Neuro-QOL and the NIH Toolbox: implications for epilepsy

The impact of neurological disorders on the lives of patients is often far more complex than what is measured in routine examination. Measurement of this impact can be challenging owing to a lack of brief, psychometrically sound and generally accepted instruments. Two NIH-funded initiatives are developing assessment tools, in English and Spanish, which address these issues, and should prove useful to the study and treatment of epilepsy and other neurological conditions. The first, Neuro-QOL, has created a set of health-related quality of life measures that are applicable for people with common neurological disorders. The second, the NIH Toolbox for the Assessment of Neurological and Behavioral Function, is assembling measures of cognitive, emotional, motor and sensory health and function that can be used across all ages, from 3 to 85 years. This article describes both the projects and their potential value to epilepsy treatment and research.

Rationale for health-related quality of life measurement

The assessment of health-related quality of life (HRQL) is particularly important in chronic diseases, including many neurological conditions, where curative therapies may not be available, and the focus of care is on minimizing the negative impact of the disease and its treatment. This impact can be on any aspect of cognitive, physical, social, sensory or emotional functioning. Epilepsy, for example, has life consequences that extend well beyond the experience of having seizures. The unpredictability of when a seizure occurs is one of the main causes of disability among people with epilepsy. It limits the ability of the individual to drive a motorized vehicle, pilot aircraft, work at heights and many more occupations. Seizures also disrupt hormonal circadian cycles, leading to decreased fertility in women, decreased arousal in both genders and other sexual dysfunctions. In addition to seizures, many patients suffer from the side effects of medication, such as sedation, nausea, weight change, double vision, tremor, cognition and memory problems[1]. Surgical interventions to control intractable seizure activity have additional risks[2].

Epilepsy is also associated with missed school days and academic underachievement in children, increased risk of learning and behavior problems, un- or under-employment in adults, restricted social activity, stigma, anxiety and depression for children and adults[3–4]. These outcomes are not well captured in traditional clinical measures such as neurological examination and laboratory or radiological results. Therefore, measures that complement traditional assessments, and directly elicit a patient or caregiver report regarding their experience of the disease and treatment are necessary in order to gain a more complete understanding of how patients are being affected.

There are other existing HRQL assessment tools appropriate for use with people who have epilepsy. Generic (i.e., those suitable for use across different clinical and healthy populations) instruments include the Medical Outcomes Study Short Form-36 for adults[5], and the Child Health Questionnaire for children[6]. Some well-validated epilepsy-specific measures for adults are the Well-Being Scale[7], the Liverpool Quality of Life Battery[8], the Quality of Life in Epilepsy Inventory[9] and the Quality of Life in Epilepsy Instruments (86, 31 and 10 item)[10–11]. For children or their parent/proxys, instruments include the Quality of Life in Pediatric Epilepsy[12], the Quality of Life in Childhood Epilepsy Questionnaire[13], the HRQL measure for Children with Epilepsy[14], the Impact of Childhood Illness Scale[15] and the Impact of Childhood Neurologic Disability Scale[16]. Adolescent self-report measures include the Adolescent Psychosocial Seizure Inventory[17] and the Quality of Life in Epilepsy Instrument–Adolescents[18]. While a review of

**KEYWORDS:** behavioral function, epilepsy, health-related quality of life, patient reported outcomes, quality of life
these instruments is beyond the scope of this article, each has strengths and weaknesses, which may make it more or less useful depending on the goals of assessment (see [18,19] for reviews of pediatric and adolescent measures and [20] for a review of adult measures). For example, each of these HRQL assessment tools was developed to measure particular aspects of quality of life, with content often guided by the theoretical focus of the developers and the method of generating (e.g., through the involvement of experts, patients and families) domains and items [21]. Thus, when choosing a HRQL instrument, it is important to review the areas assessed to help ensure that they are consistent with the purpose of the assessment. Despite the availability of these HRQL assessment tools, there remains a need to develop a broad instrument that evaluates multiple dimensions of health and quality of life across the spectrum of neurological disorders. At the same time, it is critical that the ideal HRQL tool is brief enough for its routine utilization in the neurological clinic. In this article, we describe this tool and its application to the care of people with epilepsy.

