Supporting Statement A for

Quality of Life Outcomes in Neurological Disorders (NINDS)

Project Officer: Claudia S. Moy, PhD

IC: National Institute of Neurological Disorders and Stroke (NINDS)

Address: Neuroscience Center, room 2214

6001 Executive Blvd., MSC 9520

Bethesda, MD 20892-9520

Voice: 301/496-2789

Fax: 301-480-1080

Email: moyc@ninds.nih.gov

LIST OF ATTACHMENTS

Attachment 1 – Expert Neurology Professionals Consulted

Attachment 2 – Public comment on federal register

Attachment 3 – Panel Testing Company Invitation to Participate

Attachment 4 – Panel Testing Company Privacy Policy

Attachment 5 – ENH IRB Report

Attachment 6 – ENH IRB Approval Amendment

Table of contents

A.

Quality of Life Outcomes in Neurological Disorders (NINDS)

Part A: Supporting Statement for OMB Review

A. JUSTIFICATION

A1. Circumstances Making the Collection of Information Necessary

NINDS requires the development of a tool to assess patient-centered measures for persons with neurological diseases which can be incorporated by the research community as primary and secondary outcomes in clinical trials. This initiative is authorized by the Public Health Service Act, 42 U.S.C. 285(j). The mission of NINDS is to reduce the burden of neurological disease - a burden borne by every age group, by every segment of society, by people all over the world. In support of this mission, NINDS conducts, fosters, coordinates, and guides research on the causes, prevention, diagnosis, and treatment of neurological disorders and stroke, and supports basic research in related scientific areas.

The NINDS Clinical Trials program has been established to support all clinical studies of the neurological disorders and to support development of scientific methodology to advance clinical research in neurology. A key component of the clinical research agenda is conduct of clinical trials to evaluate potential intervention and prevention strategies for neurological disorders. Evaluation of these strategies depends on defining and measuring appropriate outcome variables. Typically, these variables comprise clinical or functional outcomes (such as death, recurrence of stroke, seizure frequency, muscle strength measures, etc.). Many of these traditional clinical or functional measures of disease status do not adequately represent the full scope of the impact of disease on an individual with a chronic neurological disorder. More subjective components of patients' functioning, such as social, psychological, and mental well-being, may be more important components of disease impact. Measurement of patient-oriented outcomes is a particular concern in clinical trials, where small differences in clinical measurements or imaging results may not translate into important benefit to the patients. These measures are essential in order to

provide the full picture of the effect of intervention on patients. Currently, there is no consensus in the neurology clinical trials community about the best tools or approaches; this lack of consensus results in the inability to compare the relative burden of various neurological conditions to each other or, more importantly, to compare the relative benefits of one treatment over another on the same patient-centered outcome.

The project "Quality of Life Outcomes in Neurological Disorders**"** (Neuro-QOL) was implemented to develop a psychometrically robust health-related quality of life (HRQL) measurement tool that is accepted by the neurology clinical trials and clinical research community. Specific goals of the project are:

1. To develop a core set of questions that will address dimensions of HRQL that are universal to patients with chronic neurological disease.
2. To develop supplemental questions or modules that address additional concerns that may be specific to particular groups of patients defined by disease, age, or other factors.

The resulting instruments will become part of the NINDS clinical trials toolkit that will be incorporated into all relevant NINDS-funded clinical trials in the future.

One component of the development of this tool is pilot testing of questionnaire components in patients with neurological disorders as well as participants from the general population. The purpose of this testing is not to develop estimates of quality of life or well-being in patients or the general population but to use data from pilot interviews to determine appropriate instrument structure.

A2. Purpose and Use of the Information

The information from the proposed data collection will be used by NINDS (through a contract with Evanston Northwestern Health Care) to develop, structure, and refine the required HRQL measurement instrument. The proposed information collection will support psychometric testing of HRQL item banks and testing of the Spanish translation of the final questionnaires.

In order to accomplish this, NINDS proposes to recruit 6000 adults (2000 Spanish-speaking) and 3000 children (1000 Spanish-speaking) via an internet-based interface; each individual will be asked to provide sociodemographic and background health information, such as age, gender, educational level and previous diagnosis of common health conditions. After completing these background questions (see Appendix 1), respondents will be asked approximately 100 questions (range = 84-113, depending upon form administered) regarding their physical, mental or social health. Questions are sorted by health domain and "batched" into forms (A through F, See Appendix 2) such that each respondent is asked about a single health area (i.e., one of the six forms in Table A.16-2 and A.16-3). For example, 1,500 adults will complete the social health banks (role performance and role satisfaction; 100 questions), and 1,500 adults will complete the physical function banks (mobility and upper extremity function; 95 items), and so forth (see Table A.16-2).

