Quality of Life Research, 4, pp. 115–134

Advances in methods for assessing the impact of epilepsy and antiepileptic drug therapy on patients' health-related quality of life

A. K. Wagner,* S. D. Keller, M. Kosinski, G. A. Baker, A. Jacoby, M.-A. Hsu, D. W. Chadwick and J. E. Ware Jr

The Health Institute, New England Medical Center Box # 345, 750 Washington Street, Boston, MA 02111, USA; Tel.: (617) 636-8623; fax: (617) 636-8628 (A. K. Wagner, S. D. Keller, M. Kosinski, J. E. Ware Jr); University Department of Neurosciences, Walton Hospital, Liverpool, UK (G. A. Baker, D. W. Chadwick); Center for Health Services Research, University of Newcastle Upon Tyne, UK (A. Jacoby); Schering-Plough Corporation, Kenilworth, NJ, (M.-A. Hsu); School of Public Health, Harvard University, Boston, MA, USA (J. E. Ware Jr)

We studied 31 previously validated and newly developed generic and epilepsy-specific scales to evaluate their usefulness for assessing the impact of epilepsy and anti-epileptic drug (AED) therapy on health-related quality of life (HRQOL). Included were the MOS SF-36 Health Survey, additional measures of mental health, cognition, epilepsy-specific perception of control, behavioural problems, distress, worries and experiences, the Liverpool Epilepsy Impact and Seizure Severity scales, and a patient-completed symptom checklist. Questionnaires were completed twice by 136 patients on AED therapy in a multicentre study in the UK. Validity was assessed in relation to disease severity, defined as time since last seizure, and to patient-reported symptoms. Statistical analyses to estimate the contribution of HRQOL information of each scale relative to that of others were conducted. The 171-item questionnaire could be completed by out-patients with epilepsy with good data quality. With few exceptions, generic and epilepsy-specific measures satisfied psychometric tests of hypothesized item groupings and scale score reliability (internal consistency and test–retest reliability) and differentiated well between groups of patients differing in time since last seizure and in symptom impact, regardless of time since last seizure. However, scales differed widely in their validity in discriminating between groups of patients known to differ clinically. The SF-36 Role Physical scale best discriminated among groups differing in disease severity. The epilepsy-specific Mastery, Impact, Experience, Worry, Distress, and Agitation scales were among the 10 best measures in discriminating among groups differing in disease severity. Generic measures, especially measures of social and role functioning and mental health, were best at differentiating groups of patients differing in symptom impact. Recommendations are offered for concepts and specific scales most likely to be useful in future studies of the HRQOL burden of epilepsy and the HRQOL benefits of AED therapy.

Key words: Epilepsy, health-related quality of life, health status assessment, MOS SF-36 Health Survey, seizure control, seizure severity, symptoms, validity.

Introduction

Epilepsy and anti-epileptic drug (AED) therapy have a major impact on the daily lives of patients. It has been suggested that the psychosocial problems observed among patients with epilepsy are more handicapping than the seizures themselves.[1] The majority (61%) of people with epilepsy who responded to a recent Roper poll indicated that seizures prevent them from doing things they want to do. They reported significant problems with holding jobs, seeking work, and going to school.[2]

AED therapy, by decreasing seizure frequency and possibly severity, has the potential to ameliorate the psychosocial consequences of the disease. However, therapy may itself cause new problems in daily living because of adverse effects, interactions with other drugs, frequent blood sampling, feelings of dependency on a potentially life-long medication regimen, and financial cost associated with long-term therapy. Half of the respondents to the 1992 Roper Poll indicated that they were unhappy about the side effects caused by medications, and 59% felt that they had no choice but to accept the level of seizure control and side effects from medications.[2]

* To whom correspondence should be addressed.

While clinicians and researchers have long been concerned about the effects of epilepsy and its treatment on the daily lives of people,[3] these effects are not routinely monitored in either clinical trials or daily practice. How individuals perceive that seizures and/or adverse effects from antiepileptic medication influence the quality of their lives is only assessed informally. In the past, a major barrier to the routine standardized assessment of the health-related quality of life (HRQOL) of patients with epilepsy has been the lack of measurement tools that are practical, reliable and valid for use in this population.

The goal of medical care for most patients is the achievement of a more effective life and the preservation of functioning and well-being.[4] The patient is the best source of information regarding the achievement of these goals.[5] His or her point of view in monitoring the quality of medical care outcomes has increasingly been recognized[6] and standardized patient surveys have been developed to assess it. During the past 50 years, both the techniques of constructing patient-based outcomes measures and their content have been greatly advanced. Psychometric theory and methods to develop HRQOL measures have led to the development of reliable and valid measures of functioning and well-being, namely patient-completed HRQOL questionnaires.[7,8]

HRQOL questionnaires can be broadly divided into generic and specific questionnaires. Generic measures assess concepts that are relevant to everyone, including ability to function in everyday life and emotional well-being.[8] They are not specific to any age, disease, or treatment group. Since generic measures are designed to be broadly applicable across populations which differ in diseases and treatments, they allow for comparisons across different conditions.[9,10] Thereby, they provide the opportunity to estimate the HRQOL burden of a particular condition in relationship to that of another, or to compare the HRQOL of people with a specific disease to that of people without a chronic condition.

Specific measures are designed to focus on the impact of a particular disease and its treatment or on specific concepts and domains.[11] Rather than choosing one type of measure, a modular approach, including generic and disease-specific batteries, has been recommended as the most desirable course in assessing HRQOL.[11,12] Through this approach, more concepts can be represented.

Recent advances have contributed to the development of several new instruments to assess the impact of epilepsy on patients' lives.[13–16] However, there are still unanswered questions. What is the best mix of generic and specific measures? Can HRQOL measures detect the HRQOL impact of both, seizures and symptoms related to treatments? What is the burden of epilepsy relative to that of other conditions? What is the benefit of antiepileptic therapy in HRQOL terms? Can HRQOL measures differentiate the benefits of different treatments?

The goal of this study was to evaluate generic and specific measures of HRQOL to determine their usefulness in assessing the HRQOL burden of epilepsy and the burden and benefits of treatments. We replicated previous analyses of a generic HRQOL measure, the Medical Outcomes Study (MOS) 36-item Health Survey (SF-36)[17,18] and of epilepsy-specific measures. We examined the ability of generic and specific HRQOL measures to discriminate between groups of patients differing in seizure control. We also examined the ability of generic and specific HRQOL measures to discriminate between groups of patients differing in symptoms often associated with AED treatment. We extended the range of epilepsy-specific domains of HRQOL assessed. Finally, we estimated how much each HRQOL scale contributes to the differentiation between groups differing in time since last seizure and self-reported symptom impact.

Methods

Questionnaire development

Domains of life affected by either epilepsy and/or AED therapy were identified based on clinical experience, informal discussions with patients and a review of the psychosocial and clinical literature, including information about developmental AED. Where available, we chose previously validated scales to assess relevant domains. New items and scales were created to assess domains for which no previously developed measures were found. The study questionnaire contained 171 items which were used to construct 31 multi-item scales and a symptom checklist. They were divided into the following four modules: (1) the SF-36 as a core measure of general health, (2) additional generic measures of health, (3) epilepsy- and AED therapy-specific measures, and (4) items to measure the occurrence and impact of symptoms. Appendix 1 lists the scales and subscales, and numbers of items per scale, and identifies their sources.

Measures

General health core measure. The SF-36 includes one multi-item scale measuring each of eight health con-

cepts: physical functioning, role limitations due to physical health problems, bodily pain, general health perceptions, vitality (energy/fatigue), social functioning, role limitations due to emotional health problems, and mental health (psychological distress and well-being). In addition, one item asks for a comparison of present health with health one year ago.[17] The UK version of the SF-36, which modifies the wording of six items[18–22] was used. Three of the SF-36 scales were augmented by adding one item each. To reduce the percentage of people scoring at the lowest scale level ('floor'), the two SF-36 Role Functioning scales were augmented with one additional question each which asks if the respondent is unable to work or perform other regular daily activities at all. These scales were scored with and without the additional items to maintain comparability with published data and to evaluate their added value. One item ('I have been feeling bad lately')[23,24] was added to the assessment of general health perceptions to replicate previous tests of the Current Health scale.[23,24] In light of previous data,[14] it was hypothesized that a scale of perceptions of current health alone would have a stronger relationship to time since last seizure than the more general SF-36 General Health scale which combines current and future health perceptions.

Additional measures of general health. To evaluate whether the effects of epilepsy and its treatment are concentrated in particular psychological distress and well-being concepts, the five-item Mental Health scale of the SF-36 (MHI-5) was supplemented with additional items from the 18-item Mental Health Inventory (MHI-18, which includes the MHI-5),[25,26] and the UCLA Loneliness scale.[27] This permitted tests of whether any of the mental health subscales are more related to epilepsy. A global item to address general well-being[25,28] was included for patients to indicate their satisfaction and happiness with life in general.

Proven cognitive functioning items selected from the Sickness Impact Profile (SIP),[25,29] the Psychiatric Epidemiology Research Interview (PERI),[25,30] and the MOS Cognitive Functioning scale[25] were included. The dichotomous response choices of the SIP items were changed to categorical ratings to increase the number of levels of functioning that can be differentiated.

Epilepsy-specific measures. The epilepsy-specific module included measures of feelings of personal control, the impact of epilepsy and its treatment on daily life, distress and behavioural problems related to epilepsy, as well as a measure of patient-perceived seizure severity. The generic Internal Locus of Control scale[31] which has successfully been used in studies of epilepsy,[32,33] was modified to create an epilepsy-specific Mastery scale addressing patients' feelings of helplessness and control with regard to their epilepsy and their seizures. A list of specific areas of epilepsy and AED treatment on patients lives was created. We included a slightly modified version of the previously validated Liverpool Epilepsy Impact scale[34] to address the impact of epilepsy and AED therapy on the individual's relationships with friends and family, social life, employment, health, self esteem, plans for the future, and standard of living. Newly created epilepsy experience questions, focusing on effects not captured in the Liverpool Impact scale, were added. These covered epilepsy-related fears, embarrassment, restrictions, and the effects of having to take medications regularly and to have regular blood tests. Newly created epilepsy worry questions addressed patients' worries about work, social life, and medication effects, in relationship to their epilepsy and AED treatment.