Neuro-QOL

Neuro-QOL [101] is a multisite, multiyear project funded by the National Institute of Neurological Disorders and Stroke (NINDS). The project, which began in 2004, has developed a set of instruments, in both English and Spanish, that can be used to assess the HRQL of adults and children with common neurological diseases. The Neuro-QOL system is comprised of measures that evaluate concerns common to many or most neurological conditions, as well as ones that assess symptoms and concerns specific to only some neurological diseases (Table 1). This allows for both cross-disease comparisons and more in-depth assessment of disease-specific effects. This initial set of assessment tools has been validated for use with adults suffering from epilepsy, Parkinson’s disease, amyotrophic lateral sclerosis, stroke or multiple sclerosis, and children with epilepsy or muscular dystrophy. The instruments will become publicly available in October 2010.

### Development of the Neuro-QOL system

From the outset, the Neuro-QOL team actively solicited input from experts, patients and caregivers through surveys (n = 89), interviews (107 with experts and 63 with patients) and focus groups (seven with patients and four with caregivers). This input was used to identify the characteristics that potential end-users wanted to see in an HRQL measure and to select those areas of HRQL (domains and subdomains) most important for the assessment of people with the targeted diseases. After the domains and subdomains were identified, the next step was to assemble sets of questions (item pools) that comprehensively assessed each area. This was accomplished through an iterative process, in which pre-existing or newly developed items were reviewed in multiple stages by experts, patients and caregivers for content coverage, readability, translatability, relevance and other criteria, and modified at each step as necessary to meet those criteria. After this process was complete, these item pools underwent the first wave of field testing.

<table>
<thead>
<tr>
<th>Domain</th>
<th>Adult</th>
<th>Pediatric</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical function</td>
<td>Lower Extremity (Mobility)</td>
<td>Lower Extremity (Mobility)</td>
</tr>
<tr>
<td></td>
<td>Upper Extremity (Fine Motor, ADLs)</td>
<td>Upper Extremity (Fine Motor, ADLs)</td>
</tr>
<tr>
<td></td>
<td>Sleep Disturbance</td>
<td>Pain</td>
</tr>
<tr>
<td></td>
<td>Fatigue</td>
<td>Fatigue</td>
</tr>
<tr>
<td>Mental function</td>
<td>Depression</td>
<td>Depression</td>
</tr>
<tr>
<td></td>
<td>Anxiety</td>
<td>Anxiety</td>
</tr>
<tr>
<td></td>
<td>Stigma</td>
<td>Stigma</td>
</tr>
<tr>
<td></td>
<td>Positive Affect and Well-being</td>
<td>Anger</td>
</tr>
<tr>
<td></td>
<td>Applied Cognition – General Concerns</td>
<td>Applied Cognition – General Concerns</td>
</tr>
<tr>
<td></td>
<td>Applied Cognition – Executive Function</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Emotional and Behavioral Dyscontrol</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Communication Difficulty</td>
<td></td>
</tr>
<tr>
<td>Social function</td>
<td>Ability to Participate in Social Roles and Activities</td>
<td>Social Relations – Interactions with Peers and Adults</td>
</tr>
<tr>
<td></td>
<td>Satisfaction with Social Roles and Activities</td>
<td></td>
</tr>
</tbody>
</table>

ADLs: Activities of daily living.
The purpose of the first wave of testing was to gather sufficient data to perform item response theory (IRT) analyses of the items. Constructing Neuro-QOL instruments using modern measurement models, such as IRT, endows the instruments with certain advantages [22]. For example, it allows them to be brief, yet still precise and valid, qualities that are important for acceptance by the neurology community. Using IRT methodology, sets of items can be calibrated along a continuum that covers the full range of the construct to be measured. From this calibrated set, or ‘item bank’, items can be selected to make up ‘short forms’ (typically consisting of six to eight questions) that fit the goals of the user. For example, if a user wished to assess fine motor function in a group of patients who tended to have poor or very poor fine motor function, then he/she might select items that cluster near the lower end of the motor bank. Item banks also provide the foundation for computerized adaptive testing, a specialized type of computer-based testing that enables frequent assessments and immediate feedback with minimal burden on patients and precise evaluation of patients at the individual level [22]. Users can administer short, unique tests to every individual, with reliability and scores that are equivalent to longer, fixed-length assessments. IRT also allows comparisons of patients and questionnaire items across multiple instruments by equating the instruments along a common measurement continuum [23]. Therefore, an additional benefit of using IRT methodology is that a given score on a Neuro-QOL measure can be directly linked to what that individual would score on any existing HRQL measure that has been equated to the Neuro-QOL instruments.