A3. Use of Information Technology and Burden Reduction

All information will be collected via the Internet using an online panel constructed by an online survey company (most likely IPSOS) and automatically downloaded to a relational database. By eliminating submission of paper forms, respondent burden is reduced, and data collection is more efficient (e.g., data on forms will not need to be re-entered). The items to be tested will be transferred to the survey layout team at the panel testing company and subsequently entered into their online administration system.

A4. Efforts to Identify Duplication and Use of Similar Information

Health-related quality of life (HRQL) is a subjective measure of the impact of disease and therapy for disease that incorporates both positive and negative health states. HRQL domains may include physical status and functioning, psychological status and well-being, social interaction, role performance, physical symptoms, economic impact, and spiritual status. Depending on the characteristics of the disorder and of the patient, these domains may have varying importance. Inclusion of HRQL measures as outcomes in clinical trials is a relatively new development in neurology. HRQL measures have long been used in the area of cancer research, and more recently in other types of diseases, including cardiology and eye disease. These outcomes are needed in order to detect important changes in patients' functioning and well-being in clinical trials, especially when traditional clinical endpoints are poor markers of patient benefit. Increasingly, investigators applying for NINDS grant support to conduct clinical trials are incorporating "quality of life" measures into their trial design; however, due to lack of consensus on appropriate measures and methods in neurology, there has been little consistency in study design within and across disease categories.

In order to evaluate the actual needs in neurology trials with respect to HRQL measurement, the Project Officer has reviewed the literature in this area to determine the current state of the research. This review indicates that within a select set of disease groups (e.g., epilepsy, Parkinson's disease, ALS), several distinct HRQL instruments exist, and that for many other diseases, there has been no published instrument. Through discussions with quality of life experts at international meetings, as well as program staff at other NIH institutes, the lack of and need for a cohesive measurement tool for use in neurological disorders has been identified.

In preparation of this project concept, the Project Officer attended meetings of the International Society

for Quality of Life Research and discussed the current status of HRQL research with a variety of experts, including individuals who had developed instruments for other diseases as well as individuals with an interest in neurology measures. The concept was further discussed with Dr. Leon Ellwein, who developed a similar contract at the National Eye Institute in the mid 1990's. That contract resulted in a widely-accepted instrument that is being used in nearly every clinical trial sponsored by the NEI. Further input was provided by the concept reviewers, Dr. Carol Mangione (who headed the NEI contract), and Drs. Andrew Siderowf and Zachary Simmons, HRQL experts and practicing neurologists. In addition, Dr. Joseph Lipscomb of NCI and Drs. Michael Weinrich and Louis Quatrano of NICHD were consulted during the development of the solicitation.

The actual data proposed to be collected do not currently exist. The data will be unique to the instrument that is currently being developed and will feed back into the psychometric evaluation of the item banks and instruments. Data from the general population are necessary to provide a basis and a context for calibrating results obtained with these new health questionnaires. Whenever possible, to avoid duplication, items from existing HRQL measures and item banks (e.g., from the NIH Patient Reported Outcomes Measurement Information System (PROMIS)) will be used. However, in order to construct an item bank that covers the full range of each construct, the entire group of items proposed for each bank must undergo IRT analyses to determine how well they function and where they fit along the construct continuum. These analyses require that a minimum of 500 people respond to each item within the bank.

A5. Impact on Small Businesses or Other Small Entities

No small businesses will be involved in this data collection.

A6. Consequences of Collecting the Information Less Frequently

This information will be collected only once.

A7. Special Circumstances Relating to Guidelines of 5 CFR 1320.5

This project fully complies with 5 CFR 1320.5.

The data collection in this project will support the development of a measurement instrument (i.e., methodological development) and will not provide results that can or should be generalized to any population. Respondent burden is estimated to be 0.50 hours per respondent.

A8. Comments in Response to the Federal Register Notice and Efforts to Consult Outside Agency

Expert neurology professionals 1) provided input regarding the most important areas of HRQL to assess for neurological conditions in general and for specific diseases; and 2) reviewed candidate items for construct coverage and for clarity. When they identified gaps in item bank content and/or poorly worded items, they often suggested additional or alternative items to include in the item banks. The experts pointed us in the right direction, but they cannot give confirmation that the questions we have selected and written are the best set of questions, nor can they help us obtain the necessary item statistics we need to proceed with delivering the product. Attachment 1 lists the names and affiliations of those consulted.