Based on the four health distress items from the MOS,[24] we constructed a two-item Epilepsy Distress scale that asks about despair, discouragement, fears and worries due to epilepsy. Two questions from the Health Insurance Study[35] were modified to ask patients about behavioural problems due to their epilepsy.

To define severity of the disease from the patient's point of view, the Liverpool Seizure Severity scale[36] was included. Two subscales, the Ictal and Percept scales measure patients' perceptions of ictal and post-ictal effects and of control over seizures, respectively, during the preceding 4 weeks. Patients with more than one seizure type complete the scale twice, once for what they identify as their 'minor' seizures and once for their 'major' seizures.

Symptom checklist. Sixteen items were added to measure the frequency and impact of symptoms often associated with currently available or investigational AED therapy. Symptoms which were captured by separate scales such as cognitive impairment, mood disturbances, and sedation were not repeated in the Symptom Checklist. The Symptom Checklist described symptoms in lay terms. Patients indicated whether or not they had double or blurred vision, slurred speech or trouble with their speech, unsteadiness, involuntary movements, tremors or shaky hands, trouble sleeping, slowed reactions, headaches, weight gain or loss, stomach upset or nausea, vomiting, change of hair texture or hair loss, skin problems such as acne or rash, impotence, or decreased libido. Response choices to symptom questions were (1) No, I do not have the symptom, (2) Yes, but not a problem, (3) Yes, somewhat of a problem, and (4) Yes, a big problem.

Symptom responses were also analysed dichotomously (no, yes). The questionnaire ended with two open-ended questions to identify any additional epilepsy- or AED treatment-related issues that bother patients.

Questionnaire scoring. When tests of scaling assumptions justified doing so, we used Likert's method of summated ratings[37] to score each scale. Items were recoded where required to achieve the same direction of scoring within the same scale. Raw scale scores were computed by summing across items in the same scale. Except for the Seizure Severity scales, if $\leqslant$ 50% of items were missing, a person-specific estimate (mean of non-missing items) was substituted for the missing items. Except for the Seizure Severity and Epilepsy Impact scales (to maintain consistency with published data for these scales), raw scores were transformed to a 0–100 scale, where a score of 0 indicates the least favourable state of health, 100 indicates the most favourable state of health, and scores in between represent a percentage of the total possible score in each HRQOL domain.

Patient sample

One hundred and thirty-six ambulatory patients from three clinics in the UK [Walton Hospital, Liverpool (n = 87), Bootham Park Hospital, York (n = 29), and Chorley Health Center, Chorley (n = 20)] participated in the study. Patients were included if they were $\geqslant$ 15 years old, had a diagnosis of partial onset seizures (PS) or primarily generalized tonic-clonic seizures (GTC), received carbamazepine (CBZ) or valproic acid (VPA) monotherapy (patients on polytherapy were included if they had more than 3 PS or 1 GTC during the past 3 months) and had no change in their AED therapy for at least 3 months prior to enrollment. Patients were excluded if they lacked basic proficiency in reading English (as assessed informally by the investigator), had undergone epilepsy surgery during the past 12 months, received non-antiepileptic drug therapy with known central nervous system effects, or required treatment for a concurrent psychiatric illness. Patients gave oral informed consent.

Fifty-four males and 82 females with a mean age of 34.9 years (range 15–78 years), a mean age of seizure onset of 21.8 years (range 1–71), and a mean duration of epilepsy of 14.1 years (range 0–56 years) participated in the study. According to the International League Against Epilepsy (ILAE) 1981 Seizure Classification,[38] 79% of patients had partial onset seizures. Twenty-one percent had primarily generalized tonic-clonic seizures. Information about primarily generalized tonic-clonic seizures was deduced from patients' ILAE Classification of Epileptic Syndromes,[39] since primarily generalized tonic-clonic seizures were not specified in the data collection form. The majority of patients (59.7%) had experienced their most recent seizure during the 3 months prior to the HRQOL assessment. At least one medical or psychological problem in addition to epilepsy was identified in the medical history of 108 (80%) patients. VPA or CBZ monotherapy was the AED regimen for 35% and 37% of patients, respectively. Polytherapy regimens included either CBZ or VPA and either vigabatrin or lamotrigine. About half of the study patients were married and living with a spouse or significant other. Twenty-three percent were working full-time, 28.2% were unemployed, and the rest were working part-time, working in unpaid positions (as a student or homemaker), or were retired.

Questionnaire administration

Patients completed the HRQOL questionnaire once during a regularly scheduled clinic visit. Demographic and clinical information (medical history, seizure history and classification, AED therapy, non-AED therapy and adverse events) was collected on standardized case report forms during this visit. To estimate test–retest reliability,[40] the questionnaires were mailed to all patients about 14 days after the first assessment to complete at home. A 2-week time interval was chosen so that the likelihood of new events occurring between assessments would be minimized. However, the interval would be long enough to minimize recall of former responses to the questionnaire. Almost all (131) patients completed the second assessment, on average 12.5 days after the first.

Statistical analysis

Three basic criteria were used to evaluate all HRQOL scales: practicality, psychometric tests of scaling assumptions and construct validity, and empirical validity in relation to clinical criteria. Practicality refers to the ease with which patients could complete the questionnaire as evidenced by time to completion as well as the completeness of data and consistency of response patterns. The psychometric tests of the questionnaire pertain to how the scales should be scored, whether they define enough levels, and the reliability of the scores. For the SF-36 scales, factor analysis was performed to evaluate construct validity in light of

what is known about the factor structure of the SF-36 from previous studies.[41,42] Validity of all scales was assessed in relation to both disease severity defined as time since last seizure and the frequency of symptoms possibly associated with antiepileptic drug treatment.

Data quality and respondent burden. Quality of the data was evaluated in two ways: (1) the percentage of patients who completed all items within each scale as well as the percentage of patients who had computable HRQOL subscales (i.e. those for whom at least half of the items were completed), and (2) the frequency of inconsistent responses across 15 pairs of SF-36 items in the Response Consistency Index.[18,43] The impact of educational level on data quality was also assessed. Missing data, logical response patterns, test—retest correlations and internal consistency reliability statistics were calculated for patients differing in educational level: patients with < 11 years of education, patients with 11 years of education, and patients with $\geqslant 12$ years of education. To assess respondent burden, the time patients needed to complete the questionnaire was calculated from patient-indicated starting and completion times, written on the questionnaire; and patients' responses to an open-ended comment question were evaluated.

Psychometric evaluation. Most psychometric analyses were conducted to identify those scales which had good measurement properties for this sample of epilepsy patients. The augmented SF-36 Role Functioning scales were hypothesized to achieve a lower percentage of people scoring at the lowest level.

Summated-rating scale assumptions. The method of summated ratings assumes that items in the same scale can be aggregated without score standardization or item weighting.[37] To avoid standardization, items should have roughly equivalent means and standard deviations. To avoid weighting, items should be equally representative (that is, have roughly equivalent relationships to) the underlying scale dimension. In addition, items should correlate substantially (> 0.40[44] corrected for overlap[45]) with their hypothesized scales.

Item discriminant validity and scaling success rates. Item discriminant validity is supported when the correlation between each item and its hypothesized scale is larger than its correlation with competing scales. Tests of item discriminant validity were summarized into item scaling success rates which are the percentage of successful tests. (Differences between correlations of two standard errors were considered significant. The standard error of the correlation is approximately equal to the reciprocal of the square root of the sample size.)

Reliability of scale scores. Internal consistency reliability of each scale score was estimated using Cronbach's alpha.[46] Test–retest reliability[40] was estimated by computing the product-moment and intraclass correlation coefficients to assess the relationship between scores from the same scale over a 2-week period. Scale reliabilities of $\geqslant 0.50$[44] or $\geqslant 0.70$[47] have been suggested for scales under development for use in group level analyses, whereas scale reliabilities of $\geqslant 0.90$ have been suggested for scales used with individual data.[47]

Features of score distributions. Scale score distributions were evaluated by computing the percentage of people achieving either the highest score (ceiling effect) or lowest score (floor effect). Scales with a large percentage of people scoring at either the floor or ceiling may not be appropriate for this patient sample.

Psychometric validity. It was hypothesized, based on previous work,[41,42] that physical and mental higher-order health constructs would explain the great majority of the covariance between SF-36 scale scores. Evidence for these higher-order health constructs was obtained using principal components analysis. Besides the noteworthy statistical advantages which are well-documented in the literature,[48,49] there are practical reasons for choosing principal components analysis over factor analysis. First, the results would maintain comparability with previous studies and second, principal components analysis leads to results which are robust across methods of extraction and rotation. Two components were extracted from the correlations among SF-36 scale scores and were rotated to orthogonal simple structure using the varimax method.[50] The pattern of scale–factor correlations was compared with previous studies.[41,42] No hypotheses were formulated for the higher-order structure of the entire battery. Therefore, principal components analysis was performed only on SF-36 scales.

Validity in relation to clinical criteria. The validity of each scale was assessed by comparing groups known to differ in terms of two clinical variables: time since last seizure and symptom impact. Two hypotheses were tested: (1) people with epilepsy who have been seizure-free for longer periods of time have better HRQOL than those who have experienced seizures more recently; (2) people who experience symptoms

have lower HRQOL than those who are symptom-free. With one exception, hypotheses were not formulated regarding which scales would be most sensitive to clinical criteria. The exception was the Current Health scale, which was hypothesized to be more sensitive to seizure recency than the SF-36 General Health scale.

Correlational analyses. To test the hypothesis that people who experience symptoms have lower HRQOL than those who are symptom-free, mean HRQOL scale scores for those with and without each of the 16 reported symptoms were compared.