The first wave of testing was conducted online using large, diverse (e.g., stratified by gender, economic status and race/ethnicity) samples from general and clinical populations. The general sample consisted of 3000 adults (2000 English-speaking and 1000 Spanish-speaking) and 1500 children (1000 English-speaking and 500 Spanish-speaking). The clinical sample included 553 adults diagnosed with epilepsy, stroke, amyotrophic lateral sclerosis, multiple sclerosis or Parkinson’s disease, and 59 children with epilepsy or muscular dystrophy. Both samples were recruited through an online panel company. Following psychometric analysis of these data, calibrated short forms were constructed from the item banks, with higher scores indicating a greater degree or level of the construct being measured. These short forms were then validated alongside external validity measures and clinical criteria during a clinical validation study of individuals with epilepsy and the other Neuro-QOL conditions. Short forms were administered at three time points: baseline, 1 week (to evaluate test–retest reliability) and 6 months (to evaluate responsiveness). Validation measures, which included generic instruments given to members of all disease groups as well as disease-specific measures, were administered at baseline and 6 months. In this article, we present initial evidence of construct validity for epilepsy patients, namely the associations between baseline Neuro-QOL Short Forms and baseline Quality of Life in Epilepsy (QOLIE)-31 scores from our clinical validation sample of adults diagnosed with epilepsy and between selected Neuro-QOL Short Forms and seizure frequency. Data analysis is ongoing, and additional psychometric evidence for reliability, validity and responsiveness will be presented in future articles. The QOLIE-31 is a HRQL survey for adults (>18 years of age) with epilepsy [24,25]. Derived from the QOLIE-89, this scale contains domains that include seizure worry, emotional well-being, energy/fatigue, cognition, medication effects, social effects, health status and overall quality of life. Raw scores are converted to a 0–100 scale, with higher scores indicating a better quality of life. The scale has demonstrated good internal consistency (Cronbach’s alphas range from 0.77 to 0.85 for subscales), test–retest reliability (r; ranging from 0.64–0.85), construct validity [26] and responsiveness to change [25,27].

Clinical validation results of the Neuro-QOL Short Forms: adult epilepsy

A total of 119 adults diagnosed with epilepsy (mean age = 47.3 years, standard deviation = 16.9) completed Neuro-QOL Short Forms and external validation measures. Half of this sample was male (50%) and the majority of participants were Caucasian (85.8%) and African American (10.6%). Half of the sample (50.4%) reported having a seizure within the past 4 weeks, and 37.2% of the sample reported that their seizures were severe or very severe. Participants who reported a seizure within the past 4 weeks prior to testing demonstrated worse Neuro-QOL scores for anxiety, depression and stigma, which were found to be statistically significant (p < 0.01). Furthermore, participants who denied having a seizure within the
past 4 weeks reported better scores on Positive Affect and Well-being, and Ability to Participate in Social Roles and Activities, which were also found to be significant (p < 0.05).

The two instruments (Neuro-QOL and QOLIE-31) were highly correlated with each other. The majority of Pearson correlation coefficients were statistically significant (p < 0.001) and demonstrated relationships between Neuro-QOL and QOLIE-31 measures in expected magnitudes and directions (Table 2), suggesting good convergent validity.

