire based on symptoms and risk factors. *Diabetes Care* **20**, 491-496.

Seven Sexual Symptoms/SSS

Abel, G. G., Becker, J. V., Cunningham-Rathner, J., Mittelman, M., and Primack, M. (1982). Differential diagnosis of impotence in diabetics: the validity of sexual symptomatology. *Neurology & Urodynamics* **1**, 57-69. 1982.

Symptom questionnaire

Johansen, J., Claudi, T., and Holtedahl, K. (1999). Insulin treatment for poorly regulated diabetic patients in general practice. Better regulation and symptom relief? *Scand.J Prim.Health Care* **17**, 244-249.

Youth measures**Diabetes Adjustment Scale/DAS**

Sullivan, B. J. (1979). Adjustment in diabetic adolescent girls: I. Development of the Diabetic Adjustment Scale. *Psychosomatic Medicine* **41**, 119-126.

Diabetes Hassles Scale

Kanner, A. D., Lee, D. K., Rountree, L., and Wilson, A. M. (1986). Relationship of psychological stress and glycaemic control in diabetic children. *Diabetes* **35**, 443.

Modified quality-of-life measure for youths/DQoLY)

Ingersoll, G. M. and Marrero, D. G. (1991). A modified quality-of-life measure for youths: psychometric properties. *Diabetes Educ.* **17**, 114-118.

Self-care inventory

la Greca, A. M., Follansbee, D., and Skyler, J. S. (1990). Developmental and behavioral aspects of diabetes management in youngsters. *Children's Health Care* **19**, 132-139.

Self-efficacy for Diabetes Scale

Grey, M., Boland, E. A., Davidson, M., Yu, C., Sullivan, B. S., and Tamborlane, W. V. (1998). Short-term effects of coping skills training as adjunct to intensive therapy in adolescents. *Diabetes Care* **21**, 902-908.

Other

Health Status Measure for older African-American women with type 2 diabetes

Elasy, T. A., Samuel-Hodge, C. D., DeVellis, R. F., Skelly, A. H., Ammerman, A. S., and Keyserling, T. C. (2000). Development of a health status measure for older African-American women with type 2 diabetes. *Diabetes-Care* **23**, 325-329.

Parents Diabetes Quality of Life Questionnaire

Faulkner, M. S. and Clark, F. S. (1998). Quality of life for parents of children and adolescents with type 1 diabetes. *Diabetes-Educator* **24**, 721-727.

Personal Models of Diabetes Interview/PMDI¹⁰

Hampson, S. E., Glasgow, R. E., and Foster, L. S. (1995). Personal models of diabetes among older adults: relationship to self-management and other variables. *Diabetes-Educator* **21**, 300-307.

Diabetes Clinic Satisfaction Questionnaire (DCSQ)

Treatment Satisfaction Questionnaire (DTSQ)

Bradley, C. E. "Handbook of psychology and diabetes: A guide to psychological measurement in diabetes research and practice." Chur, Switzerland: Harwood Academic Publishers.

UK Prospective Diabetes Study Group Questionnaire/UKPDS-37¹¹

Mehta, Z. (1999). Quality of life in type 2 diabetic patients is affected by complications but not by intensive policies to improve blood glucose or blood pressure control. *Diabetes-Care* **22**, 1125-1136.

¹⁰ Not for self-completion

¹¹ Although claiming to be specific, the questionnaire used by the UKPDS Group was in fact derived from various questionnaires that were not disease-specific.