Analysis of variance. Multivariate analyses of variance (MANOVA) were conducted to control for inflation of Type I error rate due to multiple comparisons. Univariate analyses of variance (ANOVA) to determine which scales differentiated among clinically defined groups were only performed when MANOVA tests were significant. To test the hypothesis that people with epilepsy who have been seizure-free have better HRQOL than those who have experienced seizures more recently, mean HRQOL scores were compared on all scales for four mutually exclusive groups of patients who differed according to the time since their last seizure: patients who had at least one seizure during the week preceding the HRQOL assessment; patients who were seizure-free for > 1 week, but 3 months at most, patients who were seizure-free for > 3 months but 6 months at most and patients who were seizure-free for > 6 months (seizure-free patients).

In the clinical management of patients with epilepsy, the dose of usually one AED is titrated to seizure control or toxicity, whichever comes first. Patients whose seizures are not controlled may have achieved their highest tolerable level of one or more AED. Likewise, seizure-free patients may be controlled at the expense of experiencing side effects from their AED regimen. Therefore, seizure status confounds the analysis of symptom impact in the entire study population. To evaluate the usefulness of the HRQOL scales in detecting the impact of symptoms apart from the HRQOL impact of seizures, we compared the mean HRQOL scores of patients who experienced each symptom to those free of the symptom within two subgroup of patients: those who were seizure-free for at least the past 6 months and those who had at least one seizure during the past week. Symptom scores were dichotomized and mean scores on HRQOL scales were compared between the two groups. The HRQOL scores of patients answering 'no' to whether they had a symptom were compared to those of all patients who answered 'yes', regardless of the degree to which they were bothered by the symptom.

Analyses were conducted on adjusted (controlling for age, gender and education) and unadjusted means. Since the results for adjusted and unadjusted means were equivalent (data available upon request), only the analyses for unadjusted means are reported.

Relative validity. The usefulness of the HRQOL scales in detecting the effects of seizures and of symptoms was compared by computing relative validity (RV) coefficients.[42,51,52] Each RV coefficient is a ratio of two F-statistics. The F-statistic for the test of unadjusted mean differences in each scale in comparing groups of patients differing in time since last seizure was divided by the F-statistic for the best scale. RV provides an estimate of how much more or less valid each scale is relative to the best scale in a test involving discriminating between groups differing in clinical characteristics (known groups validity[53]). RV coefficients are higher when a scale captures a larger difference between groups and/or estimates group means with less error.

Incremental validity. To evaluate whether the epilepsy-specific scales add to the variance explained in comparisons of groups differing in disease severity as defined by the time since last seizure over and above that explained by general health status measures, we conducted tests of incremental validity.[54] For each of the epilepsy-specific scales (except the seizure severity scales), we tested the significance of the increment in variance explained with each epilepsy-specific scale beyond that explained by the generic SF-36 health status scales.

Results

Respondent burden and data quality

Patients required approximately 40 min on average (s.d. 23 min) to complete the 171-item study questionnaire. Two open-ended questions were included, as well as a section for comments, which was responded to by 86 patients (63%). Patients' comments were mostly favourable, indicating that they appreciated being asked about how their epilepsy affected their lives. Two respondents criticized the length of the questionnaire. Four mentioned that some questions were confusing or complicated. Six respondents reported problems deciding whether some questions asked about their epilepsy or other health problems, or a particular seizure type.

Data quality was high. Missing data were rare; at

Table 1. Descriptive statistics and features of score distributions for HRQOL scales

HRQOL Scales	K[1]	Mean	Median	Range of observed scores	Standard deviation	% Ceiling	% Floor
General HRQOL–SF-36 scales							
Physical functioning	10	83.5	95.0	10–100	23.6	38.6	0.0
Role physical	4	67.1	100.0	0–100	39.5	51.2	17.3
Bodily pain	2	78.2	84.0	0–100	25.9	48.8	0.9
General health perceptions	5	64.3	67.0	5–100	24.8	4.8	0.0
Vitality	4	57.1	60.0	0–100	22.9	2.4	0.8
Social functioning	2	75.1	75.0	12–100	26.2	39.4	0.0
Role emotional	3	66.7	100.0	0–100	42.0	57.5	21.3
Mental health (MHI-5)	5	66.5	72.0	8–100	21.0	1.6	0.0
Augmented role physical[2]	5	71.3	100.0	0–100	35.2	50.8	6.3
Augmented role emotional[2]	4	71.6	100.0	0–100	36.9	57.1	10.3
General HRQOL – additional scales							
Mental health (MHI-18)	17	71.4	72.9	10.5–98.8	18.9	0.0	0.0
Anxiety	4	67.0	70.0	0–100	22.0	8.6	0.8
Depression	4	72.5	75.0	15–100	20.1	11.7	0.0
Behavioural/emotional control	5	75.1	76.0	20–100	20.4	14.8	0.0
Positive well-being	5	59.9	66.7	5–100	21.8	2.3	0.0
Emotional ties	4	80.6	85.0	25–100	19.3	21.9	0.0
Current health	4	60.9	65.0	0–100	26.2	4.7	1.6
MOS cognition	6	76.8	86.7	3–100	23.4	17.2	0.0
Confusion	2	84.7	90.0	10–100	21.0	44.5	0.0
Thinking	2	68.4	75.0	0–100	25.6	14.8	2.3
Concentration	2	61.2	62.5	0–100	26.4	15.1	4.0
Attention	2	76.6	80.0	10–100	22.4	31.3	0.0
Psychomotor function	3	80.0	86.7	0–100	21.4	26.6	1.6
Epilepsy-specific HRQOL scales							
Mastery	6	54.7	55.6	6–94	18.6	0.0	0.0
Impact[3]	8	14.7	13.0	8–32	6.1	0.8	21.1
Experience	13	75.6	79.5	20–100	19.7	7.8	0.0
Worry	9	80.1	86.1	19–100	19.0	17.2	0.0
Agitation	2	71.2	70.0	0–100	27.1	26.0	3.1
Distress	2	74.7	80.0	0–100	26.0	26.8	1.6
Seizure severity scales							
Percept[4]	8	23.7	24.0	16–32	3.3	1.6	0.0
Ictal[5]	12	25.7	26.0	12–42	8.3	0.0	0.0

n = 136 total sample size for all scales except Seizure Severity Scales (n = 65).
[1] K = number of items per scale/subscale.
[2] Augmented role scales include one additional item to lower the floor of the distribution.
[3] Not linearly transformed, higher score indicates greater impact of epilepsy (range = 8–32).
[4] Not linearly transformed, higher score indicates more severe seizures (range = 7–32).
[5] Not linearly transformed, higher score indicates more severe seizures (range = 10–48).

least 89% of items were completed for each scale and at least 91% of patients had computable data for each scale using scoring rules documented elsewhere.[55] Completely consistent responses across 15 pairs of items in the SF-36 were provided by 85.2% of patients. This is slightly lower than the percentage observed for a representative sample of the US population (90.3%).[18] There was no difference in the percentage of missing data, the consistency of responses, or test–retest and internal consistency reliability coefficients across groups differing in educational level.

Psychometric properties of scale data

Table 1 presents descriptive statistics and score distributions for the HRQOL scales. (Note that the MHI-18 was estimated from 17 items for some analyses in order to evaluate the item referring to loneliness as part of the Emotional Ties scale). Higher standard deviations were observed for the SF-36 Role Functioning scales, as was the case in the US general population[18] and among chronically ill patients from the MOS.[52] The distributions of responses to most scales

indicated that patients used the full range of responses.

Although a fairly high percentage of patients scored at the top of the SF-36 Role Functioning scales, the percentages of epilepsy patients at the ceiling were lower than in the general population.[18] Noteworthy ceiling effects were also observed for the SF-36 Bodily Pain and the generic Confusion scales, indicating that few patients with epilepsy have problems with these health concepts. Comparisons of the percentages of patients scoring at the floor on the original and augmented SF-36 Role Functioning scales indicate that the added items lowered the percentages of patients scoring at the floor (from 17.3% to 6.3% for Role Physical, and from 21.3% to 10.3% for Role Emotional; see Table 1).

Table 2 summarizes results of tests of scaling assumptions and presents reliability estimates. Correlations between items and their scales supported the hypothesized item groupings for 27 of the 30 scales, including all of the SF-36 scales. Results were also favourable for all additional generic scales with the exception of the cognition scales. Hypothesized item groupings were not supported for the Thinking and Attention scales. All epilepsy-specific scales passed scaling tests with the exception of the Percept scale. The aggregation of the Thinking, Attention or Percept items into scales was not supported.

Results generally support the reliability of scale scores from the questionnaire. With the exception of the Percept scale, the minimum criterion for internal consistency reliability for scales under early evaluation (Cronbach's alpha > 0.50[44]) were met for all scales (coefficients ranged from 0.43 to 0.92). Coefficients equalled or exceeded the more stringent 0.7 criterion[47] for 28 of 30 scales. Those scales with the lowest internal consistency were the Attention (0.51) and Percept (0.43) scales. Except for those scales which were heterogeneous in content (such as the Percept scale), test–retest reliabilities, while sufficiently high, were generally somewhat lower than internal consistency estimates (test–retest correlations 0.55–0.89). These results suggest that some change may have occurred between administrations over time.

Consistent with previous research on the SF-36,[41,42] the principal components analysis confirmed the two-factor higher-order structure of the SF-36 scales (data not reported). These components were interpreted as physical and mental health based on their relationship to SF-36 scales (i.e., the Physical Functioning scale loaded highest on the 'physical' component and the Mental Health scale loaded highest on the 'mental' component). These results, which are available from the first author upon request, provide evidence of the construct validity of the SF-36 in the epilepsy population in the UK.

