PATIENT-REPORTED OUTCOME MEASUREMENT GROUP, OXFORD

A STRUCTURED REVIEW OF PATIENT-REPORTED OUTCOME MEASURES (PROMs) FOR EPILEPSY

Report to the Department of Health, 2009

Patient-Reported

O U T C O M E

Measures

A STRUCTURED REVIEW OF PATIENT-REPORTED OUTCOME MEASURES FOR PEOPLE WITH EPILEPSY: AN UPDATE 2009

Nicola Davies
Elizabeth Gibbons
Ray Fitzpatrick
Anne Mackintosh

Patient-reported Outcome Measurement Group
Department of Public Health
University of Oxford

Copies of this report can be obtained from:

Elizabeth Gibbons
Senior Research Officer
Patient-reported Outcome Measurement Group

tel: +44 (0)1865 289402
e-mail: elizabeth.gibbons@dphpc.ox.ac.uk

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

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

CONTENTS

A STRUCTURED REVIEW OF PATIENT-REPORTED OUTCOME MEASURES FOR PEOPLE WITH EPILEPSY: AN UPDATE 2009

A STRUCTURED REVIEW OF PATIENT-REPORTED OUTCOME MEASURES FOR PEOPLE WITH EPILEPSY: AN UPDATE 2009

EXECUTIVE SUMMARY

Aims of the report

The aims of this report are to review the evidence of patient-reported outcome measures (PROMs) for people with epilepsy and to provide recommendations to the Department of Health of PROMs for epilepsy that could potentially be used on a large scale population basis, combining good measurement properties with the likelihood of modest burden to respondents in order not to jeopardise response rates. A literature review of relevant PROMs resulted in the identification of a short-list of both generic and epilepsy-specific instruments which were then presented to a multidisciplinary panel for discussion. The literature review of the evidence-base and the discussions of the multi-disciplinary panel underpin final recommendations to the Department of Health.

The Patient-reported Outcome Measurement Group previously submitted a report to the Department of Health of evidence of Patient-reported Outcome Measures (PROMs) for chronic conditions (Fitzpatrick et al., 2006)[1]. The report included a review of evidence regarding PROMs for epilepsy with some recommendations.

The methods of the review are described and the results of the search including sources and search terms used to identify specific published research. Details of this updated evidence are presented firstly for generic PROMs evaluated with people with epilepsy, followed by epilepsy-specific PROM results. The report concludes with discussion and recommendations.

Results and short-list of PROMs for people with epilepsy

The previous review reported evidence for the following generic PROMs:

a) EQ-5D
b) Health Utilities Index (HUI)
c) Nottingham Health Profile (NHP)
d) Q-TWIST
e) SF-36
f) SF-12
g) Sickness Impact Profile (SIP)

This update identified further evidence of performance for the EQ-5D, the HUI, and the SF-36 which were identified in the previous review. Two additional generic PROMs were identified, with reporting of some evidence of performance:

h) SF-6D
i) WHO Quality of Life assessment instrument - abbreviated version (WHOQOL-BREF)

[1] A structured review of PROMs for epilepsy, 2006 can be downloaded at http://phi.uhce.ox.ac.uk/

Eight condition-specific instruments were previously identified:

a) Epilepsy Surgery Inventory-55 (ESI-55)
b) Katz Adjustment Scale
c) Liverpool Quality of Life (LQOL) Battery and Seizure Severity Scale
d) Quality of Life in Epilepsy-89 (QOLIE-89)
e) Quality of Life in Epilepsy-31 (QOLIE-31)
f) Quality of Life in Epilepsy-10 (QOLIE-10)
g) Side-Effect and Life Satisfaction (SEALS) Inventory
h) Washington Psychosocial Seizure Inventory (WPSI)

This update identified further evidence of performance for the QOLIE-89, QOLIE-31 and the QOLIE-10.

Short-listed PROMs for discussion

Based on volume of evaluations and good measurement and operational characteristics, the following PROMs were presented to a multidisciplinary panel for further consideration:

1. EQ-5D
2. SF-36
3. ESI-55
4. QOLIE-89
5. QOLIE-31

Recommendations

On the basis of appraisal by the PROM Group, and ratings and comments from the panel, the **SF-36** is the preferred generic measure of health status and the **QOLIE-31** for the measurement of epilepsy-specific quality of life.

On the basis of appraisal of evidence by the PROM Group, and taking account of ratings and comments from the panel, the SF-36 and EQ-5D are considered suitable as generic measures in epilepsy. However, it was noted that the QOLIE epilepsy-specific instruments included items from the SF-36 and therefore duplication of items may increase completion burden to patients if used in conjunction with QOLIE. Given its brevity and the fact that it yields UK-derived preferences, the EQ-5D is therefore recommended for use in combination with a condition-specific PROM. . The QOLIE-31 is recommended as an epilepsy- specific instrument. These two measures used together will provide complementary evidence of health status of people with epilepsy in the context of potential population-level applications in the NHS. In making the final selection of PROMs considered suitable for piloting in the NHS, the DH will consider salient factors in addition to the evidence and multidisciplinary panel comments.

Chapter 1: INTRODUCTION & METHODS

Background

Patient-reported outcome measures (PROMs) offer enormous potential to improve the quality and results of health services. They provide validated evidence of health from the point of view of the user or patient. They may be used to assess levels of health and need in populations, and in users of services, and over time they can provide evidence of the outcomes of services for the purposes of audit, quality assurance, and comparative performance evaluation. They may also improve the quality of interactions between health professionals and individual service users.

Lord Darzi's Interim Report on the future of the NHS recommends that patient-reported outcome measures (PROMs) should have a greater role in the NHS (Darzi 2007). The new Standard NHS Contract for Acute Services, introduced in April 2008, includes a requirement to report from April 2009 on patient-reported outcome measures (PROMs) for patients undergoing Primary Unilateral Hip or Knee replacements, Groin Hernia surgery or Varicose Vein. Furthermore, Lord Darzi's report 'High Quality Care for All' (2008) outlines policy regarding payments to hospitals based on quality measures as well as volume. These measures include PROMs as a reflection of patients' experiences and views. Guidance has now been issued regarding the routine collection of PROMs for the selected elective procedures (Department of Health, 2008).

The Patient-reported Outcome Measurement Group previously submitted a report to the Department of Health of evidence of Patient-reported Outcome Measures (PROMs) for chronic conditions, carer impact, and patient perceptions of quality (Fitzpatrick et al., 2006). The report included a review of evidence regarding PROMs for epilepsy with some recommendations.

The aim of this report is two-fold: to provide an update of more recently published evidence for PROMs in epilepsy and to provide as clear recommendations as possible to Department of Health of PROMs that could be used on a potentially large scale population basis to assess health status of people with epilepsy to provide evidence relevant to determining the quality of services in the NHS. Recommended instruments would need to combine good measurement properties with the likelihood of low burden to respondents in order not to jeopardise responses rates. An additional consideration would be the availability of a PROM which yielded preference values derived from a UK source. As widely recommended, a strategy of combining a generic measure with a condition-specific measure was considered the most appropriate way of assessing complementary aspects of health status.

This current update review and recommendations draws on the existing evidence for each PROM up to 2006 but only provides fuller descriptive details of measurement and operational evidence which has emerged since 2006.

The full body of evidence was presented to a multidisciplinary panel for discussion. Details of their discussion and views are reported in Appendix F. The PROMs review group considered the combination of the full review of evidence and the multidisciplinary panel's views before reaching its own conclusions and recommendations (Chapter 4).

Structure of the report
The methods of the review update are described and the results of the search, including sources and search terms used to identify specific published research. Details of this updated evidence are presented firstly for generic PROMs evaluated with people with epilepsy, followed by epilepsy-specific PROM results. The report concludes with discussion and recommendations for short-listed PROMs.

Methods for the review update
Methods adopted were as described in previous reviews performed by the PROM group, Oxford. Comprehensive searches were conducted; articles retrieved were assessed for relevance and checked by another reviewer; and evidence of measurement performance and operational characteristics were abstracted for each PROM identified. Assessment and evaluation of the PROMs was performed independently by two reviewers adapting the London School of Hygiene appraisal criteria (Appendix B) outlined in their review (Smith et al., 2005). These criteria were modified for our reviews. The final short-listing of promising PROMs to formulate recommendations was based on these assessments and discussion between reviewers. The most promising PROMs were then presented to a multidisciplinary panel for final agreement.

Search terms and results: identification of articles
The methods for searching were conducted using three main sources.

The primary source of evidence was the bibliography compiled by the PROM group[2]. At the time of the review update, the PROM database comprised 16,054 records (up to December 2005) downloaded from several electronic databases using a complex search strategy (Appendix A.i). These records had been assessed as eligible for inclusion in the bibliography and assigned keywords including disease-group (e.g. neurological). The keywords were searched using 'neurol*', then the title and abstract of all records were searched using the terms 'epilep*' OR 'seizure.'

A further 14,296 records covering the period January 2006-July 2007 had been downloaded using a revised search strategy (Appendix A ii b) but not assessed or keyworded. The terms 'epilep*' OR 'neurol*' OR 'seizure' from title OR abstract were performed.

Supplementary searches included scanning the reference lists of key articles, checking instrument websites, where found, and drawing on other bibliographic resources. Hand searching of titles of key journal was conducted from October 2006 to October 2008. The following journals were selected:

- Epilepsia
- Epilepsy Research
- Health and Quality of Life Outcomes
- Medical Care
- Quality of Life Research
- The Patient: Patient-Centred Outcomes Research

[2] Available online at http://phi.uhce.ox.ac.uk

In addition, PubMed records for the past two years (i.e. September 2006-2008) were searched using the term 'epilepsy' and the names of the instruments identified in the previous review and this update.

All abstracts were reviewed. Articles assessed as meeting the review inclusion criteria were retrieved and reviewed in full. Of these, 26 articles were included in the review. The earlier review included 71 articles and combined with this update evidence, the total number of articles providing evidence is 97. Results are presented in Table 1.

Table 1: Number of articles identified by the literature review

Source	***Results of search***	***No. of articles considered eligible***	***Number of articles included in review***
PROM database: original search (up to December 2005) **Total number = 16,054**	103	25	2
Additional PROM database search (January 2006-July 2007) **Total number = 14,296**	384	8	7
Supplementary searching	-	-	17
TOTAL	-	-	**26**
2006 review			**71**
TOTAL			**97**

Chapter 2: Generic PROMs evaluated with people with epilepsy

The previous review reported evidence for the following PROMs:

a) EQ-5D
b) Health Utilities Index (HUI)
c) Nottingham Health Profile (NHP)
d) Q-TWIST
e) SF-36
f) SF-12
g) Sickness Impact Profile (SIP)

This update identified further evidence of performance for the EQ-5D and the SF-36. Two additional generic PROMs, the SF-6D and the WHOQOL-BREF, were identified with reporting of some evidence of performance.

Full details of the development, domains and scoring methods are detailed in Appendices C and D.

a) EQ-5D

Three UK studies describe the evaluation of the EQ-5D in the previous review, providing mainly negative evidence to support its use with this patient group. Three further studies, one from the UK, were identified in this update, offering limited but more favourable evidence on the construct validity and responsiveness of the EQ-5D.

In a UK study, the EQ-5D demonstrated discriminant validity according to anti-epileptic drug type and seizure frequency in patients receiving adjunctive therapy (Selai et al., 2005). The measure was also able to discriminate epilepsy patients experiencing sleep disturbance from those with no sleep disturbance (Xu et al., 2006). However, the measures' discriminant validity for seizure control has been found to be inferior to other generic, preference-based measures (Langfitt et al., 2006).

Some ceiling effects have been reported (34%) (Langfitt et al., 2006).

b) Health Utilities Index III (HUI-3)

Two studies describe the evaluation of the HUI-3 in the previous review, providing some evidence to support its measurement and operational properties. One further study was identified in this update, offering further evidence for the construct validity of the HUI-3.

The HUI-3 could discriminate between patients with and without seizures two years after baseline evaluation (Langfitt et al., 2006).