The 60-day Federal Register notice was published on September 24, 2007 (page 54269). Three responses were received. One merely requested copies of the data collection plans and instruments; these were provided. A second respondent merely asked whether there is website or .pdf for members of the American Society for Experimental NeuroTherapeutics to refer to; there was none to report. The only

comment on the content of the notice was from an individual who signed his/her email "b. schau" and stated an opposition to the project based on his/her assertion that no action was proposed and that non-English-speaking individuals should not be included. The text of the email is attached (Attachment 2).

A9. Explanation of Any Payment or Gift to Respondents

There will be no payment or gift to respondents.

A10. Assurance of Confidentiality Provided to Respondents

All information will be collected by the online survey company, most likely IPSOS, from their member panels. IPSOS will send only de-identified data to CORE, the NINDS contractor for this project. The invitation to participate in the study (Attachment 3) describes the purpose and requirements of the study and describes their policy regarding personal information. In addition, each prospective panelist receives a copy of the IPSOS Privacy Policy (Attachment 4) when invited to become a panel member. IPSOS does not share participants' identifying information with any third party except to comply with legal requirements or under special circumstances as described in their privacy document.

This project has been approved by the Evanston Northwestern Healthcare IRB (FWA 00003000) and has been granted a waiver of consent form for this portion of the study. A copy of the most recent IRB project approval is included as Attachment 5 with the IRB approval of the Wave 1 procedures and waiver of signed consent included as Attachment 6.

A11. Justification for Sensitive Questions

Some questions, such as those related to death and suicidal thoughts or feelings, are sensitive but are included because they address areas of HRQL that 1) patients and caregivers have themselves identified as being significantly affected by neurological disease and its treatment and 2) are legitimate targets for clinical and other treatment. Therefore, these questions are necessary to include in order to meet NINDS'

requirement to develop a measure that evaluates important patient-reported outcomes and can be used in neurology clinical trials. We will also ask about income (for the purpose of describing our sample), but attempt to make this inquiry less intrusive by asking respondents to choose an income "bracket" (e.g. "Between $20,001 and $40,000") rather than to give a specific amount. Members of IPSOS online panels will be emailed an invitation to participate in the study (Attachment 3). This invitation describes the purpose of the project and what they are being asked to do, and assures them IPSOS will not release identifying information to third parties.

Confidentiality is protected by having all assessments completed on IPSOS's own testing platform and secure server and by providing only de-identified data to CORE.

A12. Estimates of Hour Burden Including Annualized Hourly Costs

The estimates of hour burden provided below are based on the contractor's experience during a similar item bank development project, PROMIS, within which respondents completed similar types of items. With respect to time costs, we do not know the professional backgrounds or employment status of respondents and therefore we have no knowledge about hourly wages. We have chosen $25 as an average for the purpose of calculating a time-cost estimate.

Type of Respondents	Number of Respondents	Frequency of Response	Average Time per Response	Annual Hour Burden
Adults	6000	1	0.5	3,000
Children	3000	1	0.5	1,500
Totals	9000			4,500

A.12 – 2 ANNUALIZED COST TO RESPONDENTS				
Type of Respondents	**Number of Respondents**	**Frequency of Response**	**Hourly Wage Rate**	**Respondent Cost**
Adults	6000	1	$25	$75,000
Children	3000	1	$0	$0
			Total	$75,000

A13. Estimate of Other Total Annual Cost Burden to Respondents or Recordkeepers

No other costs to respondents are involved in this data collection.

A14. Annualized Cost to the Federal Government

The cost of this proposed one-time pilot data collection is $163,625, expected to be completed in less than one year. [The total cost of the entire Neuro-QOL project is $4,402,119 over 5 years, or approximately $880,000 per year. This includes all developmental phases of the project.]

A15. Explanation for Program Changes or Adjustments

This is a new collection of information.

A16. Plans for Tabulation and Publication and Project Time Schedule

A.16 – 1 Project Time Schedule for Phase II	
Activity	**Time Schedule**
Letters sent to respondents	1 – 2 weeks after OMB approval
Online data collection	0.5 -2 months after OMB approval
IRT analyses	2 – 6 months after OMB approval
Development of short forms	7 – 8 months after OMB approval
Field testing with clinical sample	9 – 20 months after OMB approval
Analyses of field testing data	21 – 22 months after OMB approval
Packaging of final measurement system	23 – 25 months after OMB approval
Delivery of measurement system and associated documentation to NINDS	26 months after OMB approval

Phase I of Neuro-QOL, which involved identifying the HRQL domains to target for item bank development, constructing the item banks, and translating all items into Spanish, has been completed. The adult and pediatric generic item banks and the targeted scales developed in Phase I are to be tested in Phase II. Phase II involves field testing of the banks in the general population for the purpose of item calibrations (Wave 1 testing), development of short forms from the full item banks, and then evaluation of these short forms in the target populations for reliability, validity, responsiveness to change and usefulness of proxy data (Wave 2 testing). We are currently seeking OMB approval only for the Phase II, Wave 1 information collection from the general population described below.