Validity in relation to clinical criteria

Comparisons among patient groups differing in time since last seizure. Patients who were seizure-free for > 6 months tended to have higher (more favourable) average HRQOL scores than patients with more recent seizures, as hypothesized (see Table 3). There was a general trend for those patients who experienced seizures to have lower HRQOL scale scores than those who were seizure-free, with the HRQOL decrement being directly related to the time since last seizure: the more recent the seizures the more pronounced the effect on HRQOL. Differences between seizure-free patients and those experiencing a seizure in the past week were most pronounced. When compared to seizure-free patients, patients who had at least one seizure during the preceding week had significantly lower HRQOL scores on all SF-36 scales, all additional measures of psychological distress and well-being, both Current and General Health scales, four of the six measures of cognitive functioning, and all epilepsy-specific measures. These results support the validity of these scales as measures of the HRQOL impact of disease severity, defined as time since last seizure. A comparison of the F-statistics in Table 3 also supports the hypothesis that the Current Health scale (0.42) was more sensitive to the impact of time since last seizure than was the more general SF-36 General Health scale (0.29).

The epilepsy-specific scales were among the 10 best scales as defined by their relative validity in discriminating among patients differing in time since last seizure, although the most valid scale for this purpose was the SF-36 Role Functioning (Physical) scale. The scales with highest validities relative to Role Physical were the epilepsy-specific Impact (0.84), Experience (0.72), and Agitation (0.80) scales. Exceptions included the Seizure Severity scales (which would not be hypothesized to differentiate patients differing in seizure recency). Tests of incremental validity, however, indicated that only the Agitation and Impact scales added significantly to the amount of variation in disease severity (defined as time since last seizure) explained by the SF-36. (Data available upon request.)

Comparisons among patients differing in symptoms. As hypothesized, symptoms had a negative and significant impact on HRQOL. Median correlations between

Table 2. Results of item scaling tests and reliability estimates for HRQOL scales

HRQOL Scales	K[1]	Median item–scale correlations		Scaling test	Scale reliability	
		Item-internal consistency[2]	Item-discriminant validity[3]	Scaling success rate (%)[4]	Internal consistency[5]	Test–retest[6]
General HRQOL—SF-36 scales						
Physical functioning	10	0.73	0.28	93.8	0.92	0.77
Role physical	4	0.69	0.44	100.0	0.87	0.72
Bodily pain	2	0.70	0.43	100.0	0.79	0.74
General health perceptions	5	0.67	0.45	100.0	0.85	0.85
Vitality	4	0.72	0.47	100.0	0.87	0.71
Social functioning	2	0.64	0.55	100.0	0.78	0.71
Role emotional	3	0.77	0.50	100.0	0.87	0.65
Mental Health (MHI-5)	5	0.61	0.42	97.5	0.81	0.80
Augmented role physical	5	0.70	0.44	100.0	0.85	0.75
Augmented role emotional	4	0.72	0.50	100.0	0.85	0.66
General HRQOL—additional scales						
Anxiety	4	0.63	0.42	97.7	0.82	0.81
Depression	4	0.58	0.48	93.2	0.80	0.80
Behavioural/emotional control	5	0.54	0.36	85.5	0.76	0.78
Positive well-being	6	0.61	0.40	97.0	0.82	0.78
Emotional ties	4	0.52	0.30	93.8	0.74	0.76
Current health	4	0.69	0.45	86.4	0.83	0.82
MOS cognition	6	0.76	0.36	100.0	0.92	0.69
Confusion	2	0.68	0.41	93.8	0.81	0.62
Thinking	2	0.53	0.51	65.6	0.70	0.74
Concentration	2	0.68	0.46	93.8	0.81	0.63
Attention	2	0.34	0.39	31.3	0.51	0.55
Psychomotor function	3	0.65	0.40	93.8	0.80	0.66
Epilepsy–specific HRQOL scales						
Mastery	6	0.42	0.28	83.3	0.70	0.74
Impact[7]	8	0.69	0.44	100.0	0.89	0.82
Experience	13	0.56	0.34	94.7	0.86	0.84
Worry	9	0.55	0.29	98.4	0.83	0.74
Agitation	2	0.66	0.51	92.3	0.79	0.73
Distress	2	0.71	0.49	92.3	0.82	0.70
Seizure severity scales						
Percept[8]	8	0.20	0.11	68.8	0.43	0.76
Ictal[9]	12	0.57	0.22	100.0	0.88	0.88

[1] K = number of items per scale/subscale.
[2] Median correlation between items and hypothesized scales.
[3] Median correlation between items and other scales.
[4] Number of hypothesized higher/total number of correlations.
[5] Internal consistency reliability (Cronbach's Alpha)
[6] Intraclass correlation coefficient. Note: intraclass and product moment correlation coefficients were almost identical.
[7] Not linearly transformed, higher score indicates greater impact of epilepsy (range = 8–32).
[8] Not linearly transformed, higher score indicates more severe seizures (range = 7–32).
[9] Not linearly transformed, higher score indicates more severe seizures (range = 10–48).

symptom and HRQOL scale scores ranged from −0.21 to −0.42 (see Table 4). The only symptom which did not show at least a moderate association with at least one HRQOL domain is weight loss. Because only four subjects reported problematic weight loss, the study had limited potential to detect such an association.

The median correlations across symptoms reveal that the Mental Health (for MHI-5, $r = -0.35$; for MHI-17, $r = -0.42$), Vitality ($r = -0.35$), and the Psychomotor Functioning ($r = -0.40$) scales were the HRQOL scales most strongly related to symptoms. Most highly related to mental health were

Table 3. Mean HRQOL scale scores for seizure-free patients and deviation scores for patients with seizures in the last 6 months

	Seizure free > 6 months (n = 41)		Seizure free > 3 months ≤ 6 months (n = 8)		Seizure free > 1 week ≤ 3 months (n = 33)		≥ 1 seizure past week (n = 52)			
HRQOL scales	$\bar{x}$	SE	$\bar{x}$ Dev.	SE	$\bar{x}$ Dev.	SE	$\bar{x}$ Dev.	SE	F	RV[1]
General HRQOL – SF-36 scales										
Physical functioning	93.4	3.6	−3.4	8.8	−14.9**	5.4	−16.0**	4.9	4.3**	0.33
Role physical	90.6	5.5	−12.5	13.5	−23.2**	8.2	−46.9***	7.6	12.9***	1.00
Bodily plan	84.1	4.0	2.9	9.9	−4.3	6.0	−13.2*	5.5	2.3	—
General health	71.9	3.8	−1.0	9.4	−5.7	5.8	−16.9**	5.3	3.7*	0.29
Vitality	66.0	3.4	−7.2	8.5	−7.7	5.2	−18.7***	4.8	5.2**	0.40
Social functioning	89.0	3.8	−15.6	9.5	−15.6**	5.8	−24.8***	5.3	7.4***	0.57
Role emotional	89.2	6.1	−22.5	15.0	−25.5**	9.1	−42.2***	8.5	8.3***	0.64
Mental health (MHI-5)	76.6	3.1	−11.6	7.6	−9.6*	4.6	−20.3***	4.3	7.6***	0.59
Augmented role physical	91.5	3.5	−11.5	10.1	−21.2**	6.2	−39.9***	5.5	11.3***	0.87
Augmented role emotional	91.2	3.6	−19.3	14.7	−23.8**	6.8	−36.0***	5.7	7.8***	0.60
General HRQOL—additional scales										
Mental health (MHI-17)	76.9	2.7	−12.3†	6.7	−8.3*	4.0	−14.1***	3.7	4.9**	0.38
Anxiety	75.9	3.4	−15.2†	8.3	−11.0*	5.0	−14.6**	4.7	3.7*	0.29
Depression	80.2	3.1	−15.2*	7.6	−10.5*	4.6	−12.3**	4.3	3.5*	0.27
Behavioural/emotional control	82.4	3.1	−8.9	7.7	−9.2*	4.7	−13.0**	4.3	3.1*	0.24
Positive well-being	66.9	3.1	−7.5	7.6	−3.3	4.6	−16.7***	4.9	5.0**	0.39
Emotional ties	86.3	3.0	−8.2	7.4	−7.6†	4.5	−10.1*	4.1	2.1	–
Current health	71.3	3.9	−4.6	9.7	−9.7†	5.9	−21.5***	5.4	5.4**	0.42
MOS cognition	83.5	3.6	−5.2	9.0	−12.9*	5.5	−8.9†	5.1	2.1	–
Confusion	91.7	3.2	−8.0	8.0	−14.4**	4.8	−8.3†	4.5	3.1*	0.24
Thinking	77.7	3.9	−6.8	9.7	−12.1*	5.8	−17.2**	5.4	3.5*	0.27
Concentration	68.6	4.2	0.2	10.2	−9.1	6.2	−14.3*	5.8	2.3	–
Attention	84.4	3.5	−10.6	8.5	−11.7*	5.2	−12.1*	4.8	2.7*	0.21
Psychomotor function	86.7	3.3	−13.3	8.2	−11.1*	5.0	−8.8*	4.6	2.2	–
Epilepsy–specific HRQOL scales										
Mastery	62.5	2.8	−2.8	6.9	−8.5*	4.2	−15.5***	3.9	5.6***	0.43
Impact[2]	11.0	0.9	3.3	2.1	4.3***	1.3	6.8***	1.2	10.8***	0.84
Experience	87.5	1.7	−13.8*	4.3	−14.8***	3.8	−20.2***	3.1	9.8***	0.72
Worry	89.6	2.5	−19.5*	4.5	−16.1***	3.7	−12.0**	2.7	6.4***	0.50
Agitation	88.0	3.9	−13.0	9.5	−24.4***	5.8	−27.8***	5.4	10.2***	0.80
Distress	87.5	3.9	−6.3	9.5	−22.9***	5.8	−18.2***	5.4	6.4***	0.50

[1] RV = Relative validity (see text) – Note that RV not estimated if F-statistic was not significant.
[2] Not linearly transformed, higher score indicates greater impact of epilepsy (range = 8–32).
*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; † $p < 0.1$.