Ceiling effects have been reported to be below 10% (Langfitt et al., 2006).

c) Nottingham Health Profile (NHP)

Three studies were identified which evaluated the NHP in the previous review, reporting good internal consistency but offering unfavourable evidence on other measurement properties. No further studies evaluating the NHP were identified in this update.

d) Q-TWIST
One methodological study was identified which evaluated the Q-TWIST (quality-adjusted time without symptoms and toxicities) in the previous review, reporting no empirical evidence to support psychometric or operational performance. No further evidence was obtained for the Q-TWIST method in this update.

e) SF-36
Nine studies describe the evaluation of the SF-36 in the previous review with all measurement and operational criteria reported. One study was identified in this update, offering limited evidence of construct validity of the SF-36.

The SF-36 Total and eight scales failed to demonstrate discriminant validity in post-surgery seizure free epilepsy patients and age- and sex-matched norms, with the exception being for the Social Function subscale which was substantially less than predicted (Spencer et al., 2007).

f) SF-12
One study describes the evaluation of the SF-12 in the previous review with little measurement and operational criteria reported. No further studies of measurement properties were identified in this update.

g) Sickness Impact Profile (SIP)
One study was identified which evaluated the SIP in the previous review, reporting favourable, but limited, evidence to support the measurement properties of the instrument. No further evidence for the SIP was identified in this update.

h) SF-6D
No evidence was identified in the previous review for the SF-6D. One study providing evidence of the psychometric properties of the SF-6D was identified in this update.

In a US comparison of five generic, preference-based measures, the SF-6D had the strongest association with seizure severity and seizure control, supporting the construct validity of the measure, as these are the epilepsy characteristics that are known to be associated with differences in HRQoL (Langfitt et al., 2006).

The SF-6D was found to discriminate between patients with and without seizures two years after baseline evaluation (Langfitt et al., 2006).

The baseline distribution of SF-6D scores was normally distributed and closely resembled that of the QOLIE-89 (Langfitt et al., 2006).

i) WHO Quality of Life assessment instrument – abbreviated version (WHOQOL-BREF)
No evidence was identified in the previous review for the WHOQOL-BREF. One study providing evidence of the psychometric properties of the WHOQOL-BREF was identified in this update.

The WHOQOL-BREF did not discriminate between those with epilepsy and those with the comparable neurocardiogenic syncope, but did discriminate both of these diseases from controls. The Physical Health, Psychological, and Environment

domains demonstrated discriminant validity between these two diseases and controls, but the Social Relations domain was not significantly different between the three groups (Santhouse et al., 2007).

The WHOQOL-BREF did not discriminate according to seizure frequency (Santhouse et al., 2007).

Chapter 3: Condition-specific PROMs

Eight condition-specific instruments were included in the previous review:

a) Epilepsy Surgery Inventory-55
b) Katz Adjustment Scale
c) Liverpool Quality of Life (LQOL) Battery and Seizure Severity Scale
d) Quality of Life in Epilepsy-89 (QOLIE-89)
e) Quality of Life in Epilepsy-31 (QOLIE-31)
f) Quality of Life in Epilepsy-10 (QOLIE-10)
g) Side-Effect and Life Satisfaction (SEALS) Inventory
h) Washington Psychosocial Seizure Inventory

This update identified further evidence of performance for the QOLIE-89, QOLIE-31 and the QOLIE-10.

Full details of the development, domains and scoring methods are detailed in Appendices C and D.

a) Epilepsy Surgery Inventory-55 (ESI-55)
Seven studies were identified which evaluated the ESI-55 in the previous review, making it one of the most used and tested instruments reviewed. One further study reporting evidence for the responsiveness of the ESI-55 was identified in this update.

The following domains of the ESI-55 demonstrated responsiveness to patients following temporal lobectomy in those patients who scored low to medium pre-surgery: Health Perceptions; Energy Fatigue; Social Function; Cognitive Function and Role Limitations due to physical problems (Rose et al., 1996). No changes were observed in patients scoring high on the ESI-55 pre-surgery.

b) Katz Adjustment Scale
One study evaluating the Katz Adjustment Scale was identified in the previous review, offering little evidence to support psychometric and operational properties. No further studies evaluating the Katz Adjustment Scale were identified in this update.

c) Liverpool Quality of Life (LQOL) Battery and Seizure Severity Scale
Nine studies evaluating the LQOL Battery and Seizure Severity Scale were identified in the previous review, offering only adequate evidence supporting measurement properties. No further studies reporting on the measurement properties of the LQOL were identified in this review.

d) Quality of Life in Epilepsy-89 (QOLIE-89)
Thirteen studies were identified which evaluated the QOLIE-89 in the previous review, with all measurement and operational criteria reported and offering support for the QOLIE-89 being one of the most psychometrically sound instruments reviewed. A further nine studies have been identified in this update, offering further evidence on the psychometric and operational strengths of the QOLIE-89.

The QOLIE-89 converged with scores on the POMS only when personality factors (as measured via the Minnesota Multiphasic Personality Inventory (MMPI)) were not

taken into consideration (Testa et al., 2007). In another study with medication-resistant patients, the QOLIE-89 demonstrated convergent validity with POMS Depression/Dejection as well as with the Adverse Events Profile (Szaflarski et al., 2006; Gilliam et al., 2006). The QOLIE-89 also converged with the Neurological Disorders Depression Inventory for Epilepsy (NDDI-E) (Gilliam et al., 2006). Strong and significant correlations have been observed between the QOLIE-89 and Washington Psychosocial Seizure Inventory (WPSI) (Pramuka et al., 2007).

The QOLIE-89 could significantly discriminate between those with psychogenic non-epileptic seizures (PNES) with epileptic seizures (ES), as expected (Testa et al., 2007). Further support for the discriminant validity of the QOLIE-89 has been shown in terms of discriminating between patients who had received cerebral resective surgery as opposed to vagus nerve stimulation (VNS) or medication; however, the measure could not discriminate between the latter two (McGlone et al., 2008). The Memory subscale could not discriminate according to medication type (Hamberger et al., 2007), but the Cognitive Distress subscale could discriminate medication usage in patients who had not experienced seizures post-surgery (Spencer et al., 2007). In the latter study, no other subscales could make this discrimination. Physical Health and Epilepsy-Targeted Health subscales in this cohort could discriminate between those patients with greater time seizure free, whilst Cognitive Distress or Mental Health could not.

QOLIE-89 Total and subscale scores were significantly responsive in patients following surgery at 6 months (Spencer et al., 2007). Furthermore, changes in QOLIE scores were sensitive to the time since last seizure. Further support for the responsiveness of the QOLIE-89 is evidenced by Pramuka et al. (2007), who demonstrated significant improvements in the Role Limitations-Emotional score in patients receiving a self-management intervention; however, no significant differences were evident in overall quality of life.

The Memory subscale failed to demonstrate responsiveness in patients receiving different medications (Hamberger et al., 2007) and VNS (McGlone et al., 2008). However, in the former study, the authors did note that the medication might only be beneficial in particular epilepsy syndromes, and thus the medication may be ineffective as opposed to the instrument.

Significant main effects of time have been found for the QOLIE-89 Total score (McGlone et al., 2008; Langfitt et al., 2006).

Testa et al. (2007) used a 10-point change in HRQOL to interpret their findings, this cut-off point being considered a clinically important difference with regard to the QOLIE-89.

Szaflarski et al. (2006) report 98% completion rates for the total QOLIE-89, demonstrating minimal missing data.

e) Quality of Life in Epilepsy-31 (QOLIE-31)
Seven studies were identified which evaluated the QOLIE-31 in the previous review, with evidence being good but not as extensive as the parent measure (QOLIE-89). A further seven studies have been identified in this update, offering further evidence supporting the robustness of QOLIE-31.

The construct validity of the QOLIE-31 is supported via convergence with depression scores as measured via the Beck Depression Inventory (BDI-II), ability to drive, and improved seizure control (Tracy et al., 2007). Convergent validity was also supported for the Emotional Well-Being, Energy/Fatigue, Cognitive Functioning, Medication Effects, Social Functioning, and Health Concept subscales of the QOLIE-31 and BDI-II scores. This convergence was found to be direct and not mediated by other variables, such as seizure type and frequency. In terms of physiological outcomes, the Seizure Worry and Social Functioning domains of the QOLIE-31 correlated highly significantly with seizure severity (Harden et al., 2007).

The QOLIE-31 could not discriminate according to medication type (Hamberger et al., 2007), but did demonstrate an ability to discriminate according to age (Sachedo et al., 2006) and depressed mood (Tracy et al., 2007). In a UK study with temporal lobe epilepsy patients, discriminant validity was not supported for patients with and without ictal fear (the most commonly reported emotional aura in this type of epilepsy) (Reynders et al. 2005).

QOLIE-31 Total score improved by 10 points from baseline in patients completing adjunctive therapy and by 15 points from baseline in patients completing lamotrigine monotherapy (Kustra et al., 2005). Support for the responsiveness of the QOLIE-31 composite and subscale scores (with the exception of Energy/Fatigue) has also been demonstrated in an open-label, single-arm study whereby patients' current medication was changed to oxcarbazepine monotherapy (Sachedo et al., 2006). In another open-label study, the QOLIE-31 Total and seven subscales demonstrated responsiveness in patients switching from immediate-release carbamazepine to extended-release carbamazepine (Ficker et al., 2005). The QOLIE-31 did not demonstrate responsiveness to change in patients receiving medication in localization-related epilepsy, in contrast to significant effect sizes in the QOLIE-89 Memory subscale (Hamberger et al., 2007).

Concern has been raised as to the strong influence of mood states, such as depression and anxiety, on QOLIE scores. Tracy et al. (2007) found that the influence of depression on QOLIE scores was greater than the influence of illness factors, such as seizure control, thus raising important interpretation issues.

Sachedo et al. (2006) and Kustra et al. (2005) used a 10% change from baseline as a criterion to denote clinically noticeable differences in QOLIE-31 scores.

Sachedo et al. (2006) report a completion rate of 91% for patients taking part at baseline and follow-up.

f) Quality of Life in Epilepsy-10 (QOLIE-10)
Two studies were identified which evaluated the QOLIE-10 in the previous review, offering limited evidence to support measurement properties when compared to the

QOLIE-89 and QOLIE-31. A further nine studies have been identified in this update, offering further evidence for the construct validity and responsiveness of the QOLIE-10.

Scores on the QOLIE-10 correlated with number of clinic visits (Bautista et al., 2008) and satisfaction with care (Bautista et al., 2007), being more predictive of frequency of clinic visits than the number of seizures.

Multivariate regression analysis provides support that QOLIE-10 scores can discriminate patients who had two or less emergency room visits over the previous year compared to those with more than two visits (Bautista et al., 2008). The discriminant validity of the QOLIE-10 increased when combined with seizure frequency. The QOLIE-10 also discriminated patients on different doses of gabapentin (Beran et al., 2001). In terms of treatment outcomes, Cramer (2001) demonstrated the QOLIE-10 to be able to discriminate between responders and non-responders of VNS.

Responsiveness has been supported by significant change in QOLIE-10 scores to concomitant sleep disturbances (Xu et al., 2006), adjunctive therapy with Levetiracetam (Somerville et al., 2007), VNS (Cramer, 2001), seizure frequency and in-patient admissions (Bautista et al., 2008). Significant scores changes have also been found in response to gabapentin as adjunctive therapy for partial seizures (Bruni, 1998; Bruni, 1999; Beran et al., 2001).

As with the QOLIE-31, concern has been raised regarding the strong influence of mood states on QOLIE scores. Cramer et al. (2005) found clinically important reductions (25-29%) in QOLIE-10 scores in epilepsy patients with mild anxiety or depression, which declined further with moderate and severe levels of these moods. This raises important interpretation issues.

g) Side-Effect and Life Satisfaction (SEALS) Inventory

Three studies were identified which evaluated the SEALS in the previous review, offering little evidence of psychometric and operational testing. No further studies evaluating the SEALS with epilepsy have been identified in this update.

h) Washington Psychosocial Seizure Inventory (WPSI)

Four studies were identified which evaluated the WPSI in the previous review, reporting little evidence in favour of the instrument when compared to other PROMs reviewed. One further study supporting the construct validity of the WPSI has been identified in this update.

Strong and significant correlations have been observed between the WPSI and the QOLIE-89 (Pramuka et al., 2007).

Chapter 4: DISCUSSION

2006 Review
The 2006 review identified seven generic and eight epilepsy-specific instruments. Details of the content, domains and scoring are outlined in Appendices C and D.

From the previous review, the most frequently evaluated generic PROM was the SF-36. The most comprehensively evaluated epilepsy-specific PROMs were the ESI-55, QOLIE-89, QOLIE-31, WPSI, and the LQOL Battery and Seizure Severity Scale.