Associations between Neuro-QOL Forms of emotional distress (Anxiety, Depression and Stigma) and the QOLIE-31 Emotional Well-being Subscale were in the moderate-to-strong range (r of -0.66, -0.71 and -0.52, respectively) while relations with the QOLIE-31 Seizure Worry Subscale were also in the moderate range (r between -0.42 and -0.55). The Neuro-QOL Positive Psychological Function Short Form was moderately associated with the QOLIE-31 Emotional Well-being and overall quality of life (r of 0.69 and 0.67, respectively). Moderate correlations were observed between Neuro-QOL Social Role Performance and Satisfaction and the QOLIE-31 Social Function (r of 0.58 and 0.52, respectively). In terms of associations with physical function, the Neuro-QOL Mobility and Upper Extremity forms demonstrated moderate associations with the QOLIE-31 total score (r of 0.42 and 0.43, respectively). Neuro-QOL measures of applied cognitive function (Executive Function and General Concerns) produced moderate-to-strong correlations with the QOLIE-31 Cognition Subscale (r of 0.65 and 0.75, respectively), while Neuro-QOL’s Communication Difficulty Scale was also moderately associated to this (r of 0.57). Finally, the Neuro-QOL Fatigue and Sleep Disturbance measures demonstrated moderate associations with the QOLIE-39 Fatigue Subscale (r of -0.65 and -0.52, respectively).

**Neuro-QOL: implications for epilepsy research & treatment**

Health-related quality of life assessment has become an accepted, and almost required, end point in clinical trials. Information regarding HRQL and other patient reported outcomes is important when evaluating treatments, and can be a useful aid to patients and providers as they decide between various therapy options. Neuro-QOL is poised to become a valuable tool for clinical researchers in evaluating treatments. It was developed in a manner consistent with the recent US FDA guidance on patient reported outcomes instrument development [102] and early results indicate that it has good reliability and validity, and other positive characteristics (e.g., brevity, flexibility in administration, suitability for generic and targeted assessment, and minimal administrator and participant burden) that will enable it to be readily incorporated into epilepsy and other clinical research in the field of neurology.

Table 2. Pearson correlations between Neuro-QOL and QOLIE-31 scores.

<table>
<thead>
<tr>
<th>QOLIE-31</th>
<th>Neuro-QOL Short Forms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Anxiety</td>
</tr>
<tr>
<td>Total score</td>
<td>-0.642</td>
</tr>
<tr>
<td>Cognitive</td>
<td>-0.441</td>
</tr>
<tr>
<td>Energy/Fatigue</td>
<td>-0.528</td>
</tr>
<tr>
<td>Emotional Well-being</td>
<td>-0.663</td>
</tr>
<tr>
<td>Medication Effects</td>
<td>-0.410</td>
</tr>
<tr>
<td>Overall QOL</td>
<td>-0.502</td>
</tr>
<tr>
<td>Social Function</td>
<td>-0.505</td>
</tr>
<tr>
<td>Seizure Worry</td>
<td>-0.545</td>
</tr>
</tbody>
</table>

Correlation is significant at the 0.001 level (two-tailed) unless indicated as NS. ADLs: Activities of daily living; NS: Not significant; QOL: Quality of life; QOLIE: Quality of life in epilepsy.
There has been significant interest in using HRQL and other patient reported outcome measures in clinical practice for a variety of purposes, including screening, monitoring treatment progress and quality of care, improving communication between patients and providers, and between care team members, making care more patient-centered, and as an aid in decision-making [28]. However, implementation of patient reported outcomes in clinical practice is challenging and lags behind the use of patient reported outcomes in research. Barriers to implementation include logistical difficulties, need for a supportive environment/infrastructure and uncertainty of providers that collecting HRQL information will be clinically useful [29]. The same characteristics that make Neuro-QOL useful in the research setting should facilitate its use in the clinical setting. For example, it supports computerized data collection and scoring, which can overcome some of the logistical barriers. When released, its reporting system will be designed to make scores easy to understand and interpret, which can help alleviate barriers caused by the clinicians’ misconceptions. Finally, when implemented as a computerized adaptive test (CAT), scores will be sufficiently precise for change to be tracked at the individual level rather than only the group level, a property that many existing, non-IRT-based HRQL measures do not have.

While evaluating HRQL is important for many types of research and clinical uses, many studies require different kinds of measures. We will now describe an NIH initiative to develop measurement tools to assess other aspects of health.