Table 4. Correlations between symptoms and HRQOL scales and estimates of relative validity

HRQOL Scales	Symptom Correlation[1] Median	Range	Most highly correlated symptom	F[2]	RV[3]
General HRQOL – SF-36 scales					
Physical functioning	−0.21	0.06 to − 0.44	unsteadiness	6.8***	0.41
Role physical	−0.21	0.07 to − 0.48	unsteadiness	5.2***	0.32
Bodily pain	−0.28	0.01 to − 0.49	vomiting	8.8***	0.53
General health perceptions	−0.29	−0.06 to − 0.45	stomach upset	4.3***	0.26
Vitality	−0.35	0.06 to − 0.48	unsteadiness	8.0***	0.48
Social functioning	−0.32	0.02 to − 0.53	unsteadiness	16.5***	1.00
Role emotional	−0.30	−0.01 to − 0.54	unsteadiness	14.1***	0.85
Mental health (MHI-5)	−0.35	−0.08 to − 0.50	react slowly	15.0***	0.91
Augmented role physical	−0.21	0.02 to − 0.47	unsteadiness	4.9***	0.30
Augmented role emotional	−0.30	−0.07 to − 0.61	unsteadiness	12.5***	0.76
General HRQOL–additional scales					
Mental health (MHI-17)	−0.42	−0.12 to − 0.49	invol. movements	15.3***	0.93
Anxiety	−0.35	−0.08 to − 0.43	invol. movements	9.0***	0.55
Depression	−0.33	−0.02 to − 0.45	invol. movements	10.8***	0.65
Behavioural/emotional control	−0.32	−0.09 to − 0.47	react slowly	11.2***	0.68
Positive well-being	−0.28	0.01 to − 0.42	unsteadiness	9.7***	0.59
Emotional ties	−0.22	−0.11 to − 0.32	invol. movements	5.7***	0.35
Current health	−0.31	−0.03 to − 0.44	stomach upset	3.9***	0.24
MOS cognition	−0.30	−0.14 to − 0.61	unsteadiness	16.5***	1.00
Confusion	−0.27	−0.12 to − 0.47	react slowly	7.6***	0.46
Thinking	−0.30	−0.13 to − 0.55	unsteadiness	11.7***	0.71
Concentration	−0.26	−0.10 to − 0.41	invol. movements	10.1***	0.61
Attention	−0.32	−0.17 to − 0.53	react slowly	15.5***	0.94
Psychomotor function	−0.40	−0.13 to − 0.58	unsteadiness	13.3***	0.81
Epilepsy–specific HRQOL scales					
Mastery	−0.22	0.05 to − 0.48	unsteadiness	6.1***	0.37
Impact	−0.24	0.47 to − 0.01	invol. movements	13.1***	0.79
Experience	−0.26	0.42 to − 0.07	unsteadiness	5.1***	0.31
Worry	−0.27	0.37 to − 0.04	slurred speech	5.1***	0.31
Agitation	−0.27	−0.08 to − 0.52	invol. movements	12.2***	0.74
Distress	−0.25	−0.05 to − 0.40	unsteadiness	6.6***	0.40

[1] Correlation between the scale and any of the 16 symptoms.
[2] Sensitivity of the scale to the simultaneous impact of all 16 symptoms.
[3] Estimates of Relative Validity (see text).
*** $p < 0.001$

unsteadiness, involuntary movements, trouble sleeping, reacting slowly, and headaches. Neurological symptoms were also most highly related to HRQOL problems in the area of psychomotor functioning.

The majority of scales with highest relative validity to differentiate significantly between groups of patients differing in the impact of 13 symptoms were generic HRQOL scales. The most valid scales for this purpose were the SF-36 Social Functioning scale and the MOS Cognition scale. Relative validity coefficients were also high for the SF-36 Role-Emotional (0.85), and Mental Health (0.91) scales. Of the additional generic scales, relative validities were high for the generic Attention (0.94) and Psychomotor Functioning scales (0.81).

Tables 5 and 6 present analyses of the impact of the five most prevalent symptoms on the HRQOL of seizure-free patients and patients with seizures during the past week, respectively. These data demonstrate that, despite the small sample sizes, controlling for seizure recency, patients who reported unsteadiness, hand tremors, slowed reactions, headaches, or nausea had lower HRQOL scores on most of the scales than those who did not report those symptoms. In many instances, despite small sample sizes, those experiencing symptoms had significantly lower HRQOL than those not experiencing symptoms.

Open-ended questions. Content analysis of patients' responses to two open-ended questions revealed very

Table 5. Comparison of mean HRQOL scale scores for patients with and without selected symptoms[1]: seizure-free patients[2]

HRQOL scales	SYMPTOMS									
	Unsteadiness		Hand tremors		React slowly		Headaches		Upset stomach	
	No (n = 32)	Yes (n = 9)	No (n = 31)	Yes (n = 10)	No (n = 29)	Yes (n = 12)	No (n = 21)	Yes (n = 20)	No (n = 31)	Yes (n = 10)
General HRQOL – SF-36 scales										
Physical functioning	95.3	85.6	95.3	87.5	92.9	94.5	90.8	96.0	93.9	91.7
Role physical	93.8	78.1	95.0	77.5*	89.7	93.2	92.5	88.7	90.3	91.7
Bodily pain	86.7	74.9	86.9	75.4	85.5	80.7	86.2	81.9	86.6	76.4
General health	74.3	62.1	75.1	62.4	74.7	64.5	75.2	68.6	76.4	56.6*
Vitality	68.7	56.1	70.0	53.5*	69.0	58.7	70.2	61.5	70.2	53.0*
Social functioning	92.6	76.4	92.7	77.5**	91.4	83.3	90.5	87.5	92.3	78.8*
Role emotional	91.7	79.2	94.4	73.3	90.8	84.8	85.0	93.3	86.0	100.0
Mental health (MHI-5)	77.0	75.1	78.7	70.0	79.0	70.7	81.1	71.8†	79.0	69.2
Augmented role physical	95.0	77.5*	96.0	78.0*	91.0	92.7	93.0	90.0	90.9	93.3
Augmented role emotional	93.7	81.2	95.8	77.5*	92.2	88.6	87.5	95.0	88.7	100.0
General HRQOL – additional scales										
Mental health (MHI-17)	78.2	72.0	78.6	71.7	79.0	71.8	80.9	72.7	79.4	69.1
Anxiety	78.1	67.8	78.2	68.5	78.5	69.6	81.9	69.6*	77.9	69.5
Depression	82.0	73.9	81.3	77.0	82.1	75.8	85.5	74.8*	82.3	74.0
Behavioural/emotional control	83.6	78.1	84.9	74.7	87.3	70.6**	86.1	78.6	85.8	72.0
Positive well-being	68.1	62.8	67.9	64.0	67.4	65.8	70.2	63.5	70.0	57.5
Emotional ties	85.5	89.4	89.2	77.5	89.1	79.6	88.6	84.0	88.9	78.5
Current health	73.6	63.1	75.2	59.3	75.1	62.1	75.1	67.3	75.3	59.0
MOS cognition	88.5	65.6*	87.7	70.3	90.6	66.4**	85.7	81.2	88.5	68.0**
Confusion	95.0	80.0	91.9	91.0	95.2	83.3	91.0	92.5	93.5	86.0
Thinking	83.7	56.7**	82.7	62.5	85.4	59.2***	80.6	74.8	82.4	63.3*
Concentration	72.6	50.0*	71.4	57.8	77.7	45.4***	75.6	61.2	72.0	58.7
Attention	89.4	66.7**	90.0	67.0*	90.0	70.8**	87.1	81.5	88.7	71.0*
Psychomotor function	92.9	64.4**	90.8	74.0	94.0	68.9***	86.3	87.0	91.4	72.0**
Epilepsy–specific HRQOL scales										
Mastery	66.2	49.7**	66.8	49.4**	62.6	62.3	63.9	61.1	64.3	56.9
Impact[3]	10.1	14.1*	10.9	11.4	10.2	13.0	10.1	11.9	10.6	12.3
Experience	88.9	82.6	87.4	87.9	89.1	83.8	88.6	86.4	87.8	86.7
Worry	92.0	81.1	93.0	79.1	92.8	81.9	91.2	87.9	92.0	82.2
Agitation	89.4	82.5	89.4	83.3	91.7	78.2*	94.3	81.1*	90.3	81.0
Distress	87.5	87.5	88.1	85.6	87.9	86.4	95.2	79.0**	88.3	85.0

[1] Symptoms within 4 weeks prior to HRQOL assessment response choices 1) No; 2) Yes, but not a problem; 3) Yes, somewhat of a problem; 4) Yes, a big problem.
[2] Epilepsy patients without seizures for at least 6 months prior to HRQOL assessment.
[3] Not linearly transformed, higher score indicates greater impact of epilepsy (range = 32).
*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; † $p < 0.10$.