Based on this evidence, the previous review recommended the following most extensively used and tested PROMs for use with people with epilepsy: the SF-36 as a generic PROM for the broad evaluation of health, and the ESI-55 and QOLIE-89 for condition-specific evaluations of health.

Review update 2009

Generic PROMs
Table 2 summarises the psychometric criteria and operational characteristics of the generic PROMs included in the review using the appraisal criteria found in Appendix B. The previous review reported evidence for the EQ-5D, HUI, NHP, Q-TWIST, SF-36, SF-12, and HUI. It is clear from the appraisal of the evidence, as detailed in Table 2, that the SF-36 remains the most frequently used generic PROM overall with people with epilepsy, with good measurement and operational performance reported.

The EQ-5D is the second most utilised generic measure in epilepsy patients and more studies reporting on the EQ-5D as opposed to the SF-36 were identified in this update. Although this update offers more favourable evidence for the EQ-5D than the previous review, there still remain a number of concerns regarding the psychometric properties of this measure, particularly in terms of evidence demonstrating high ceiling effects. However, it needs to be considered further because of its ability to generate utilities.

Whilst further support has been obtained for the construct validity of HUI, the evidence is too limited to make recommendations for this instrument with epilepsy patients.

The SF-6D was identified in this update, demonstrating construct validity, responsiveness, and correlations with the condition-specific QOLIE-89. Further studies are needed, utilising this instrument with epilepsy patients.

The WHOQOL-BREF was identified in this update with very limited evidence of performance.

No further evidence was found for the NHP, Q-TWIST, SF-12, or SIP.

Table 2: Appraisal of psychometric and operational performance of generic PROMs for people with epilepsy

PROM (no. of studies)	Reproducibility	Internal consistency	Validity: Content	Validity: Construct	Responsiveness	Interpretability	Floor/ceiling/ precision	Acceptability	Feasibility
EQ-5D (6)	0	n/a	_	+	+	0	_	_	0
HUI (3)	0	0	0	+	0	+	+	0	0
NHP (3)	0	+	0	+	_	0	0	0	0
Q-TWIST (1)	0	0	0	0	0	0	0	0	0
SF-36 (10)	+	++	+	++	+	0	_	+	0
SF-12 (1)	0	0	0	+	+	0	0	0	0
SF-6D (1)	0	0	0	+	+	0	+	0	0
SIP (1)	0	+	0	+	0	0	0	_	0
WHOQOL-BREF (1)	0	0	0	+	0	0	0	0	0

Psychometric and operational criteria

0 *not reported*

— *no evidence in favour*

\+ *some limited evidence in favour*

++ *some good evidence in favour*

+++ *good evidence in favour.*

Epilepsy-specific PROMs

Table 3 summarises the psychometric criteria and operational characteristics of the condition-specific PROMs included in the review using the Appraisal criteria outlined in Appendix B. The previous review reported and recommended the ESI-55 and QOLIE-89 based on the number of evaluations reporting good measurement and operational performance. With the evidence that has emerged from this review, the recommended instruments for short-listing are the QOLIE-89, QOLIE-31 and ESI-55 (Table 3).

The majority of evidence obtained for this update was for the QOLIE instruments – versions 10, 31, and 89. The QOLIE-89 remains the most robust in terms of psychometric performance, but an increase in evidence for the shorter QOLIE-31 might make this a more pragmatic option for use in the NHS. There is evidence that the additional 58 items in the QOLIE-89 do not significantly improve its ability to detect real clinically important change. Evidence supporting the measurement properties of the even shorter QOLIE-10 is limited in comparison to its parent versions.

The ESI-55 was one of the most extensively used and tested condition-specific PROMs in the previous review. However, this update has found limited further evidence to support the instruments psychometric and operational properties. The instrument is primarily utilised as an epilepsy surgery outcome and thus further testing is needed in more diverse epilepsy contexts.

Limited support was obtained for the construct validity of the WPSI.

No further evaluations were identified for the following PROMs, which were reported in the previous review: Katz Adjustment Scale, LQOL Battery and Seizure Severity Scale, or SEALS Inventory.

RECOMMENDATIONS

Tables 2 and 3 summarise the evidence of measurement and operational performance applying the adapted appraisal criteria of the PROMs identified in these reviews. Based on this appraisal, the following instruments were presented to a multidisciplinary panel for discussion (see Appendix F).

Generic measures:

1. EQ-5D
2. SF-36

Epilepsy-specific measures:

3. ESI-55
4. QOLIE-89
5. QOLIE-31

The multi-disciplinary panel were favourable toward the SF-36 as generic measure of health status and the QOLIE-31 as an epilepsy specific measure. However, it was noted that the QOLIE instruments included items from the SF-36 and therefore duplication of items if the two instruments were used together may increase

completion burden to patients. They felt that neither instrument should be used in isolation if the full range of patient experience is to be captured. Having in mind an overall strategy of a generic and condition-specific measure being used in combination to assess complementary aspects of health status, and also having in mind the need for an approach that reduces the volume of questions and likely burden of responding, the current review recommends the combination of EQ-5D and the QOLIE-31 for use in potentially large scale population studies. The multidisciplinary panel commented on the ease of use of EQ-5D. The simplicity and the brevity of the EQ-5D, make it likely that it will not adversely influence response rates. The fact that it yields UK-derived preference values makes it an attractive generic measure providing complementary evidence on health status alongside QOLIE-31.

Table 3: Appraisal of psychometric and operational performance of condition-specific PROMs for people with epilepsy

PROM (no. of studies)	Reproducibility	Internal consistency	Validity: Content	Validity: Construct	Responsiveness	Interpretability	Floor/ceiling/ precision	Acceptability	Feasibility
ESI-55 (8)	0	+	0	++	++	0	0	0	+
Katz Adjustment Scale (1)	0	+	+	+	+	0	0	0	0
LQOL (9)	_	++	0	++	_	0	+	+	0
QOLIE-89 (22)	++	++	++	++	++	+	+	+	+
QOLIE-31 (14)	+	+	++	+	++	+	+	+	0
QOLIE-10 (11)	+	_	0	++	++	0	0	0	_
SEALS (3)	+	0	0	_	+	0	0	0	0
WPSI (5)	+	+	+	+	_	0	0	0	0

Psychometric and operational criteria

0 *not reported*

— *no evidence in favour*

\+ *some limited evidence in favour*

\++ *some good evidence in favour*

\+++ *good evidence in favour*

Chapter 5: Epilepsy-specific PROMs for children

The measurement of health-related QoL in adults with epilepsy is now well-established and the PROMs identified in this review have been developed primarily for use with this population. Very few evaluations have applied these in paediatric populations and, furthermore, the development of epilepsy-specific measures for children is rare.

Two of the most commonly reported paediatric epilepsy-specific PROMs are:

a) Health-Related Quality of Life in Children with Epilepsy
b) Quality of Life in Epilepsy Inventory for Adolescents (QOLIE-AD-48)

a) HRQoL in Children with Epilepsy (Ronen et al., 2003)
The HRQoL in Children with Epilepsy is a 25-item measure developed for children with epilepsy aged 8-15 years. Content validity is supported via the use of patient focus groups in the development of the instrument (Ronen et al., 1999; Ronen et al., 2001). The development study for this instrument (Ronen et al., 2003) reports adequate scaling properties with no significant floor or ceiling effects, internal consistency of above 0.70 on all subscales apart from the Quest for Normality subscale, and test-retest reliability of 0.59 or higher over a two-week period. Convergent validity with a number of clinical criteria, including health-care utilisation and seizure severity, has been demonstrated. The developers have also created a parent-proxy measure, acknowledging that at times proxy outcomes are necessary with this population. The HRQoL in Children with Epilepsy measure has been acknowledged as involving a comprehensive development process with further evaluation of item discrimination across scales being an important area for future development (Cowan & Baker, 2004).

b) Quality of Life in Epilepsy Inventory for Adolescents (Cramer et al., 1999)
The QOLIE-AD-48 is the adolescent version of the QOLIE family of instruments evaluated in this review. It is a 48-item self-report scale designed for adolescents aged 11-17 years. Content validity is supported through the use of literature reviews, reviews of existing measures, patient focus groups, and expert professional opinion, in the development of the instrument (Cramer at al., 1999). The development study for this instrument (Cramer at al., 1999) reports no indication of floor or ceiling effects, internal consistency of above 0.70 for the summary score and for all subscale scores apart from the Health Perceptions subscale, and test-retest reliability of 0.83 over a four-week period. Discriminant validity between groups differing in seizure severity has been demonstrated. A further study provides support for the discriminant validity of the QOLIE-AD-48 in terms of age, socio-economic status, and severity of neurotoxicity symptoms (Devinsky et al., 1999). As noted in the evaluated adult QOLIE measures, this instrument incorporates a generic measure of health-related QoL, an approach which has been recommended in the development of paediatric PROMs assessing chronic conditions (Levi & Drotar, 1998). The QOLIE-AD-48 has been reported as meeting many of the psychometric criteria necessary for a robust instrument with further investigation of the construct validly of the scale on larger populations being an important area for future development (Cowan and Baker, 2004).

Many paediatric-specific generic PROMs have been developed and measurement performance is widely reported. Examples include the Child Health Questionnaire (CHQ), Child Health and Illness Profile (CHIP), KIDSCREEN, Pediatric Quality of Life Inventory (PedsQL), and TNO-AZL[3] Children's Quality of Life (TACQOL). As part of a systematic review of measures of QoL for children, Eiser and Morse (2001) conclude that only three of 19 generic measures fulfil very basic psychometric criteria, these being the CHQ, PedsQL, and HUI. However, the evidence-base for these measures with children with epilepsy is beyond the scope of this discussion.

There are no reports of evaluations with children for the other epilepsy-specific PROMs included in this review.

The remit of this chapter was to examine whether adult epilepsy-specific PROMs could be applied with children in the NHS. There is very little evidence to support recommending any of the adult PROMs identified in this review for use with children. Moreover, the conceptual and methodological challenges of adopting adult PROMs with children strengthen such a judgement. There are, however, a few paediatric PROMs which show promise for their utilisation in measuring epilepsy outcomes in children, such as the HRQoL in Children with Epilepsy and QOLIE-AD-48. If the latter were to demonstrate the psychometric properties of its adult counterpart, this could enhance our understanding of patient-reported outcomes in children.

[3] Toegepast Natuurwetenschappelijk Onderzoek-Academisch Ziekenhuis Leiden

APPENDIX A

i. Sources for PROM bibliography

1. AMED: Allied and Complementary Medicine Database
2. Biological Abstracts (BioAbs)
3. BNI: British Nursing Index Database, incorporating the RCN (Royal College of Nursing) Journals Database
4. CINAHL: Cumulative Index to Nursing and Allied Health Literature
5. Econlit - produced by the American Economic Association
6. EMBASE - produced by the scientific publishers Elsevier
7. MEDLINE - produced by the US National Library of Medicine
8. PAIS: Public Affairs Information Service
9. PsycINFO (formerly PsychLit) - produced by the American Psychological Association
10. SIGLE: System for Information on Grey Literature in Europe
11. Sociofile: Cambridge Scientific Abstracts Sociological Abstracts Database
12. In addition, all records from the journal ‘Quality of Life Research’ are downloaded via Medline.

ii. PROM Bibliography search strategy

a. records to December 2005 (downloads 1-12)

((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)) adj2 (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) adj5 ((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)) adj2 (index or indices or instrument or instruments or measure or measures or questionnaire$ or profile$ or scale$ or score$ or status or survey$))))

b. records from January 2006 (download 13)

(((acceptability or appropriateness or component$ analysis or comprehensibility or effect size$ or factor analys$ or factor loading$ or feasibility 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 valid$ or valuation or weighting$)

and

(HRQL or HRQoL or QL or QoL or qualit$ of life or quality adjusted life year$ or QALY$ or disability adjusted life year$ or DALY$ or COOP or SF-12 or SF-20 or SF-36 or SF-6 or standard gamble or summary score$ or time trade off or health index or health profile or health status or ((patient or self$) adj (rated or reported or based or assessed)) or ((disability or function$ or subjective or utilit$ or well?being) adj2 (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) adj5 (HRQL or HRQoL or QL or QoL or qualit$ of life or quality adjusted life year$ or QALY$ or disability adjusted life year$ or DALY$ or COOP or SF-12 or SF-20 or SF-36 or SF-6 or standard gamble or summary score$ or time trade off or health index or health profile or health status or ((patient or self$) adj (rated or reported or based or assessed)) or ((disability or function$ or subjective or utilit$ or well?being) adj2 (index or indices or instrument or instruments or measure or measures or questionnaire$ or profile$ or scale$ or score$ or status or survey$)))))

APPENDIX B: Psychometric criteria

Appraisal of PROMs

The methods that will be used for assessing the performance of PROMs were developed and tested against multidisciplinary consensus and peer review. They focus on explicit criteria to assess reliability, validity, responsiveness, precision, acceptability, and feasibility. A pragmatic combination of the criteria developed and used in previous reports to DH by the Oxford and LSHTM groups will be used.