**Beyond HRQL: the NIH Toolbox for the Assessment of Neurological & Behavioral Function**

The NIH Toolbox [103], a Blueprint for Neuroscience Initiative that began in fall 2006, aims to develop a set of brief and well-validated instruments, in English and Spanish, to assess cognitive, emotional, motor and sensory function across all ages from 3 to 85 years [30]. The NIH Toolbox is intended for use in epidemiological and longitudinal studies to identify those aspects of cognition, emotion, motor and sensation that are associated with optimal function and health, as well as for use in large-scale intervention and prevention trials. Currently, there are many ways in which data on function are collected. By adopting the use of a standard set of publicly available tools, the NIH will enable aggregation of data from multiple studies and comparisons across studies, greatly enhancing the value of information collected in any one project. The NIH Toolbox currently includes 47 primary and supplemental instruments assessing the following constructs:

- Cognition – attention, executive function, processing speed, working memory, episodic memory and language
- Emotional health – positive affect, negative affect, social relationships, and stress and coping

<table>
<thead>
<tr>
<th>Table 2. Pearson correlations between Neuro-QOL and QOLIE-31 scores (cont.).</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>QOLIE-31</strong></td>
</tr>
<tr>
<td><strong>Applied Cognition – Executive Function</strong></td>
</tr>
<tr>
<td>Total score</td>
</tr>
<tr>
<td>Cognitive</td>
</tr>
<tr>
<td>Energy/fatigue</td>
</tr>
<tr>
<td>Emotional Well-being</td>
</tr>
<tr>
<td>Medication Effects</td>
</tr>
<tr>
<td>Overall QOL</td>
</tr>
<tr>
<td>Social Function</td>
</tr>
<tr>
<td>Seizure Worry</td>
</tr>
</tbody>
</table>

Correlation is significant at the 0.001 level (two-tailed) unless indicated as NS.
ADLs: Activities of daily living; NS: Not significant; QOL: Quality of life; QOLIE: Quality of life in epilepsy.
* Motor function – locomotion, strength, nonvestibular balance, endurance and dexterity

* Sensory function – vision, audition, vestibular balance, taste, olfaction and somatosensation

Most of the instruments are objective measures of function (e.g., use of a pegboard test to evaluate dexterity). However, as emotional experience is primarily subjective, the Emotional Health Battery is comprised of self- and proxy-report measures of well-being. In addition, there are supplemental self-report measures of vision- and hearing-related HRQL.

As with Neuro-QOL, input from stakeholders and potential end-users was solicited early and repeatedly during the development process. This input guided the selection of constructs to assess, as well as the length and format of measures. The final Toolbox will consist of four domain (cognition, motor function, emotional health and sensation) batteries, each requiring an average of 30 min to administer (20 min for children ages 3–5 years), with a total of 2 h administration time for the entire Toolbox. Supplementary instruments will also be available for investigators wishing to more extensively evaluate a particular domain. Many of the measures are administered by a computer and all will have computerized scoring. Training materials will be available and test administrators will not require any specialized educational background. Instruments were developed using modern psychometric methods, such as IRT, when possible. They have undergone or will undergo calibration and evaluation of reliability and validity in samples ranging from 100–7500 people per instrument. A national norming study will commence in early 2011, with the NIH Toolbox scheduled to become publicly available in early 2012.

The NIH Toolbox: implications for epilepsy

The NIH Toolbox is intended to be brief, minimally burdensome to respondents and administrators, relatively low in cost, psychometrically sound, free of intellectual property issues, and appropriate for use across a wide age range and with diverse populations (e.g., English and Spanish speakers). All these qualities are expected to make its use attractive to investigators because they lessen the barriers (e.g., cost, unknown validity and uncertainty regarding what instrument to select) to measurement. This is true for those already planning to assess aspects of cognitive, emotional, motor or sensory function and, perhaps more importantly, to investigators who might not assess areas of function were it not for the availability of the Toolbox. This should lead to a greater number of studies collecting standard data on more areas of function, which, since these data can be aggregated and directly compared, significantly increases the likelihood of making new discoveries, and identifying currently unknown relationships between function and health, and function and disease. In the study of epilepsy for example, inclusion of the multidimensional Toolbox in longitudinal epidemiological research could reveal new predictors or risk factors for developing intractable epilepsy, as well as currently unsuspected long-term outcomes. This could lead to new prevention strategies as well as additional treatment targets. Similarly, using Toolbox instruments when evaluating treatments for epilepsy could reveal a broader range of treatment effects than it is typically possible to do in a single study or a few studies. This kind of finding, in turn, may lead to adjustment of antiepileptic drug medications in clinical practice.