Table 6. Comparison of mean HRQOL scale scores for patients with and without selected symptoms[1]: patients having seizures in the last week[2]

HRQOL scales	SYMPTOMS									
	Unsteadiness		Hand tremors		React slowly		Headaches		Upset stomach	
	No (n = 15)	Yes (n = 24)	No (n = 22)	Yes (n = 20)	No (n = 20)	Yes (n = 22)	No (n = 15)	Yes (n = 27)	No (n = 29)	Yes (n = 15)
General HRQOL – SF-36 scales										
Physical functioning	83.3	74.0	82.5	72.2	90.0	64.3***	84.0	73.2	83.1	62.9*
Role physical	61.7	34.3*	46.6	41.3	61.8	23.9**	56.7	35.7	52.6	20.8*
Bodily pain	78.0	68.1	77.2	64.0	76.8	61.9	80.5	64.8	77.2	53.3*
General health	69.5	47.3**	61.5	47.6	68.6	42.2**	71.4	45.0**	64.4	31.1***
Vitality	62.7	39.1***	54.5	38.8*	58.2	37.5**	60.0	39.6**	53.8	30.8***
Social functioning	79.2	56.9**	67.0	60.0	75.7	52.8**	81.7	54.0***	71.1	42.7**
Role emotional	68.9	34.6**	57.6	35.0	57.9	30.3*	66.7	34.5*	52.9	27.8
Mental health (MHI-5)	67.2	50.2**	64.2	46.0**	63.2	46.6**	68.0	48.9**	59.4	43.0**
Augmented role physical	65.3	44.3*	53.6	49.5	66.3	40.0*	61.3	46.4	57.9	38.6*
Augmented role emotional	75.0	44.6**	62.5	47.6	64.5	47.9*	73.3	45.5*	60.3	44.6
General HRQOL – additional scales										
Mental health (MHI-17)	70.5	58.4*	69.6	54.0**	66.5	56.1*	71.6	57.1**	65.3	51.8*
Anxiety	69.3	57.4*	71.1	49.0***	65.0	55.0†	70.7	55.4*	64.0	50.8†
Depression	75.7	64.0	74.5	59.2*	72.4	61.6	78.7	61.4**	70.8	55.8*
Behavioural/emotional control	76.8	64.9*	75.1	62.0*	72.5	64.2	74.5	66.1	70.0	64.0
Positive well-being	60.7	44.8**	56.8	43.6*	57.6	44.6**	60.7	44.8**	55.5	40.0***
Emotional ties	80.0	74.1	79.1	72.8	77.1	74.2	77.3	75.1	77.2	73.3
Current health	64.8	42.1**	57.0	41.8	64.1	35.4***	67.8	39.0***	59.4	24.4***
MOS cognition	80.2	70.6	81.1	65.7**	84.0	64.9***	81.1	70.2	78.5	63.1*
Confusion	89.3	80.0	87.3	78.0	86.8	79.1	88.7	80.0	87.9	70.8**
Thinking	65.8	55.9	66.4	52.0*	66.3	52.2*	68.2	55.0	64.0	48.3*
Concentration	55.8	51.8	62.5	43.7**	55.9	48.3	56.7	51.3	53.0	54.2
Attention	82.0	66.3**	79.5	63.5**	84.2	61.8***	79.3	67.5*	78.6	57.5***
Psychomotor function	86.2	72.6**	80.9	73.7	82.8	73.0*	84.9	73.3*	80.9	70.0*
Epilepsy–specific HRQOL scales										
Mastery	51.8	45.1	46.9	45.6	51.4	41.9	57.4	40.7**	47.1	43.0
Impact[3]	15.3	19.1*	17.3	18.8	14.7	20.6**	15.3	19.4*	17.4	19.6
Experience	81.1	60.1***	70.5	64.0	74.2	62.0*	79.0	61.2**	70.3	61.3
Worry	81.8	75.5	79.7	75.6	82.3	74.1	84.1	74.3	78.9	75.2
Agitation	72.7	54.4*	65.9	52.5	64.2	55.9	73.3	52.5**	64.8	45.0*
Distress	83.3	63.3**	71.8	65.0	73.2	63.2	84.7	60.4**	71.7	58.3

[1] Symptoms within 4 weeks prior to HRQOL assessment response choices 1) No; 2) yes, but not a problem; 3) Yes, somewhat of a problem; 4) Yes, a big problem.
[2] Epilepsy patients with ≥ 1 seizure in the week prior to HRQOL assessment.
[3] Not linearly transformed, higher score indicates greater impact of epilepsy (range = 8–32).
*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; † $p < 0.10$.

little new information. However, four patients (3% of the sample) indicated being bothered by prejudice from others. The study questionnaire asked patients about the impact of epilepsy and its treatment on their relationships, but did not include an assessment of perceived or enacted stigma.

Discussion

We have studied 31 generic and epilepsy-specific scales to evaluate their usefulness for assessing the impact of epilepsy and AED therapy on HRQOL. We conclude that: (a) a 171-item questionnaire can be completed in about 40 min on average by outpatients with epilepsy with good data quality; (b) both generic and epilepsy-specific measures satisfy psychometric criteria, with few exceptions; (c) generic and epilepsy-specific HRQOL scales were valid in relation to clinical criteria of disease severity and symptoms; and (d) disease severity and symptoms affect different HRQOL concepts. These conclusions and their implications are discussed below.

Data quality and feasibility

Although respondents completed a 171-item questionnaire with good data quality, there are good reasons to shorten the questionnaire. A shorter questionnaire would be more acceptable to respondents, would cost less to administer and process, and would provide more latitude for inclusion of other modules needed for a particular study. Results from data quality, psychometric, and validity analyses should be taken into account to identify the trade offs associated with the inclusion or exclusion of measures.

Psychometric tests of scaling assumptions

Consistent with previous studies of people with epilepsy[14] and other medical conditions,[52] psychometric tests indicate that all SF-36 scales met the scaling assumptions underlying their scoring. Test–retest reliability results of this study confirm previous estimates for the SF-36 based on internal consistency methods and show that they are generalizable to people with epilepsy. The components analysis confirmed the physical and mental health measurement model underlying the SF-36 and demonstrates its generalizability to the epilepsy population. All but two of the additional generic HRQOL scales and subscales met scaling assumptions and reliability requirements. Modifications of the Internal Locus of Control scale achieved the goal of focusing on mastery in relationship to epilepsy and did not compromise the psychometric properties of the scale. The internal consistency reliability of the epilepsy-specific Mastery scale compares favourably with the original version of the scale.[33] Consistent with previous analyses,[35,36] we found satisfactory reliability of the Epilepsy Impact scale and the Ictal subscale of the Seizure Severity scale, however not the Percept subscale. The newly created epilepsy-specific measures (Experience and Worry scales) demonstrated high reliability.

Tests of validity using clinical criteria

Proof that scaling assumptions are satisfied is necessary but insufficient for establishing the usefulness of a measure. The usefulness of the measure in capturing the impact of the disease and treatment under study must also be demonstrated. A measure of HRQOL for people with epilepsy must be shown to be empirically valid in tests using clinical criteria that define the severity of the disease and the impact of treatment. In this study, validity of the HRQOL measures was assessed with regard to two clinical criteria: time since last seizure and self-reported symptom status. Time since last seizure has previously been linked to patients' functioning and well-being.[32,56] Most of the study measures discriminated between patients who were seizure-free for > 6 months and those who recently had seizures, as well as between patients differing in self-reported symptom status, regardless of seizure control. However, the scales differed widely in their precision in detecting clinically defined differences. For example, the SF-36 Role Physical scale differentiated best among people differing in seizure control, while the SF-36 Role Emotional scale performed at a 64% level, relative to the best measure. The sensitivity of the role scales to the impact of seizures is particularly noteworthy because these scales are among the least precise scales studied due to their dichotomous response format. Because role disability constitutes an important HRQOL domain for people with epilepsy, priority should be given to improving measures of this important outcome. Adding two items to the two SF-36 Role Functioning scales in this study resulted in an improvement of the scale score distribution. This improvement will be even more important for studies of epilepsy patients who are more disabled than those participating in this study. Substantial gains with a multicategorical

response format for the SF-36 Role Functioning scales have been demonstrated.[57]

The Current Health scale was found to be more sensitive to the impact of seizures than the SF-36 General Health scale. The addition of only one item to the SF-36 allows the scoring of both the Current Health scale and the original General Health scale. This addition is recommended for future studies of epilepsy patients.

Comparisons of the 5-item and 17-item mental health measures indicated that for eight of 14 symptoms (blurred vision, slurred speech, involuntary movements, hand tremors, weight gain, stomach upset, vomiting and skin problems) the longer measure improved empirical validity up to 88% for purposes of detecting HRQOL differences associated with self-reported symptoms. Thus, a smaller sample size or more powerful statistical tests are possible when the longer measure is used. In addition, with the MHI-17, four different mental health scales can be constructed to identify the impact of seizures and symptoms in terms of anxiety, depression, loss of behavioural/emotional control and positive affect. Therefore, the longer form is recommended for future studies of people with epilepsy. The one MHI-18 emotional ties item (which had been split off in the analyses to compute the separate Emotional Ties scale) should be included to create the MHI-18 to maintain comparability with previous research.[26]

The HRQOL measures most affected differed substantially across the two clinical criteria studied. The six epilepsy-specific HRQOL scales (Mastery, Impact, Experience, Worry, Agitation, and Distress) tested in this study were among the 10 best scales in discriminating between patients who differed in severity as defined by time since last seizure. By contrast, generic HRQOL scales were best at differentiating between patients who did and did not report problems with the 16 symptoms studied. These findings underscore the importance of using multiple clinical criteria for measurement validation. The criteria chosen should fit the intended purpose of measurement. Measures that respond most to differences in disease severity, as defined here, are not likely to be most responsive to the suspected side effects of treatment. The converse also appears to be true.

The absence of a patient-based measure of drug toxicity has been identified as a clear deficiency of an HRQOL model for people with epilepsy,[16] and research to develop such a measure has begun.[58] In this study, we attempted to validate HRQOL measures with regard to patient-reported symptoms frequently associated with AED therapy.

Note that although there may be conceptual overlap between specific symptoms (such as hand tremors) and HRQOL concepts (such as anxiety), we tried to minimize this by excluding from the symptom checklist those symptoms that were assessed by individual HRQOL scales. It is our opinion that information about the frequency and severity of specific symptoms is necessary to understand individuals' evaluations of their health in general.[59] That is, patient-based reports of specific symptoms can be used to interpret a more global health effect such as reflected in the general HRQOL scales. This topic, which is beyond the scope of this paper, is discussed elsewhere.[59]

Although mostly physical, patient-reported symptoms were generally shown in this study to be most correlated with measures of psychological distress and well-being. This finding further suggests that thorough assessment of mental health may be required in order to detect HRQOL differences between AED that may differ in their side-effect profiles. However, these conclusions assume that different AED produce differences in the symptom impact studied.