The appraisal framework focuses on psychometric criteria and PROMs must fulfil some or all to be considered as a short-listed instrument. Practical or operational characteristics are also assessed (acceptability and feasibility).

Once evidence has been assessed for eligibility, records considered as inclusions will be assembled for each PROM identified. Measurement performance and operational characteristics will be appraised independently by two reviewers using the following rating scale, and inter-rater reliability calculated.

Psychometric evidence:
– = evidence does not support criteria
0 = not reported or no evidence in favour
+ = some limited evidence in favour
++ = some good evidence in favour, but some aspects do not meet criteria or some aspects not reported
+++ = good evidence in favour

PROMs for which there are strong psychometric properties will be judged in terms of operational characteristics and clinical credibility.

Appraisal criteria (adapted from Smith et al., 2005 and Fitzpatrick et al., 1998; 2006)

Appraisal component	Definition/test	Criteria for acceptability
Reliability		
Test-retest reliability	The stability of a measuring instrument over time; assessed by administering the instrument to respondents on two different occasions and examining the correlation between test and re-test scores	Test re-test reliability correlations for summary scores 0.70 for group comparisons
Internal consistency	The extent to which items comprising a scale measure the same construct (e.g. homogeneity of items in a scale); assessed by Cronbach's alpha's and item-total correlations	Cronbach's alphas for summary scores ≥0.70 for group comparisons Item-total correlations ≥ 0.20
Validity		
Content validity	The extent to which the content of a scale is representative of the conceptual domain it is intended to cover; assessed qualitatively during the questionnaire development phase through pre-testing with patients. Expert opinion and literature review	Qualitative evidence from pre-testing with patients, expert opinion and literature review that items in the scale represent the construct being measured Patients involved in the development stage and item generation
Construct validity	Evidence that the scale is correlated with other measures of the same or similar constructs in the hypothesised direction; assessed on the basis of correlations between the measure and other similar measures	High correlations between the scale and relevant constructs preferably based on a priori hypothesis with predicted strength of correlation
	The ability of the scale to differentiate known-groups; assessed by comparing scores for sub-groups who are expected to differ on the construct being measured (e.g a clinical group and control group)	Statistically significant differences between known groups and/or a difference of expected magnitude
Responsiveness	The ability of a scale to detect significant change over time; assessed by comparing scores before and after an intervention of known efficacy (on the basis of various methods including t-tests, effect sizes (ES), standardised response means (SRM) or responsiveness statistics	Statistically significant changes on scores from pre to post-treatment and/or difference of expected magnitude
Floor/ceiling effects	The ability of an instrument to measure accurately across full spectrum of a construct	Floor/ceiling effects for summary scores <15%
Practical properties		
Acceptability	Acceptability of an instrument reflects respondents' willingness to complete it and impacts on quality of data	Low levels of incomplete data or non-response
Feasibility/burden	The time, energy, financial resources, personnel or other resources required of respondents or those administering the instrument	Reasonable time and resources to collect, process and analyse the data

APPENDIX C: GENERIC PROMS USED IN EPILEPSY

This appendix provides a brief description of the nine generic health status instruments reviewed in the update on patient-reported instruments for epilepsy, and summarises their origins, development, and content. Content, format, and health status domains included are set out in table form at the end of this appendix.

a) EuroQol-EQ-5D(The EuroQol Group, 1990; revised 1993)

The European Quality of Life instrument (EuroQol) was developed by researchers in five European countries to provide an instrument with a core set of generic health status items (The EuroQol Group, 1990; Brazier et al., 1993). Although providing a limited and standardized reflection of HRQoL, it was intended that use of the EuroQol would be supplemented by disease-specific instruments. The developers recommend the EuroQol for use in evaluative studies and policy research; given that health states incorporate preferences, it can also be used for economic evaluation. It can be self or interview-administered.

Existing instruments, including the Nottingham Health Profile, Quality of Well-Being Scale, Rosser Index, and Sickness Impact Profile were reviewed to inform item content (The EuroQol Group, 1990). There are two sections to the EuroQol: the EQ-5D and the EQ thermometer. The EQ-5D assesses health across five domains: anxiety/depression (AD), mobility (M), pain/discomfort (PD), self-care (SC), and usual activities (UA), as shown in Table 4. Each domain has one item and a three-point categorical response scale; health 'today' is assessed. Weights based upon societal valuations of health states are used to calculate an index score of –0.59 to 1.00, where –0.59 is a state worse than death and 1.00 is maximum well-being. A score profile can be reported. The EQ thermometer is a single 20 cm vertical visual analogue scale with a range of 0 to 100, where 0 is the worst and 100 the best imaginable health.

b) Health Utilities Index (Feeny et al., 1995)

The Health Utilities Index (HUI) was designed as a comprehensive measure of health status and health-related quality of life. The Health Utilities Index (Mark 3) is a system composed of a health status classification which defines 972,000 discrete health states, and a preference, or utility, function which can be used to calculate the desirability for each health state. The HUI3 health status classification was developed by Feeny et al. (1995) to assess capacity on eight dimensions: vision, hearing, speech, ambulation, dexterity, emotion, cognition, and pain/discomfort. The utility function reflects community preferences and scores each unique health state on a scale ranging from 0 (death) to 1 (perfect health).

c) Nottingham Health Profile (Hunt et al., 1980)

The Nottingham Health Profile (NHP) was developed in the UK during the 1970s for use in the evaluation of medical or social interventions (Hunt et al., 1980). Instrument content was derived from over 2000 statements given by 768 patients with a variety of chronic ailments, and other laypeople.

Part I of the instrument has 38 items across six domains: bodily pain (BP), emotional reactions (ER), energy (E), physical mobility (PM), sleep (S), and social isolation (SI), as shown in Table 4. All items are statements that refer to departures from

normal functioning, and relate to feelings and emotional state rather than change in behaviour. Respondents answer 'yes' or 'no' according to whether or not they feel the item applies to them in general. Positive responses are weighted and summed to give six domain scores between 0 and 100, where 100 denotes maximum limitation.

Part II of the NHP is less widely used and provides a brief indicator of handicap. The instrument may be self-, interview-, or telephone-administered.

d) SF-36: Medical Outcomes Study 36-item Short Form Health Survey (Ware and Sherbourne, 1992; Ware et al., 1994; Ware, 1997)

The Medical Outcomes Study (MOS) Short Form 36-item Health Survey (SF-36) is derived from the work of the Rand Corporation during the 1970s (Ware & Sherbourne, 1992; Ware et al., 1994; Ware, 1997). It was published in 1990 after criticism that the SF-20 was too brief and insensitive. The SF-36 is intended for application in a wide range of conditions and with the general population. Ware et al. (1994; 1997) proposed that the instrument should capture both mental and physical aspects of health. International interest in this instrument is increasing, and it is by far the most widely evaluated measure of health status (Garratt et al., 2002).

Items were derived from several sources, including extensive literature reviews and existing instruments (Ware & Sherbourne, 1992; Ware & Gandek, 1998; Jenkinson & McGee, 1998). The original Rand MOS Questionnaire (245 items) was the primary source, and several items were retained from the SF-20. The 36 items assess health across eight domains (Ware, 1997), namely bodily pain (BP: two items), general health perceptions (GH: five items), mental health (MH: five items), physical functioning (PF: ten items), role limitations due to emotional health problems (RE: three items), role limitations due to physical health problems (RP: four items), social functioning (SF: two items), and vitality (V: four items), as shown in Table 4. An additional health transition item, not included in the final score, assesses change in health.

All items use categorical response options (range: 2-6 options). Scoring uses a weighted scoring algorithm and a computer-based programme is recommended. Eight domain scores give a health profile; scores are transformed into a scale from 0 to 100 scale, where 100 denotes the best health. Scores can be calculated when up to half of the items are omitted. Two component summary scores for physical and mental health (MPS and MCS, respectively) can also be calculated. A version of the SF-36 plus three depression questions has been developed and is variously called the Health Status Questionnaire (HSQ) or SF-36-D. The SF-36 can be self-, interview-, or telephone-administered.

e) SF-12: Medical Outcomes Study 12-item Short Form Health Survey (Ware et al., 1995)

In response to the need to produce a shorter instrument that could be completed more rapidly, the developers of the Medical Outcomes Study (MOS) 36-item Short Form Health Survey (SF-36) produced the 12-item Short Form Health Survey (SF-12) (Ware et al., 1995).

Using regression analysis, 12 items were selected that reproduced 90% of the variance in the overall Physical and Mental Health components of the SF-36 (Table 4). The

same eight domains as the SF-36 are assessed and categorical response scales are used. A computer-based scoring algorithm is used to calculate scores: Physical Component Summary (PCS) and Mental (MCS) Component Summary scales are generated using norm-based methods. Scores are transformed to have a mean value of 50, standard deviation (SD) 10, where scores above or below 50 are above or below average physical or mental well-being, respectively.

Completion by UK city-dwellers reporting the absence of health problems yielded a mean PCS score of 50.0 (SD 7.6) and MCS of 55.5 (SD 6.1) (Pettit et al., 2001). Although not recommended by the developers, Schofield & Mishra (1998) report eight domain scores and two summary scores. The SF-12 may be self-, interview-, or telephone-administered.

Several authors have proposed simplification of the scoring process and revision of the SF-12 summary score structure, where norm-based weighting is replaced by item summation to facilitate score interpretation (Resnick & Nahm, 2001; Resnick & Parker, 2001).

f) SF-6D (Brazier et al., 2002)

The SF-6D was designed to be used in health economic analyses. The instrument is composed of six multi-level dimensions and is a preference based algorithm based on a sub-set of items from the SF-36, developed by Brazier et al. (2002). The SF-6D comes with a set of preference weights obtained from a sample of the general population. Using the valuation technique of standard gamble, members of the general population were asked to value a selection of health states from which a model has been estimated to predict all the health states described by the SF-6D.

g) Sickness Impact Profile (Bergner et al., 1976; revised: Bergner et al., 1981)

The Sickness Impact Profile (SIP) was developed in the USA to provide a broad measure of self-assessed health-related behaviour (Bergner et al., 1976; Bergner et al., 1981). It was intended for a variety of applications, including programme-planning and assessment of patients, and to inform policy decision-making (Bergner et al., 1976; Bergner et al., 1981; McDowell & Newell, 1996).

Instrument content was informed by the concept of ‘sickness’, which was defined as reflecting the change in an individual’s activities of daily life, emotional status, and attitude as a result of ill-health (McDowell & Newell, 1996). Item derivation was based on literature reviews and statements from health professionals, carers, patient groups, and healthy subjects describing change in behaviour as a result of illness. The SIP has 136 items across 12 domains: alertness behaviour (AB: ten items), ambulation (A: 12 items), body care and movement (BCM: 23 items), communication (C: nine items), eating (E: nine items), emotional behaviour (EB: nine items), home management (HM: ten items), mobility (M: ten items), recreation and pastimes (RP: eight items), sleep and rest (SR: seven items), social interaction (SI: 20 items) and work (W: nine items).

Each item is a statement. Statements that best describe a respondent’s perceived health state on the day the instrument is completed are ticked. Items are weighted, with higher weights representing increased impairment. The SIP percentage score can be calculated for the total SIP (index) or for each domain, where 0 is better health and

100 is worse health. Two summary scores are calculated: Physical function (SIP-PhysF), a summation of A, BCM, and M, and psychosocial function (SIP-PsychF), a summation of AB, C, EB, and SI. The five remaining categories are scored independently. The instrument may be self or interview-administered.