While the focus of NIH Toolbox development has been for use in research studies, there has been considerable interest in directly utilizing it in the clinical arena. Several external projects are evaluating its use with clinical populations, including patients with Parkinson’s disease, traumatic brain injury, stroke and patients undergoing neurological rehabilitation for acute brain injury. The results of these studies will help inform its use in clinical settings.

Future perspective

The NINDS recently began a common data elements (CDE) initiative to standardize data collection in clinical research. The CDE aims to increase efficiency and facilitate data sharing, which enables one to compare results across studies and aggregate information for meta-analysis or systematic review. Information on the NINDS CDE and its products, including some recommendations regarding CDEs for epilepsy, can be found at [104]. These goals are consistent with those of Neuro-QOL and NIH Toolbox. The NINDS CDE committees and working groups evaluate existing standards, and tools and make recommendations, acknowledging that this will be an ongoing process as the field changes and advances. Neuro-QOL and NIH Toolbox are creating new measures in order to further these goals. We are hopeful that in 5–10 years more studies will not only collect standard data elements but will also utilize common instruments to demonstrate outcomes in their studies – the
Health-related quality of life (HRQL) measurement is essential for understanding the full impact of a disease and its treatment on people's lives. There is a lack of well-validated, standard HRQL assessment tools for use with people suffering from neurological disease.

**Neuro-QOL**

- Neuro-QOL, a multisite, multyear project funded by the National Institute of Neurological Disorders and Stroke (NINDS), has developed a set of instruments (in English and Spanish) to assess the HRQL of adults and children with common neurological conditions.
- The Neuro-QOL measurement system is intended to be brief, reliable, valid, responsive and sufficiently consistent across the selected neurological conditions to allow for cross-disease comparison, and yet flexible enough to capture condition-specific HRQL issues.
- Neuro-QOL measures (13 adult and 11 pediatric) have been validated for use with adults suffering from epilepsy, Parkinson's disease, amyotrophic lateral sclerosis, stroke or multiple sclerosis, and children with epilepsy or muscular dystrophy.

**Clinical validation results of the Neuro-QOL Short Forms: adult epilepsy**

- Neuro-QOL Short Forms have been evaluated in clinical populations.
- Initial results from testing in a sample of adults with epilepsy indicate good convergent validity.

**Beyond HRQL: the NIH Toolbox for the Assessment of Neurological & Behavioral Function**

- The NIH Toolbox is a Blueprint for Neuroscience Initiative that is developing a set of brief and well-validated instruments, in English and Spanish, to assess cognitive, emotional, motor and sensory function in patients aged 3–85 years.
- The NIH Toolbox currently includes 47 primary and supplementary instruments assessing 21 different functional constructs. While most of the instruments directly assess function, the emotional health and some of the sensory domains include measures of HRQL and well-being.
- Normative data for the NIH Toolbox will be collected prior to its public release in 2012.

**The NIH Toolbox: implications for epilepsy**

- Use of the NIH Toolbox will enable aggregation of data from diverse studies, which can lead to identification of currently unknown predictors of and risk factors for intractable epilepsy as well as long-term outcomes.

---

**Bibliography**

Papers of special note have been highlighted as:

* of interest

** of considerable interest


* Summary and discussion of issues surrounding the use of patient reported outcomes in clinical settings.


** Provides a rationale for and purpose of the NIH Toolbox and vision of final products.

### Websites

101 Neuro-QOL – quality of life outcomes in neurological disorders


** Contains information regarding the Neuro-QOL project, including preliminary item banks and how to get access to Neuro-QOL measures.

102 US FDA, Center for Drug Evaluation and Research, Center for Biologics Evaluation and Research, Center for Devices and Radiological Health: Guidance for industry patient-reported outcome measures: use in medical product development to support labeling claims

[http://purl.access.gpo.gov/GPO/LPS13413](http://purl.access.gpo.gov/GPO/LPS13413)


[www.nihtoolbox.org](http://www.nihtoolbox.org)

** The public site for learning about and keeping up to date on the NIH Toolbox.

104 National Institute of Neurological Disorders and Stroke. NINDS Common Data Elements


** Source for information regarding the National Institute of Neurological Disorders and Stroke (NINDS) Common Data Elements and their products.