Items and scales recommended for further evaluation

Some of the scales (Attention, Concentration, and Percept Scales) fielded in this study did not satisfy psychometric or validity criteria and may not warrant widespread use in this population. The Percept subscale of the Liverpool Seizure Severity scale had low internal consistency reliability and failed tests of item discriminant validity due to heterogeneity of content. Although psychometrically sound, the Emotional Ties scale failed to differentiate between groups of patients differing in seizure control or symptom status in this study, indicating that it could be excluded without a substantial loss of information. Although tests of validity did support the inclusion of the epilepsy experience and worry questions, there seemed to be redundancy in the content of these questions. When eliminating questions for which the majority of patients responded that the issue was not relevant and combining those that referred to the same dimension, the 22 epilepsy experience and worry questions could be reduced to a set of seven experience and four worry questions to test in future studies. When shortening of the HRQOL questionnaire is desired, these considerations could justify the deletion of 35 items from the battery tested in this study, pending further evaluations.

Limitations of the study

When interpreting the findings of this study, it is important to note that the study design does not allow any conclusions regarding the causes of the HRQOL differences observed. We attempted to solve the problem of confounding of HRQOL scale scores by the impact of possible AED-related seizures and symptoms by analysing subgroups of patients separately. However, only in a prospective study can the causes of HRQOL problems be investigated. Furthermore, due to the multicolinearity observed among the patient-reported symptoms we cannot be certain that any HRQOL impact observed is due to any one symptom.

Our study does not provide information on the sensitivity of the measures tested to change in severity or symptoms over time. To ascertain that seizure control and side effects due to AED affect the HRQOL of people with epilepsy, and to explore the sensitivity of these measures to changes over time, controlled, longitudinal trials are needed. Based on the cross-sectional data gathered in this study, we would hypothesize that SF-36 scales, as well as other scales tested, could detect HRQOL changes related to changes over time in seizure control and/or symptom impact. Measures that do best in discriminating the effects of differences at a point in time should also be most responsive to the impact of those changes over time.[42,47,48] However, this principal has been questioned,[60] and in a previous study no significant differences were found for any of the scales in the Nottingham Health Profile between epilepsy patients on placebo or study medication who differed in seizure frequency.[33] It is not clear whether this reflects the efficacy of the treatment studied or measurement validity. Other studies of the SF-36 suggest that the scales that do best in cross-sectional analyses also do best in longitudinal analyses.[42,61,62]

Lastly, several domains of HRQOL, such as sleep, stigma, and sexual function, have not been assessed with multi-item scales in this study. Because of their importance to patients,[8] they warrant inclusion in future studies.

Several approaches to comprehensively measure the HRQOL of people with epilepsy have been undertaken during the past 2 years.[13–16] This study is similar to those in that it evaluates generic measures of HRQOL as well as epilepsy-specific measures. Close similarities exist between the questionnaire studied here and the recently developed Quality of Life in Epilepsy Inventory (QOLIE-89).[15] Appendix 2 provides a comparison of the content and the number of items per scale of the two questionnaires. Differences are most noticeable in the assessment of psychological distress and well-being, cognitive functioning, and epilepsy-specific areas of life. Our study adds to the currently available data by evaluating the validity of generic and specific HRQOL measures not only in relationship to the impact of seizures, but also in relationship to potential adverse effects of AED therapy. It also provides an initial estimate of the information gained from epilepsy-specific measures over and above that provided by generic scales. Further use of these and other HRQOL scales is needed to identify which will provide most relevant information for which purpose. The successful treatment of patients with epilepsy consists of establishing the best possible balance between seizure control and adverse treatment effects, and both seizures and adverse drug effects can influence quality of life. It is therefore important that HRQOL measures detect both the impact of seizures and AED-related adverse effects.

We have used statistical methods to estimate differences in empirical validity in order to better quantify the contribution from each HRQOL scale relative to that of other scales. We hope that these methods will prove beneficial in assessing the usefulness of additional scales in studies of the HRQOL of people with epilepsy. The scales evaluated and recommended here appear to be practical, reliable, and valid for use in measuring outcomes in clinical trials of AED. Hopefully they will contribute to the improvement of HRQOL among patients with epilepsy.

Acknowledgements

We thank the patients for taking time to complete the questionnaire, and the clinicians at the three study sites for their collaboration. We thank Schering-Plough Corporation for funding the fielding of this study and the preparation of this manuscript. The preparation of this manuscript was also supported by the International Quality of Life Assessment (IQOLA) Project funded by Schering-Plough Corporation and Glaxo Research Institute. We acknowledge Dr John Brazier and his colleagues who within the IQOLA Project adapted and validated the SF-36 for use in the United Kingdom. We acknowledge the Wellcome Foundation for granting permission for us to use the Seizure Severity scale in this study. We thank Drs Barbara Vickrey, Ron Hays and two anonymous reviewers for their helpful comments on earlier drafts of this manuscript.

References

1. Livingston S. Psychosocial aspects of epilepsy. In: British Epilepsy Association, ed. *Perspectives on Epilepsy 80/81.* Berkshire: British Epilepsy Association 1981: 17–27.
2. The Roper Organization. *Living with Epilepsy: Report of a Roper Poll of Patients on Quality of Life.* New York, 1992.
3. Hermann BP. Quality of life in epilepsy. *J Epilepsy* 1992; **5:** 153–165.
4. Ellwood PM. Outcomes management: A technology of patient experience (Shattuck Lecture). *New Engl J Med 1988;* **318:** 1549–1556.
5. Slevin ML, Plant H, Lynch D, *et al.* Who should measure quality of life, the doctor or the patient? *Br J Cancer* 1988; **57:** 109–112.
6. Geigle R, Jones SB. Outcomes measurement: A report from the front. *Inquiry* 1990; **27:** 7–13.
7. Wilkin D, Hallam L, Doggett MA. *Measures of Need and Outcome for Primary Health Care.* Oxford: Oxford University Press, 1992.
8. Stewart AL, Ware JE, eds. *Measuring Functioning and Well-Being. The Medical Outcomes Study Approach.* Durham, NC: Duke University Press, 1992.
9. Temkin NR, Dikmen S, Machamer J, *et al.* General versus disease-specific measures: Further work on the Sickness Impact Profile for Head Injury. *Med Care* 1992; **27:** S44–S53.
10. Kantz ME, Harris WJ, Levitsky K, *et al.* Methods for assessing condition-specific and generic functional status outcomes after total knee replacement. *Med Care* 1992; **30:** MS240–MS252.
11. Patrick DL, Erickson P. *Health Status and Health Policy: Allocating Resources to Health Care.* New York: Oxford University Press, 1993.
12. Patrick DL, Deyo RA. Generic and disease-specific measures in assessing health status and quality of life. *Med Care* 1989; **27:** S217–S232.
13. Jacoby A, Johnson A, Chadwick DW, *et al.* Psychosocial outcomes of antiepileptic drug discontinuation. *Epilepsia* 1992; **33:** 1123–1131.
14. Vickrey BG, Hays RD, Graber J, *et al.* A health-related quality of life instrument for patients evaluated for epilepsy surgery. *Med Care* 1992; **30:** 299–319.
15. Devinsky O, Hays R, Vickrey B, *et al.* Development of a quality of life instrument. *Neurology,* 1994; **44:** A141.
16. Baker GA, Smith DF, Dewey M, *et al.* The initial development of a health-related quality of life model as an outcome measure in epilepsy. *Epilepsy Res* 1993; **16:** 65–81.
17. Ware JE, Sherbourne CD. The MOS 36-Item Short-Form Health Survey (SF-36): I. Conceptual framework and item selection. *Med Care* 1992; **30:** 473–483.
18. Ware JE, Snow KK, Kosinski M, *et al. SF-36 Health Survey: Manual and Interpretation Guide.* Boston: New England Medical Center, The Health Institute, 1993.
19. Jenkinson C, Wright L, Coulter A. *Quality of Life Measurement in Health Care: A Review of Measures, and Population Norms for the UK SF-36.* Oxford: Horgan Print Partnership, 1993.
20. Brazier JE, Harper R, Jones NMB, *et al.* Validating the SF-36 Health Survey Questionnaire: New outcome measure for primary care. *Br Med* 1992; **305:** 160–164.
21. Garratt AM, Ruta DA, Abdalla MI, *et al.* The SF-36 Health Profile: An outcome measure suitable for routine use within the NHS? *Br Med J* 1993; **306:** 1440–1444.
22. Jenkinson C, Coulter A, Wright L. The Short-Form 36 (SF-36) Health Survey Questionnaire: Normative data for adults of working age. *Br Med J* 1993; **306:** 1437–1440.
23. Davies AR, Ware JE. *Measuring Health Perceptions in the Health Insurance Experiment.* Santa Monica: RAND 1982; R-2711-HHS.
24. Stewart AL, Hays RD, Ware JE. Health perceptions, energy/fatigue, and health distress measures. In: Stewart AL, Ware JE, eds. *Measuring Functioning and Well-Being: The Medical Outcomes Study Approach.* Durham, NC: Duke University Press, 1992.
25. Stewart AL, Ware JE, Sherbourne CD, *et al.* Psychological distress/well-being and cognitive functioning measures. In: Stewart AL, Ware JE, eds. *Measuring Functioning and Well-Being: The Medical Outcomes Study Approach.* Durham, NC: Duke University Press, 1992.
26. Weinstein MC, Berwick DM, Goldman PA, *et al.* A comparison of three psychiatric screening tests using Receiver Operating Characteristic (ROC) analysis. *Med Care* 1989; **27:** 593–607.
27. Russell D, Peplau LA, Ferguson ML. Developing a measure for loneliness. *J Pers Assess* 1978; **42:** 290–294.
28. Dupuy HJ. The psychological general well-being (PGWB) index. In: Wenger NK, Mattson ME, Furberg CD, Elinson J (eds). *Assessment of Quality of Life in Clinical Trials of Cardiovascular Therapies.* New York: Le Jacq, 1984.
29. Bergner M, Bobbitt RA, Carter WB, *et al.* The Sickness Impact Profile: Development and final revision of a health status measure. *Med Care* 1981; **19:** 787–805.
30. Dohrenwend BP, Shrout PE, Egri G, *et al.* Nonspecific psychological distress and other dimensions of psychopathology: Measures for use in general population. *Arch Gen Psychiatry* 1980; **37:** 1229–1236.
31. Pearlin L, Schooler C. The structure of coping. *J Health Soc Behav* 1978; **19:** 2–21.
32. Jacoby A. Epilepsy and the quality of everyday life: Findings from a study of people with well-controlled epilepsy. *Soc Sci Med* 1992; **34:** 657–666.
33. Smith D, Baker G, Davies G, *et al.* Outcomes of add-on treatment with lamotrigine in partial epilepsy. *Epilepsia* 1993; **34:** 312–322.
34. Jacoby A, Baker B, Smith D, *et al.* Measuring the impact of epilepsy: The development of a novel scale. *Epilepsy Res* 1993; **16:** 83–88.
35. Ware JE, Johnston SA, Davies-Avery A, *et al. Conceptualization and Measurement of Health for Adults in the Health Insurance Study: Vol. III, Mental Health.* Santa Monica, CA: Rand 1979; R-1978/3-HEW.
36. Baker GA, Smith DF, Dewey M, *et al.* The development of a seizure severity scale as an outcome measure in epilepsy. *Epilepsy Res* 1991; **8:** 245–251.
37. Likert RA. A technique for the measurement of attitudes. *Arch Psychol* 1932; **140:** 5–55.
38. Commission on Classification and Terminology of the International League Against Epilepsy (1981). Proposal for revised clinical and electroencephalographic classification of epileptic seizures. *Epilepsia* 1981; **22:** 489–501.