The Functional Limitation Profile (FLP) is an Anglicized version of the SIP (Patrick & Peach, 1989; McDowell & Newell, 1996). Wording and some weightings have been altered, and summary scores are calculated using different dimensions to those used in the SIP (i.e. FLP Physical summary calculated by summing A, BCM, M and HM; FLP Psychosocial summary calculated by summing RP, EB, AB, SI and SR. Several abbreviated versions of the SIP have been developed, including a 68-item version (De Bruin et al., 1992; Post et al, 1996).

h) WHOQOL-BREF (The WHOQOL Group, 1998)

The WHOQOL-BREF is a shortened version of the original WHOQOL instrument. It is an international, cross-culturally comparable quality of life assessment instrument. It assesses the individual's perceptions in the context of their culture and value systems, and their personal goals, standards, and concerns. The WHOQOL instruments were developed collaboratively in a number of centres worldwide, and have been widely field-tested (WHOQOL Group, 1998). The WHOQOL-BREF instrument comprises 26 items, which measure the following broad domains: physical health, psychological health, social relationships, and environment.

Table 4: **Generic PROMS used in epilepsy - content and scoring**

Instrument	***Domains (no. items)***	***Response options***	***Score***	***Completion (time in minutes)***
European Quality of Life Questionnaire (EuroQol-EQ5D) (5+1)	EQ-5D Anxiety/depression (1), Mobility (1), Pain/discomfort (1), Self-care (1), Usual activities (1) EQ-thermometer Global health (1)	EQ-5D Categorical: 3 options *EQ-thermometer* VAS Current health	EQ-5D Summation: domain profile Utility index (–0.59 to 1.00) *Thermometer* VAS (0-100)	Interview or self
Health Utilities Index 3 (Feeny et al., 1995) (8)	Vision, Hearing, Speech, Ambulation, Dexterity, Emotion, Cognition, Pain	Four domains have five response options and five have six response options	Global Utility index and single attribute utility scores for the eight separate dimensions	Self report, face to face, and telephone
Nottingham Health Profile (NHP) (38)	Bodily pain (BP) (8), Emotional reactions (ER) (9), Energy (E) (3), Physical mobility (PM) (8), Sleep (S) (5), Social isolation (SI) (5)	Yes/no; positive responses weighted Recall 'general' health	Algorithm Domain profile 0-100, 100 is maximum limitation	Interview Self (10-15)
SF-36: MOS 36-item Short Form Health Survey (36)	Bodily pain (BP) (2), General health (GH) (6), Mental health (MH) (5), Physical functioning (PF) (10), Role limitation-emotional (RE) (3), Role limitation-physical (RP) (4), Social functioning (SF) (2), Vitality (4) *plus 1 health transition question not included in final score*	Categorical: 2-6 options Recall: standard 4 weeks, acute 1 week	Algorithm Domain profile (0-100, 100 best health) Summary: Physical (PCS), Mental (MCS) (mean 50, s.d. 10)	Interview (mean values 14-15) Self (mean 12.6)
SF-12: MOS 12-item Short Form Health Survey (12)	Bodily pain (BP) (1), Energy/Vitality (V) (1), General health (GH) (1), Mental health (MH) (2), Physical functioning (PF) (2), Role limitation-emotional (RE) (2), Role limitation-physical (RP) (2), Social functioning (SF) (1)	Categorical: 2-6 options Recall: standard 4 weeks, acute 1 week	Algorithm Domain profile (0-100, 100 best health) Summary: Physical (PCS), Mental (MCS) (mean 50, s.d. 10)	Interview or self
SF-6D: MOS 6-item Short Form Health Survey (6)	Bodily pain (BP) (1), Energy/Vitality (V) (1), Mental health (MH) (1), Physical functioning (PF) (1), Role limitation (1), Social functioning (SF) (1)	Categorical: 3 options	Algorithm Domain profile (0-100, 100 best health)	Interview or self
Sickness Impact Profile (136)	Alertness behaviour (AB) (10), Ambulation (A) (12) Body care and movement (BCM) (23), Communication (C) (9) Eating (E) (9), Emotional behaviour (EB) (9) Home management (HM) (10), Mobility (M) (10) Recreation and pastimes (RP) (8), Sleep and rest (SR) (7) Social interaction (SI) (20), Work (W) (9)	Check applicable statements. Items weighted: higher weights indicate increased impairment Recall current health	Algorithm Domain profile (0-100%, 100 worst health); Index (0-100%) Summary: Physical (A, BCM, M), Psychosocial function (AB, C, EB, SI)	Interview (range: 21-33) Telephone: PF only (11.5) Self (19.7)
WHOQOL-BREF (26)	physical health, psychological health, social relationships, environment (26 items in total)	% responses categories	Summed scales	Self complete

Table 5: **Generic PROMs used in epilepsy - summary of health status domains** *(after Fitzpatrick et al., 1998)*

Instrument	***Instrument domains***							
	Physical function	**Symptoms**	**Global judgement of health**	**Psychological well-being**	**Social well-being**	**Cognitive functioning**	**Role activities**	**Personal Constructs**
EQ-5D (5+1)	x	x	x	x	x		x	
HUI (8)	x			x		x		
NHP (38)	x	x		x	x			
SF-36 (36)	x	x	x	x	x		x	
SF-12 (12)	x	x	x	x	x		x	
SF-6D (6)	x	x		x	x		x	
SIP (136)	x	x		x	x	x	x	
WHOQOL -BREF (26)	x			x	x			

APPENDIX D: EPILEPSY-SPECIFIC PROMs

This appendix provides a brief description of the eight condition-specific health status instruments reviewed in the update on patient-reported instruments for epilepsy, and summarises their origins, development, and content. Content, format, and health status domains included are set out in table form at the end of this appendix.

a) Epilepsy Surgery Inventory-55 (ESI-55)

The ESI-55 is a 55-item measure of health-related quality of life, designed to assess outcome of epilepsy surgery. It was constructed after a literature review, and includes the Rand SF-36 as a generic core, plus 19 epilepsy-specific items (Vickrey et al., 1992a). The ESI-55 contains 11 multi-item sub-scales of health perceptions, energy/fatigue, overall QoL, social functioning, emotional well-being, cognitive functioning, and role limitations due to emotional problems, role limitations due to memory problems, role limitations due to physical health problems, physical functioning and pain. The initial scale was tested on a small sample of epilepsy patients and then reviewed by health care professionals, before administration to a sample of epilepsy patients to evaluate its reliability and validity (Vickrey et al., 1992a).

Sub-groups of these scales can be weighted and summed to form scores for mental health, physical health and role functioning. The scale scores are weighted and summed to produce the overall score. The health perceptions sub-scale has been reported to have the greatest sensitivity in discriminating between patients varying by seizure type and frequency (Vickrey et al., 1995). The ESI-55 takes an average of 15 minutes to complete. It was reviewed by Devinsky and Vickrey (1994), Selai and Trimble (1995), Jacoby (1996), Leidy et al. (1998) and Buelow and Ferrans (2001).

b) Katz Adjustment Scale

This instrument was originally developed to measure social behaviour and adjustment of people with a diagnosis of schizophrenia (Katz and Lyerly, 1963), but has been extended and adapted for use with people with epilepsy (Vickrey et al., 1992b; and see summary by Trimble and Dodson, 1994). However the scale is completed by relative/friend proxies and not the patients themselves. Vickrey et al. (1992b) increased the items and supported the scale's validity for use with epilepsy patients. Their tested version contains 126 items (127 items should have been included but one was omitted in error), in 14 rather than 12 sub-scales. Proxies are asked to rate the patient according to 'how your relative or friend has looked to you during the past few weeks on these things'. For each item there are four response choices ('almost never' to 'almost always'). The revised instrument by Vickrey et al. (1992b) includes 14 sub-scales: Over-sensitivity/fearfulness, Social Irritability, Dependency, Acting out, Paranoia, and Abnormal thought process, Withdrawal-R, Emotional liability, Nervousness-R, Sociopathy, Bizarreness-R, Hyperactivity-R, Disorientation. The responses are summed and transformed to a 0-100 point scale. Higher values indicate better functioning. The scale has been reviewed by Hermann (1995).

c) Liverpool Quality of Life (LQOL) Battery and Seizure Severity Scale (LSSS)

The aim of the LQOL was to focus on issues relevant and important to people with epilepsy. The LQOL was initially tested using a wide range of existing and new scales. The final version consists of two epilepsy-specific subscales: adverse drug

effects (21 adverse drug effects, rated on a 4-point Likert scale; total sum scores are used in analyses) and impact of epilepsy (9 areas of life that can be affected by epilepsy or treatment, rated on 4-point Likert scales; mean scores are used in analyses). It also includes three general subscales. The first scale is affect-balance (an existing, well tested scale encompassing 5 items describing negative and 5 items describing positive states, rated dichotomously as present or absent. The second scale is sense of mastery (an existing, well tested, 7-item scale, rated on 4-point Likert scales, summed with higher scores representing higher mastery. Thirdly, there is the life fulfilment scale (10 items on areas of life, rated for importance (4-point rating scales) and then again for satisfaction, on 4-point Likert scales. The scale together with the LSSS takes up to 45 minutes to complete. It was reviewed by Hermann (1995), Leidy et al. (1998), and Buelow and Ferrans (2001).

The LSSS contains 20 clinical features or symptoms of seizures over the past 4 weeks, rated on a 4-point Likert scales (total scores range from 20-80). Two subscales measure perceived control over seizures and ictal and post-ictal (11 items) symptoms. Scores are computed by summing item scores, with scores ranging from 9-36 for perceptions, and 11-44 for the ictal scale. Higher scores indicate worse severity.

d) Quality of Life in Epilepsy-89 (QOLIE-89)

The QOLIE-89 is an epilepsy specific measure that includes the 7 sub-scales of the Rand SF-36 as a generic core. It is an extension of a 55-item, QoL questionnaire (the Epilepsy Surgery Inventory – ESI-55) which was designed for use with epilepsy surgery patients. Items judged by the investigators to be missing were included in the QOLIE. It was developed with 304 epilepsy patients and their relative/friend proxies from 25 epilepsy centers in the USA (130 men and 174 women), with a mean age of 36 years (range 17-63). It was repeated 2-3 weeks later (Devinsky et al., 1995). The questionnaire takes an average of 28.4 minutes (SD 15.6, range 6-135) to complete.

It contains 17 multi-item sub-scales/86 items plus 3 single item measures of change in health, sexual relations, overall health (the SF-36 core is supplemented by 53 items specific to epilepsy) grouped into four factors. It aims to assess physical, mental and social areas of life. Standardised methods are used to convert each item to a 0-100 score, with higher scores indicating better QoL. Subscale scores involve averaging across the items in the subscale, with the number of items as the division. The overall score is a weighted sum of the individual subscale scores. Factor-based, standardised regression coefficients (weights) are used to calculate domain scores.

The instrument initially included 99 items at administration, 86 of which, across 17 sub-scales, were retained after multi-trait scaling. Factor analysis of the 17 subscales yielded four underlying dimensions of health: an epilepsy-targeted dimension, cognitive, mental health, and physical health. Construct validity was supported by significant patient-proxy correlations, and correlations between the instrument and seizure frequency over the past year, neuropsychological tests, and emotional and cognitive function. It was reviewed by Devinsky and Vickrey (1994), Jacoby (1996), Leidy et al. (1998), Leppik (1998), and Buelow and Ferrans (2001).

e) Quality of Life in Epilepsy-31 (QOLIE-31)

The QOLIE-31 was developed, using the original dataset, from the 99 items used to develop the QOLIE-89 by Cramer et al. (1998). They selected the subscales that were

reported to be most important by people with epilepsy (as determined by an expert panel), with the result that generic topics (e.g. pain) were excluded.

Following psychometric and factor analyses of the full scale, variables with loadings of greater than or equal to 0.4 were included in the subscales for the QOLIE-31. This resulted in a 31-item questionnaire, with seven subscales, forming two factors: Emotional/Psychological Effects (seizure worry, overall QoL, emotional well-being, energy/fatigue) and Medical/Social Effects (medication effects, work-driving-social limits, cognitive function). Cross-cultural translations were developed. Analyses supported the reliability and validity of the QOLIE-31. It was reviewed by Leidy et al. (1998), and more briefly by Jacoby (1996) and Leppik (1998).

f) Quality of Life in Epilepsy-10 (QOLIE-10)

This was also developed, using the original dataset, from the 99 items used to develop the QOLIE-89 (Cramer et al., 1998). Items from the QOLIE-89 were selected for inclusion in the QOLIE-10 by an expert panel, which also identified seven domains considered to be important for epilepsy patients. The panel selected items with high item-scale correlations, and consistent or appropriate wording and sentence structure. The 10-item questionnaires covers general and epilepsy-specific areas, grouped into three factors: Epilepsy effects (memory, physical effects, mental effects of medication), Mental health (energy, depression, overall QoL), Role functioning (seizure worry, work, driving, social limits).There is some support for its reliability and validity.

g) Side-Effect and Life Satisfaction (SEALS) Inventory

The SEALS is a 50-item self completion questionnaire designed to measure satisfaction with anti-epileptic (AED) drug therapy. The original also contains an ADL sub-scale with items on frequency of daily activities, from household to social roles. A less diffuse, 38-item version is available (Gillham et al., 2000). The items relate to the patient's feelings and behaviour over the past week, and were grouped into five subscales, supported by factor analysis: General cognitive difficulties, Satisfaction/Dysphoria. Fatigue/Tiredness, Temper, and Worry, each with 4-point Likert frequency response scales. Answers are summed for each domain and for an overall score. The SEALS was designed by Brown and Tomlinson (1982) with 125 epilepsy patients and 79 people without epilepsy. Fatigue discriminated well between patients and non-patients.

h) Washington Psychosocial Seizure Inventory (WPSI)

The WPSI is the oldest instrument for the evaluation of psychosocial concerns in adults with epilepsy. It was not intended to cover broader health or QoL. The questions are anchored in actual performance in life, and assess adaptation and functioning. It has good reliability and validity when compared to clinical ratings (Dodrill et al., 1980).

The instrument was developed with a group of behavioural scientists, who compiled a list of categories of psychosocial problems they had encountered in this area. These were: Family background, Emotional adjustment, Interpersonal problems, Vocational adjustment, Financial status, Adjustment to seizures, Medicine and medical management, and overall psycho-social functioning. Item development and scaling was constructed next. After two piloting studies, 132 items with Yes/No responses

resulted. Professionals then rated 127 adults. Each subscale had to satisfy empirical requirements for inclusion. Inter-rater reliability, retest and internal consistency reliability were established, although patients were not consulted. The resulting scale has 132 items in three validity subscales and eight clinical subscales (Family background, Emotional adjustment, Interpersonal problems, Vocational adjustment, Financial status, Adjustment to seizures, Medicine and medical management, and overall psychosocial functioning). Later an item measuring QoL was added (Dodrill and Batzel, 1996). There are four profiles: 1. No problems; 2 Possible or slight difficulties; 3 Definite problems; 4 Severe or major problems. Higher scores indicate poorer adjustment. It is a lengthy instrument, taking 15-20 minutes to complete, using a trained interviewer.

The instrument has been reviewed by Hermann (1995), Jacoby (1996) and Selai and Trimble (1995). An overview of its development and widespread use was published by Dodrill and Batzel (1994), who reported 48 published papers on the WPSI.

Table 6: **Epilepsy-specific PROMs - content and scoring**

Instrument	*Domains (no. items)*		*Response options*	*Score*	*Administration/ Completion (time)*
Epilepsy Surgery Inventory-55 (ESI-55)	*5 sub-scales/55 items* 1. SF-36* 2. Cognitive function (5) 3. Role limitations (8) 4. Health perceptions (4) 5. Overall QoL (2) ** Includes the 7 sub- scales of Rand SF-36 as generic core*		Various, including 5 and 6-point scales, dichotomous responses and VAS	Three summary composite scores computed: mental functioning, physical functioning, role functioning	15 minutes to complete
Katz Adjustment Scales (adapted for epilepsy)	Original KAS-R: 12 sub-scales/76 of 127 potential items 1. General psychopathology (24) 2. Suspiciousness (4) 3. Anxiety (6) 4. Negativism (9) 5. Confusion (3) 6. Belligerence (4) 7. Withdrawal (5) 8. Bizarreness (5) 9. Hyperactivity (3) 10. Helplessness (4) 11. Verbal expansiveness (5) 12. Nervousness (4) Misc. (not used) (50)	*Revised by Vickrey et al. (1992b) for epilepsy*: 14 sub-scales/127 items 1.Oversensitivity/fearfulness (18) 2. Social (10) 3. Irritability (9) 4. Dependency (15) 5. Acting out (12) 6. Paranoia (5) 7. Abnormal thought process (5) 8. Withdrawal-R (11) 9. Emotional liability (6) 10. Nervousness-R (5) 11. Sociopathy (4) 12. Bizarreness-R (4) 13. Hyperactivity-R (4) 14. Disorientation (5) Misc (not used) (13)	4-point scales	No details found	

Instrument	*Domains (no. items)*	*Response options*	*Score*	*Administration/ Completion (time)*
Liverpool Quality of Life Batteries (LQOL Batteries) and Liverpool Seizure Severity Scale (LSSS)	LQOL: 1. Adverse drug effects Scale(21) 2. Impact of Epilepsy Scale (8) General: Affect-Balance Scale (10) 3. Mastery Scale (7) 4. Life Fulfillment Scale (20) *[Early versions included other existing psychological and health status scales]* LSSS: Seizure Severity Scale (2 sub-scales): Perceived control over seizures (9) Ictal and post-ictal symptoms (11)	4-point Likert; rating scales, dichotomous Present/Absent	Adverse drug effects summed, with higher scores indicating more problems. Impact: mean scores used; Affect-Balance: range 1-9 with higher scores indicating more positive balance; Mastery: range 7-28, with higher scores indicating greater mastery; Fulfilment: compute of difference between actual and ideal life satisfaction scores.	30-45 minutes to complete
Quality of Life in Epilepsy-89 (QOLE-89)	*17 subscales[including the 7 subscales of SF-36]/89 items [SF-36 supplemented by 53 items specific to epilepsy] grouped into four factors:* 1.Seizure-specific effects (seizure worry, health discouragement, medicine effects, work or driving or social function) 2. Cognition (language, memory, attention) 3. Physical health (role limitations/physical pain, health perceptions, or physical function) 4. Mental health (overall quality of life, emotional well-being, role limitations/emotional, social isolation, social support, and energy or fatigue) *[Original 99 items reduced to 87, two items added on overall health perception and sexual functioning to produce 89 items.* *]*	4, & 6 point Likert scales; dichotomous Yes/No; 0-10 QoL rating scale, 1-5 point VAS scale, 0-100 VAS scale	Summation and domain score (weighted) Overall score = weighted sum of subscale scores Subscale scores = mean scores across items within the subscale Index:0 = worst QoL, 100 = best QoL Higher scores represent better function on all scales	Self-report 28.4 (±15.6) minutes to complete
Quality of Life in Epilepsy-31 (QOLIE-31)	*7 sub-scales/31 items* 1. Seizure worry (5) 2. Overall QoL (2) 3. Emotional well-being (5) 4. Energy/Fatigue (4) 5. Cognitive functioning (6) 6. Medication effects (3) 7. Social functioning (5) Health status (1) [not included in total score]	5 point Likert; 0-10 QoL rating scale	Summation and domain score (weighted) Higher scores represent better function on all scales Index:0 = worst QoL, 100 = best QoL Higher scores represent better function on all scales	Self-report 15 minutes to complete

Quality of Life in Epilepsy-10 (QOLIE-10)	*7 sub-scales/10 items* 1. Seizure worry(1) 2. Overall QoL (1) 3. Emotional well-being (1) 4. Energy/fatigue (1) 5. Cognitive functioning (1) 6. Medication effects (2) 7. Social function (3)	5 point Likert; 0-10 QoL rating scale	Summation and domain score (weighted) Higher scores represent better function on all scales Index:0 = worst QoL, 100 = best QoL Higher scores represent better function on all scales	Self-report 'few minutes to complete'
Side-Effect and Life Satisfaction	5 sub-scales/ 50 *and shorter form versions; Gillam et al.,(1996) standardisation:* 1. General cognitive difficulties (17) 2. Satisfaction/Dysphoria (8) 3. Fatigue/Tiredness (5) 4. Temper(4) 5. Worry (4) *[early version included frequency of ADL)*	4-point Likert scales	Summation and domain scores	
Washington Psychosocial Seizure Inventory (WPSI)	*7 sub-scales/132 items* 1. Family background (11) 2. Emotional adjustment (34) 3. Interpersonal adjustment (22) 4. Vocational adjustment (13) 5. Financial status (7) 5. Adjustment to seizures (15) 6. Medicine/medical management (8) 7. Overall psychosocial functioning (57) *Later an item measuring QoL was added*	Dichotomous Yes/No	Summation and domain score Higher scores indicate poorer adjustment. There are four profiles: 1. No problems; 2 Possible or slight difficulties; 3 Definite problems; 4 Severe or major problems.	Interviewer 15-20 minutes to complete

Table 7: **Epilepsy-specific PROMs - health status domains** *(after Fitzpatrick et al., 1998)*

Instrument	*Instrument domains*								
	Physical function	**Symptoms**	**Global judgement of health**	**Psychological well-being**	**Social well-being**	**Cognitive functioning**	**Role activities**	**Personal constructs***	**Treatment satisfaction**
Epilepsy Surgery Inventory-55 (ESI-55)	x	x	x	x	x	x	x	x	
Katz Adjustment Scale				x		x			
Liverpool Quality of Life Batteries (LQoL batteries)		x	x	x	x		x	x	
Quality of Life in Epilepsy-89 (QOLIE-89)	x	x	x	x	x	x	x	x	
Quality of Life in Epilepsy-31 (QOLIE-31)		x	x	x	x	x	x	x	
Quality of Life in Epilepsy-10 (QOLIE-10)		x		x	x	x	x	x	
Side-Effect and Life Satisfaction (SEALS)	x	x		x		x	x	x	
Washington Psychosocial Seizure Inventory (WPSI)				x	x		x	x	x

*** Includes global (HR) QoL ratings**

APPENDIX E: Availability and licensing details for short-listed PROMs

PROMs	Country	Licensing Fee, etc.	Contact
SF-36	USA	*QualityMetric, the Medical Outcomes Trust, and the Health Assessment Lab are the co-copyright and trademark holders of the SF-36.* A commercial license is required. Permission for use is required for scholarly research. The shared licensing program allows individuals and organisations that benefit from commercial uses of the intellectual property to pay royalties or other user fees that will support the research community that made the original surveys possible.	Hospitals, Health Systems, and Clinicians can contact Victoria Hall: (800) 572-9394 ext. 253. The licensing request form can be found at http://www.qualitymetric.com/
ESI-55	Canada	Copyright by UCLA. Permission is granted for use for non-commercial purposes. Free at questionnaire and scoring manual: http://www.rand.org/health/surveys_tools/esi/index.html	*Contracts Office, RAND, 1700 Main Street, PO Box 2138, Santa Monica, CA 90407-2138, USA* *Barbara Vickrey* UCLA Department of Neurology, C-128 RNRC, 710 Westwood Plaza, Los Angeles, CA 90024-1769, USA Voice 310/206-7671, Fax 310/794-7716
QOLIE-89	USA	*No charge is made for academic and government use; a charge is made for commercial use.* *Free questionnaire and scoring manual:* http://www.rand.org/health/surveys_tools/qolie/	joyce.cramer@yale.edu
QOLIE-31	USA	*No charge is made for academic and government use; a charge is made for commercial use.* *Free questionnaire and scoring manual:* http://www.rand.org/health/surveys_tools/qolie/	joyce.cramer@yale.edu

APPENDIX F: Methods of working, membership, and conclusions of the multidisciplinary panel

Members of the multidisciplinary panel were invited to participate based on their clinical or research experience of epilepsy and special interest in patient-reported outcome measures.

The panel were sent the following documents:

- A structured review of patient-reported outcome measures for people with epilepsy: an update 2009;
- A structured review of patient-reported health instruments for people with epilepsy, 2006;
- copies of the PROMs short-listed for discussion.

The panel were sent by e-mail rating scales to judge the suitability of the questionnaire for use in the NHS for the evaluation of services. There was a section for comments. The rating scale used the following responses:

- 'not at all suitable' (score 0)
- 'to some extent unsuitable' (score 1)
- 'uncertain' (score 2)
- 'to some extent suitable' (score 3)
- 'very suitable' (score 4)

Scores for each questionnaire were ranked in order of preference. The total maximum score=36.