39. Commission on Classification and Terminology of the International League Against Epilepsy. Proposal for revised classification of epilepsies and epileptic syndromes. *Epilepsia* 1989; **30:** 389–399.
40. Ghiselli EE, Campbell JP, Zedeck S. *Measurement Theory for the Behavioral Sciences.* San Francisco: Freeman, 1981.
41. McHorney CA, Ware JE, Raczek AE. The MOS 36-Item Short-Form Health Status Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Med Care* 1993; **31:** 247–263.
42. Ware JE, Kosinski M, Bayliss MS, *et al.* Comparison of methods for the scoring and statistical analysis of SF-36 Health Survey profiles and summary measures: Results from the Medical Outcomes Study. *Med Care* 1995 (in press).
43. McHorney CA, Kosinski M, Ware JE. Comparisons of the cost and quality of norms for the SF-36 Survey collected by mail versus telephone interview: Results from a national survey. *Med Care;* **32:** 551–567.
44. Helmstadter GC. *Principles of Psychological Measurement.* New York: Appleton-Century-Crofts, 1964.
45. Howard KI, Forehand GC. A method for correcting item-total correlations for the effect of relevant item inclusion. *Educ Psychol Meas* 1962; **22:** 731–735.
46. Cronbach LJ, Warrington WG. Time-limit tests: Estimating their reliability and degree of speeding. *Psychometrika* 1951; **16:** 167–188.
47. Nunnally JC. *Psychometric Theory.* New York: McGraw-Hill, 1978.
48. Stevens J. *Applied Multivariate Statistics for the Social Sciences.* Hillsdale: Erlbaum, 1992.
49. Kim J-O, Mueller CW. Factor Analysis. *Statistical Methods and Practical Issues.* London: Sage, 1978.
50. Comrey AL, Lee HB. *A First Course in Factor Analysis.* Hillsdale: Erlbaum, 1992.
51. Liang MH, Larson MG, Cullen KE, *et al.* Comparative measurement efficiency and sensitivity of five health status instruments for arthritis research. *Arthritis Rheum* 1985; **28:** 542–547.
52. McHorney CA, Ware JE, Lu JFR, *et al.* The MOS 36-Item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Med Care* 1994; **32:** 40–66.
53. Kerlinger FN. *Foundations of Behavioural Research.* New York: Holt, Rinehart and Winston, 1973.
54. Sechrest Lee. Incremental validity. In: Jackson DN, Messick S. *Problems in Human Assessment.* New York: McGraw-Hill Book Company, 1967.
55. Medical Outcomes Trust. *How to score the SF-36 Health Survey.* Boston: Medical Outcomes Trust, 1993.
56. Chaplin JE, Lasso RY, Shorvon SD, *et al.* National general practice study of epilepsy: The social and psychological effects of a recent diagnosis of epilepsy. *Br Med J* 1990; **304:** 1416–1418.
57. Kazis LE, Clark J, Miller DR, *et al.* A modified version of a short-form health survey for assessing health-related quality of life in veterans: The Veterans Health Study. In: AHSR & FHSR. *Health Services Research: Implications for Policy, Management, and Clinical practice.* Abstract book of the 11th Annual Meeting, 1994 June 12–14; San Diego, CA: 14–15.
58. Baker GA, Francis P, Middleton A, *et al.* Development of a patient-based symptom checklist to quantify adverse events in persons receiving anti-epileptic drugs. *Epilepsia* 1993; **34:** 18.
59. Ware JE. The status of health assessment 1994. *Annu Rev Public Health* 1994; **16**: 327–354.
60. Guyatt G, Feeny D, Patrick D. Issues in quality of life measurement in clinical trials. *Controlled Clin Trials* 1991; **12:** 81S–90S.
61. Katz JN, Larson MG, Phillips CB, *et al.* Comparative measurement sensitivity of short and longer health status instruments. *Med Care* 1992; **30:** 917–925.
62. Beaton DE, Bombardier C, Hogg-Johnson S. Choose your favorite tool: A comparison of the psychometric properties of five generic health status instruments in workers with soft tissue injuries. *Qual Life Res* 1994; **3:** 50.

Appendix 1: Scales, subscales, number of items and source of items of the HRQOL questionnaire for people with epilepsy

HRQOL scales and subscales by domain	No. of items[1]	Source
General HRQOL—SF-36 scales		
Physical functioning	**10**	SF-36[17] UK version
Role functioning—physical	**4**	SF-36[17] UK version plus:
Augmented role functioning—physical	**5**	1 new item
Bodily pain	**2**	SF-36[17] UK version
General health perceptions	**5**	SF-36[17] UK version plus:
Current health	4	1 item from MOS[23,24]
Health outlook	1	
Resistance to illness	1	
Vitality	**4**	SF-36[17] UK version
Social functioning	**2**	SF-36[17] UK version
Role functioning—emotional	**3**	SF-36[17] UK version plus:
Augmented role functioning—emotional	**4**	1 new item
Mental health (MHI-5)	**5**	SF-36[17] UK version
Change in health	1	SF-36[17] UK version
General HRQOL—additional scales		
Mental health (MHI-18)	**18**	MHI-18[25,26] (includes MHI-5) plus:
Anxiety	4	
Depression	4	
Behavioural/emotional control	5	1 item from MHI 38[25,26]
Positive well-being	5	
Emotional ties	4	3 items from UCLA Loneliness scale[27]
Overall quality of life	1	MOS[24,28]
Cognition	**6**	6 item MOS cognition scale[25] plus:
Confusion	2	1 item from SIP[27]
Thinking	2	1 item from PERI[30]
Concentration	2	2 items from PERI[30]
Attention	2	1 item from SIP[29]
Memory	1	
Reasoning	1	
Psychomotor functioning	3	2 items from SIP[29]
Epilepsy-specific HRQOL scales		
Mastery	**6**	Pearlin/Schooler Mastery scale[31]
Impact	**8**	Liverpool Impact scale[34]
Experience	**13**	New items
Worry	**9**	New items
Agitation	**2**	HIS[35]
Distress	**2**	MOS[24]
Seizure severity scales		
Ictal	**12**	Liverpool Seizure Severity scale[36]
Percept	**8**	Liverpool Seizure Severity scale[36]
Symptoms	**16**	New items
Open-ended questions	2	New items

[1] Bold numbers indicate the number of items in each overall scale. Non-bold numbers indicate the number of items in each subscale. Note: item numbers do not add up to 171.

Appendix 2: Comparison of multi-item scales (subscales) and number of items per scale in the study questionnaire and the QOLIE-89, including number of virtually identical items per scale

Study questionnaire scales	No. of items	QOLIE-89 scales	No. of items	No. identical items*
Physical functioning	10	Physical functioning	10	10
Role functioning—physical	5	Role Limitations—Physical	5	4
Pain	2	Pain	2	2
General health	6	Health perception	6	5
Current health	4			
Vitality	4	Energy/fatigue	4	4
Social functioning	2	Social functioning	11	1
		Social support	4	—
		Social isolation	2	—
Role functioning—emotional	4	Role limitations—emotional	5	3
Mental health (MHI-5)	5	Emotional well-being (MHI-5)	5	5
Mental health (MHI-18 including MHI-5)	18			
Anxiety	4			
Depression	4			
Behavioural/emotional control	5			
Positive well-being	5			
Emotional ties	4			
Overall quality of life	1	Overall quality of life	2	—
Cognitive functioning		Cognitive Functioning		
MOS cognition	6			
Confusion	2			
Thinking	2			
Concentration	2	Attention/Concentration	9	3
Attention	2			
Memory	1	Memory	6	—
Reasoning	1			
Psychomotor functioning	3			
		Language*	5	—
Epilepsy mastery	6			
Epilepsy impact	8			
Epilepsy experience	13			
Epilepsy worry	9	Seizure worry	5	—
Epilepsy agitation	2			
Epilepsy distress	2	Health discouragement	2	—
Seizure severity scales				
Ictal	12			
Percept	8			
		Sexual relationships**	1	—
		Medication effects	3	—
Symptoms	16			
Number of items	171	Number of items	89	
Number of multi-item scales	31	Number of multi-item scales	17	

* Number of virtually Identical Items.
** Sexual problems and problems with speech are assessed with one question each in the symptom checklist of the study questionnaire.