The results and comments were then distributed by e-mail to the panel for further rating should they wish to change their vote.

Notes of electronic discussion March-April 2009

Generic measures

SF-36

As with other generic measures, use of the SF-36 allows comparison between the HRQoL outcomes of people with epilepsy and those of people with other long-term conditions. The SF-36 has been shown to be robust and accurate with a wide range of clinical groups; it is norm-referenced and very widely used. The measure is relatively brief and straightforward for patients to complete (although one panel member considered it over-complex), and appears acceptable to patients. It is felt to have good content validity with respect to important psychological and social impacts of epilepsy, without being intrusive. If a utility estimate is required, SF-36 data can be converted using the SF-6D.

Criticisms include floor and ceiling effects, and some doubt as to content validity. Not all questions are relevant or meaningful to people with epilepsy, and there may be an over-emphasis on physical function. Questions on daily activities do not capture the full impact of epilepsy in the majority of cases where patients have adapted to living with epilepsy. The wording of some questions admits of different meanings, leading to problems of interpretation; there are also language issues (Americanisms). Recalling different time-periods (a year ago, the past four weeks, a typical day) may

be difficult for those with memory problems, which commonly affect people with epilepsy.

EQ-5D

Advantages of the EQ-5D include its brevity, although this was also the main criticism of the measure. It was suggested the EQ-5D may be more appropriate for assessing co-morbidities than for epilepsy itself. One panel member, who strongly favours the EQ-5D, suggested the recently released 5-level (as opposed to the standard 3-level) version be considered, as it appears to have greater discriminative ability.

Disadvantages were the brevity of the instrument and a restricted range of responses, which limit its sensitivity. Some questions, e.g. mobility and usual activities, are not well-phrased for people with epilepsy; patients may report 'no problems' with these domains even though recalcitrant epilepsy can prevent, or have an adverse impact on, many activities. Asking the respondent to describe their health state 'today' may be inappropriate for people with epilepsy, as their responses will vary depending on whether they have had a seizure that day.

The evidence base for use of the EQ-5D in epilepsy appears weak, with several domains in the appraisal of psychometric and operational performance for which results are either not reported, or there is no evidence in favour of the questionnaire.

Suitability of the EQ-5D depends upon whether there is a need to estimate utilities for economic evaluations.

Generic total

FIRST RATING	**'not at all suitable' (score 0)**	**'to some extent unsuitable' (score 1)**	**'uncertain' (score 2)**	**'to some extent suitable' (score 3)**	**'very suitable' (score 4)**	**TOTAL**
SF-36	-	1	2	12	12	**27**
EQ-5D	-	3	4	9	4	**20**

Epilepsy-specific PROMs

ESI-55

It was felt that this measure may be useful in assessing the potential impact of epilepsy surgery, the purpose for which it was originally intended. However, it is possibly less relevant outside the context of surgery. One panel member with experience of using the measure commented on its good responsiveness and sensitivity to change.

Based as it is on the generic SF-36, the ESI-55 contains relatively few epilepsy-specific questions. One panel member felt there was insufficient emphasis on the important cognitive and social effects of epilepsy. Others, however, found the questions on memory, concentration, and functional language particularly relevant.

As with the SF-36, using different time-frames for question evaluation may pose difficulties for people with memory impairments, and may be an explanation for mixed findings regarding responsiveness.

QOLIE-89

Although the length of this instrument is an obvious drawback, its strength is in the broad range of its measurement, comprehensiveness, and detail. One panel member with experience of its use reported that length does not appear to make it less acceptable to people with epilepsy, probably because of its relevance. It is argued to be the best of the questionnaires shortlisted, both from the point of view of gathering data about the delivery of services and for the evaluation of clinical treatments.

The scale is overall very robust, with a strong psychometric basis. It has good convergent and discriminant validity, and responsiveness/sensitivity to change. The measure is useful for studying changes over a relatively short time-frame, as well as over the longer term; this is helpful and relevant to the epilepsy population.

One panel member suggests the QOLIE-89 has the same limitations as the SF-36 on which it is based – i.e. an over-emphasis on physical items. However, others commend the inclusion of questions on memory, language, concentration, and problems such as impact of seizures and side-effects of medication, which are highly relevant to people with epilepsy.

Disadvantages of the instrument are its length and complexity. It is time-consuming to complete which may inhibit respondents and affect response rates; there are also implications for some people with epilepsy who, due to medication or other reasons, find it difficult to concentrate for extended periods. On the other hand, one panel member suggests the number of items may help to overcome some interpretive problems.

QOLIE-31

The shorter version of the QOLIE has many of the valuable features of its ‘parent’ instrument, and is more focused and less intimidating to respondents. The length of the QOLIE-31 makes it more suitable for repeated use, with less patient burden.

The QOLIE-31 emphasises the psychological aspect of epilepsy. It also covers aspects which people with epilepsy consider highly relevant, namely cognition (memory and concentration), seizures, and medication. It is suggested that, given its psychological focus, the measure should be used with caution in the national measurement of qualities and outcomes, and may best be used in conjunction with other tools

The scale has been extensively tested and is overall robust (though perhaps less so than the longer measure), with good construct and convergent validity. However, evidence for discriminant validity is less strong, and there is mixed evidence for responsiveness to change.

Although the QOLIE-31 includes most of the health status domains addressed by longer instruments, it lacks items on physical function and could miss some adverse effects of medication because of this.

Specific Total

FIRST RATING	'not at all suitable' (score 0)	'to some extent unsuitable' (score 1)	'uncertain' (score 2)	'to some extent suitable' (score 3)	'very suitable' (score 4)	TOTAL
QOLIE-89	-	1	-	9	20	**30**
QOLIE-31	-	-	2	12	16	**30**
ESI-55	-	-	2	18	8	**28**

General points

Panel members commented that a number of issues of importance to people with epilepsy are not adequately covered by the measures presented for consideration. These include the impact of seizures, and/or the prospect of seizure occurrence, on everyday life; the physical, cognitive, and emotional side-effects, and potential long-term impact, of medication; the often troublesome experience of aura; feelings of stigma; and discrimination on the part of others. Patient satisfaction with treatment is also not addressed directly.

Problems with 'Americanised' language apply to all of the measures considered, with the exception of the EQ-5D, giving rise to potential problems of misunderstanding or misinterpretation. All three of the condition-specific measures shortlisted draw on the US version of the SF-36; it is suggested that changes made to the UK version of the SF-36 be incorporated into whichever epilepsy-specific measure is adopted, preferably along with a full UK cultural adaptation.

All three specific measures invite the respondent to write explanatory comments if they are unsure about how to answer a question. One panel member commended this aspect as, although it may introduce subjectivity and possible bias on the part of the researcher, it allows people with epilepsy to qualify their answers and enriches the information gathered.

Recommendations

Based on ratings and comments from the panel, the SF-36 is preferred to the EQ-5D as a generic measure of health status.

For the measurement of epilepsy-specific quality of life, the QOLIE-89 and QOLIE-31 are rated equally strongly.

Given that the QOLIE-89 incorporates the SF-36, this would be the instrument of choice if a single instrument is required.

If both a generic and a specific instrument are to be used, the QOLIE-31 is the recommended specific instrument.

Patient-reported Outcome Measure Rating Scale

1. On the basis of the review of evidence and your personal experience, is this questionnaire suitable for the measurement of the quality and outcomes of services for people with epilepsy? (please tick one box)

☐ Not at all suitable ☐ To some extent unsuitable ☐ Uncertain ☐ To some extent suitable ☐Very suitable

Do you have another questionnaire you could suggest?

Any additional comments

EPILEPSY OUTCOMES CONSENSUS GROUP MEMBERS.

Dr Laura Goldstein
Professor of Clinical Neuropsychology, MRC Centre in Neurodegeneration Research London

Mel Goodwin
Epilepsy Nurse Specialist, Neurology Department, Northampton

Dr. Yvonne Hart
Department of Neurology, John Radcliffe Hospital (West Wing), Oxford

Andrew Lloyd
Director, Oxford Outcomes, Oxford

Graham Scambler
Professor of Medical Sociology, UCL Medical School, London

Dr Caroline Selai
Senior Lecturer in Clinical Neuroscience, University College London - Institute of Neurology, London

Dr Henry Smithson
Academic Unit of Primary Medical Care, University of Sheffield, Sheffield

Phil Tittensor
Clinical Nurse Specialist-Epilepsy, Department of Neurology, Mid-Staffordshire NHS Foundation Trust, Stafford

Christine Wheeler
Patient representative, Northampton

PROM Group members

Ray Fitzpatrick
Professor of Public Health and Primary Care, Deputy Head, Department of Public Health, University of Oxford.

Elizabeth Gibbons
Senior Research Officer, Department of Public Health, University of Oxford.

Anne Mackintosh
Research Officer, Department of Public Health, University of Oxford.

REFERENCES

Bautista RED, Glen ET, Shetty NK. Factors associated with satisfaction with care among patients with epilepsy. *Epilepsy and Behavior* 2007; 11: 518-524.

Bautista RED, Glen ET, Wludyka PS, Shetty NK. Factors associated utilization of healthcare resources among epilepsy patients. *Epilepsy Research* 2008; 79: 120-129.

Beran RG, Berkovic S, Black A. Australian Study of Titration to Effect Profile of Safety (AUS-STEPS): high-dose gabapentin (Neurontin) in partial seizures. *Epilepsia* 2001; 42: 1335–1339.

Bergner M, Bobbitt RA, Kressel S, et al. The Sickness Impact Profile: conceptual formulation and methodology for the development of a health status measure. *International Journal of Health Service* 1976; 6(3):393-415.

Bergner M, Bobbitt RA, Carter WB et al. The Sickness Impact Profile: development and final revision of a health status measure. *Medical Care* 1981; 19(8):787-805.

Brazier JE, Jones N, Kind P. Testing the validity of the EuroQol and comparing it with the SF-36 Health Survey questionnaire. *Quality of Life Research* 1993; 2(3):169-180.

Brazier JE, Roberts j, Deverill M: The estimation of a preference-based measure of health from the SF-36. *Journal of Health Economics* 2002; 21: 271-292.

Brown SW, Tomlinson LL. Anticonvulsant side-effects: a self report questionnaire for use in community surveys. *British Journal of Clinical Practice* 1982; 18 - symposium supplement: 147-9.

Bruni J. Outcome evaluation of gabapentin as add-on therapy for partial seizures. *Canadian Journal of Neurological Sciences* 1998; 25:134–40.

Bruni J. Gabapentin as adjunctive therapy for partial seizures. *Epilepsia* 1999; 40(6): 27-28.

Buelow JM, Ferrans CE. Quality of life in epilepsy. In Kanner AM, ed. *Psychiatric issues in epilepsy: a practical guide to diagnosis and treatment. 400 pp*, pp 307-18. Philadelphia, PA, USA: Lippincott Williams & Wilkins Publishers, 2001.

Cowan J and Baker GA. A review of subjective impact measures for use with children and adolescents with epilepsy. *Quality of Life Research* 2004; 13: 1435-1443.

Cramer JA, Perrine KR, Devinsky O, Bryant-Comstock L, Meador K, Hermann BP. Development and cross-cultural translations of a 31-item Quality of Life in Epilepsy Inventory. *Epilepsia* 1998; 39: 81-8.

Cramer JA, Westbrook LE, Devinsky O, Perrine K, Glassman MB, and Camfield C. Development of the quality of life in epilepsy inventory for adolescents: The QOLIE-AD-48. *Epilepsia* 1999; 40(8):1114-1121.

Cramer JA. Exploration of changes in health-related quality of life after 3 months of vagus nerve stimulation. *Epilepsy and Behavior* 2001; 2: 460-465.

Cramer JA, Brandenburg N, Xu X. Differentiating anxiety and depression symptoms in patients with partial epilepsy. *Epilepsy and Behavior* 2005; 6: 563-569.

Darzi A. Our NHS Our Future: NHS Next Stage Review Interim Report. Department of Health, London, October 2007.

Darzi A. Our NHS Our Future: High Quality Care for All. NHS Next Stage Review Final Report. Department of Health, London, June 2008.

De Bruin AF, De Witte LP, Stvene F, Diederiks JPM. Sickness Impact Profile: the state of art of a generic functional status measure. *Social Science and Medicine* 1992*;* 35: 1003-14.

Department of Health, London, December 2008. Guidance on the routine collection of Patient Reported Outcome Measures (PROMs) for the NHS in England 2009/10.

Devinsky O,Vickrey BG. Quality of life: recent developments in the USA. In Trimble MR, Dodson WE, eds. *Epilepsy and quality of life,* New York, NY, USA: Raven Press, 1994.

Devinsky O, Vickrey BG, Cramer JA, Perrine KR, Hermann BP, Meador K *et al*. Development of the Quality of Life in Epilepsy Inventory. *Epilepsia* 1995; 36: 1089-104.

Devinsky O, *et al*. Risk factors for poor health-related quality of life in adolescents with epilepsy. *Epilepsia* 1999; 40(12):1715-1720.

Dodrill CB, Batzel LW, Queisser HR, Temkin NR. An objective method for the assessment of psychological and social problems among epileptics. *Epilepsia* 1980; 21: 123-35.

Dodrill CB, Batzel LW. The Washington Psychosocial Seizure Inventory and quality of life in epilepsy. In Trimble MR, Dodson WE, eds. *Epilepsy and quality of life*, New York, NY, USA: Raven Press, 1994.

Dodrill CB, Batzel LW. Epilepsy. *Archives of Neurology* 1996; 53: 476-7.

Eiser C, Morse R. A review of measures of quality of life for children with chronic illness. *Archives of Disease in Childhood* 2001; 84: 205-211.

EuroQol Group, The. EuroQol: a new facility for the measurement of health-related quality of life. *Health Policy* 1990; 16(3):199-208.

Feeny D, Furlong W, Boyle M, Torrance GW. Multi-Attribute Health Status Classification Systems: Health Utilities Index. *PharmacoEconomics* 1995; 7(6):490-502.

Ficker DM, Privitera M, Krauss G, Kanner A, Moore JL, Glauser T. Improved tolerability and efficacy in epilepsy patients with extended-release carbamazepine. *Neurology* 2005; 65: 593-595.

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

Fitzpatrick R, Bowling A, Gibbons E, Haywood K, Jenkinson C, Mackintosh A, Peters M. A Structured Review of Patient-Reported Measures in Relation to Selected Chronic Conditions, Perceptions of Quality of Care, and Carer Impact. Oxford: National Centre for Health Outcomes Development, November 2006.

Garratt AM, Schmidt L, Mackintosh A, Fitzpatrick R. (2002) Quality of life measurement: bibliographic study of patient assessed health outcome measures. *British Medical Journal;* 324(7351):1417-1421.

Gillham R, Bryant-Comstock L, Kane K. Validation of the Side-Effect and Life Satisfaction/SEALS inventory. *Seizure* 2000; 9: 458-63.

Gilliam FG, Barry JJ, Hermann BP, Meador KJ, Vahle V, Kanner AM. Rapid detection of major depression in epilepsy: a multicentre study. *Lancet Neurol* 2006; 5: 399-405.

Hamberger MJ, Palmese CA, Scarmeas N, Weintraub D, Choi H, Hirsch LJ. A randomized, double-blind, placebo-controlled trial of donepezil to improve memory in epilepsy. *Epilepsia* 2007; 48(7):1283-1291.

Harden CL, Maroof DA, Nikolov B, Fowler K, Sperling M, Liporace J, Pennell P, Labar D, Herzog A. The effect of seizure severity on quality of life in epilepsy. *Epilepsy and Behavior* 2007; 11: 208-211.

Hermann BP. The evolution of health-related quality of life assessment in epilepsy. *Quality of Life Research* 1995; 4: 87-100.

Hunt SM, McKenna SP, McEwen J, Backett EM, Williams J, Papp E. A quantitative approach to perceived health status: a validation study. *Journal of Epidemiology and Community Health* 1980; 34: 281-286.

Jacoby A. Assessing quality of life in patients with epilepsy. *Pharmacoeconomics* 1996; 9: 399-416.

Jenkinson C, McGee H. Health Status Measurement, a Brief but Critical Introduction. Oxford: Radcliffe Medical Press Ltd, 1998.

Katz MM, Lyerly SB. Methods for measuring adjustment and social behavior in the community: I. Rationale, description, discriminative validity and scale development. *Psychological Reports* 1963; 13: 503-35.

Kustra R, Meador K, Evans B, Leschek-Gelman L, Groenke D, Hammer A, Nanry K, Messenheimer J. Lamotrigine therapy in patients requiring a change in antiepileptic drug regimen. Seizure 2005; 14(4): 254-261.

Langfitt JT, et al. Validity & responsiveness of generic preference-based HRQOL instruments in chronic epilepsy. *Quality of Life Research* 2006; 15(5):899-914.

Leidy NK, Rentz AM, Grace EM. Evaluating health-related quality of life outcomes in clinical trials of anti-epileptic drug therapy. *Epilepsia* 1998; 39:965-77.

Leppik IE. Quality of life of people with epilepsy in the United States. *Clinical Therapeutics* 1998; 20: A13-A18.

Levi R, Drotar D. Critical issues and needs of health-related quality of life assessment in children and adolescents with chronic health conditions. In Drotar D (ed.), *Measuring Health-related Quality of Life in Children and Adolescents*, Erlbaum: NJ, 1998; 3-24.

McDowell I, Newell C. Measuring Health: a guide to rating scales and questionnaires. New York, USA: Oxford University Press, 1996.

McGlone J, Valdivia I, Penner M, Williams J, Sadler RM, Clarke DB. Quality of life and memory after vagus nerve stimulator implantation for epilepsy. *Canadian Journal of Neurological Sciences* 2008; 35: 287-297.

Patrick D, Peach H (eds). Disablement in the Community: A Sociomedical Press Perspective. Oxford: Oxford University Press, 1989.

Pettit T, Livingston G, Manela M, Kitchen G, Katona C, Bowling A. Validation and normative data of health status measures in older people: the Islington study. *International Journal of Geriatric Psychiatry* 2001; 16: 1061-1070.

Post MWM, de Bruin AF, de Witte et al. The SIP68: a measure of health related functional status in rehabilitation medicine. *Archives of Physical Medicine and Rehabilitation* 1996; 77(5):440-5.

Pramuka M, Hendrickson R, Zinski A, Van Cott AC. A psychosocial self-management program for epilepsy: A randomized pilot study in adults. *Epilepsy and Behavior* 2007; 11: 533-45.

Resnick B, Nahm ES. Reliability and validity testing of the revised 12-item Short-Form Health Survey in older adults. *Journal of Nursing Measurement* 2001; 9(2): 151-61.

Resnick B, Parker R. Simplified scoring and psychometrics of the revised 12-item Short-Form Health Survey. *Outcomes Management for Nursing Practice* 2001; 5(4): 161-6.

Reynders HJ, Broks P, Dickson JM, Lee CE, Turpin G. Investigation of social and emotion information processing in temporal lobe epilepsy with ictal fear. *Epilepsy and Behavior* 2005; 7: 419-29.

Ronen GM et al. Health-related quality of life in childhood epilepsy: The result of children's participation in identifying the components. *Developmental Medicine and Child Neurology* 1999; 41(8): 554-559.

Ronen GM et al. Health-related quality of life in childhood disorders: A modified focus group technique to involve children. *Quality of Life Research* 2001; 10(1): 71-79.

Ronen GM, Streiner DL, Rosenbaum PL. Health-related quality of life in children with epilepsy: development and validation of self-report and parent proxy measures. *Epilepsia* 2003; 44(4):598-612.

Rose KJ, Derry PA, Wiebe S, McLachian RS. Determinants of health-related quality of life after temporal lobe epilepsy surgery, *Quality of Life Research* 1996; 5: 395–402.

Sachedo RC, Gates JR, Bazil JR, Barkley GL, Tatum W, D'Souza J, McCague K. Improved quality of life in patients with partial seizures after conversion to oxcarbazepine monotherapy. *Epilepsy and Behavior* 2006; 9: 457-463.

Santhouse J, Carrier C, Arya S, Fowler H, Duncan S. A comparison of self-reported quality of life between patients with epilepsy and neurocardiogenic syncope. *Epilepsia* 2007; 48(5):1019-22.

Schofield M, Mishra G. Validity of the SF-12 compared with the SF-36 health survey in pilot studies of the Australian longitudinal study on women's health. *Journal of Health Psychology* 1998; 3: 259-271.

Selai CE, Trimble MR. Quality of life assessment in epilepsy: the state of the art. *Journal of Epilepsy* 1995; 8: 332-7.

Selai C, Trimble M, Price M. Evaluation of health status in epilepsy using the EQ-5D questionnaire: a prospective, observational, 6-month study of adjunctive therapy with anti-epileptic drugs. *Current Medical Research and Opinion* 2005; 21(5):733-740.

Smith SC, Cano S, Lamping DL, Staniszewska S, Browne J, Lewsey J, van der Meulen J, Cairns J, Black N. Patient-reported outcome measures (PROMs) for routine use in treatment centres: recommendations based on a review of the scientific evidence. *London School of Hygiene and Tropic Medicine, London, UK, and the Department of Health*, 2004.

Smith S, Cano S, Lamping D, et al. Patient reported outcome measures (PROMs) for routine use in treatment centres: recommendations based on a review of the scientific literature. Final Report to Department of Health, 2005.

Somerville ER, McLaughlin DB, Robinson MK, Berkovic SF. Adjunctive therapy of uncontrolled partial seizures with levetiracetam in Australian patients. *Epilepsy and Behavior* 2007; 11: 338-342.

Spencer SS, Berg AT, Vickrey BG, Sperling MR, Bazil CW, Haut S, Langfitt JT, Walczak TS, Devinsky O. Health-related quality of life over time since respective epilepsy surgery. *Annals of Neurology* 2007; 62: 327-334.

Szaflarski M, Meckler JM, Privitera MD, Szaflarski JP. Quality of life in medication-resistant epilepsy: The effects of patient's age, age at seizure onset, and disease duration. *Epilepsy and Behavior* 2006; 8: 547-551.

Testa SM, Schefft BK, Szaflarski JP, Yeh H-S, Privitera MD. Mood, personality, and health-related quality of life in epileptic and psychogenic seizure disorders. *Epilepsia* 2007; 48(5): 973-982.

Tracy JI, Dechant V, Sperling MR, Cho R, Glosser D. The association of mood with quality of life ratings in epilepsy. *Neurology* 2007; 68: 1101-1107.

Trimble MR, Dodson WE eds. Epilepsy and quality of life. New York, NY, USA: Raven Press, 1994.

Vickrey BG, Hays RD, Graber J, Rausch R, Engel JJ, Brook RH. A health-related quality of life instrument for patients evaluated for epilepsy surgery. *Medical Care* 1992a; 30: 299-319.

Vickrey BG, Hays RD, Brook RH, Rausch R. Reliability and validity of the Katz Adjustment Scales in an epilepsy sample. *Quality of Life Research* 1992b; 1: 63-72.

Vickrey BG, Hays RD, Engel JJ, Spritzer KL, Rogers WH, Rausch R *et al*. Outcome assessment for epilepsy surgery: the impact of measuring health-related quality of life. *Annals of Neurology* 1995; 37: 158-66.

Ware JE, Sherbourne CD. The MOS 36-Item Short-Form Health Survey (SF-36). I. Conceptual framework and item selection. *Medical Care* 1992; 30(6):473-483.

Ware J, Kosinski M, Keller SD. SF-36 Physical and Mental Health Summary Scales: A User's Manual. Boston, MA: The Health Institute, New England Medical Centre, 1994.

Ware JE, Kosinski M, Keller SD. SF-12: How to Score the SF-12 Physical and Mental Health Summary Scales. Boston, MA: The Health Institute, New England Medical Centre, 1995. Second Edition.

Ware JE. SF-36 Health Survey. Manual and Interpretation Guide. Boston, MA: The Health Institute, New England Medical Centre, 1997. Second Edition.

Ware JE, Gandek B. Methods for testing data quality, scaling assumptions and reliability: the IQOLA project approach. *Journal of Clinical Epidemiology* 1998; 51(11):945-952.

WHOQOL Group, The. The World Health Organisation Quality of Life Assessment (WHOQOL). Development and general psychometric properties. *Social Science & Medicine* 1998; 46: 1569-85

Xu X, Brandenburg NA, McDermott AM, Bazil CW. Sleep disturbances reported by refractory partial-onset epilepsy patients receiving polytherapy. *Epilepsia* 2006; 47(7): 1176-11