Phase II; Field-Testing of Instrument: Wave 1: The purpose of Wave 1 testing is to provide item calibrations (i.e., item statistics such as slope and location parameters) obtained from Item Response Theory (IRT) models. A maximum of 6000 adults (4000 English-speaking and 2000 Spanish speaking)

and 3000 children (2000 English-speaking and 1000 Spanish-speaking) will be recruited by the panel testing company for Wave 1. The majority of respondents will be randomly sampled from the general population. However, this general sample will be supplemented with participants who self-identify as having one or more current medical illnesses. This is to help ensure that all categories of all items are answered by some proportion of the sample (necessary for item calibrations), with the assumption that persons with an illness are more likely to function at the lower end of at least some traits. A detailed sampling plan is presented in section B.

The item banks will be divided amongst an initial set of four forms (A-D) for adults and two for children (E-F) as shown in Tables A.16-2 and A.16-3. Each respondent will take only the banks that comprise a single form such that each adult participant will answer a maximum of 202 questions (including sociodemographic and background questions) and each pediatric participant will answer a maximum of a 130 questions. For each trait tested by our item banks, item calibrations and norm tables will be created based on the general population sample (maximum 1,500 people per form). Following the completion of the first wave of testing, the analysis of the item bank data will be conducted and short forms of each item bank will be developed.

Table A 16-2

Adult Banks	**Form**	**Number of Survey Items Per Form**	**Total Respondents**
Social Role Performance	A	**100**	**1,500**
Social Role Satisfaction			
Mobility/Ambulation	B	**95**	**1,500**
Upper Extremity/ADLs			
Depression	C	**86**	**1,500**
Fear/Anxiety			
Positive Psychological Function			
Perceived Cognitive Function	D	**89**	**1,500**
Applied Cognitive Function			
Total enrollment			**6,000**

Table A 16-3			
Pediatric Banks	**Form**	**Number of Survey Items Per Form**	**Total Respondents**
Emotional Distress	E	**84**	**1,500**
Social Function			
Upper Extremity Function	F	**113**	**1,500**
Mobility/Lower Extremity Function			
Total enrollment			**3000**

Analyses

Factor Structure. Following data clean-up, our first task will be to examine the unidimensionality of the items in each bank. Each bank will be analyzed separately. Spearman's rho and item-total correlation will be used to examine the relationship among items; non-parameter based Mokken Scale Procedures will be used to examine item scalability and monotonicity. Residual correlation will be used to examine local dependency among items; and, finally, factor analysis (FA) will be implemented to identify the factor structure among items. FA was chosen over principal components analysis because in most cases we hypothesize an underlying (causal) latent variable represented by the pool of administered items.

Item calibrations. The sample size will enable both 1-parameter and 2-parameter analyses to be conducted (including the Graded Response Model which has been selected as a default analysis approach). These analyses will produce item calibrations for all the items in each domain, and short forms constructed for each item bank. In parallel, sufficient analyses will be conducted to create the item parameters necessary for computerized adaptive testing (CAT – albeit that CAT will not be utilized for testing as part of this contract).

Item quality control. The fit of the items to the rating scale model will be evaluated by examining the infit mean square statistic for the calibrated items. This will be done separately for each item bank. "Infit" is an information-weighted fit statistic, which is sensitive to responses to items near the person's level of

what is being measured. Infit mean square has an expected value of 1.0. This fit statistic will expose digressions in the data that violate the measurement specifications. If the specifications are not met, then the calibration is disqualified and investigation is undertaken to identify which aspects of the data are not conforming. If the infit mean square lies between .6 and 1.4 we will consider the item to have good fit. Any item with mean square infit < .6 (i.e. "over-fit") is, in and of itself, not of concern for inclusion in an item bank. Items that do not fit the model will be evaluated and reviewed by our advisory panel (comprised of the expert consultants and clinicians in the study) before the final HRQL item bank is established. Misfitting items may be retained when they are identified as clinically relevant.

Item equivalence -Stability of item parameters across conditions (Differential Item Functioning, DIF). The extent to which items in a questionnaire perform similarly across different reference groups is of critical interest when determining whether a given questionnaire can be used fairly to compare groups. Using our standard methods for evaluating equivalence, we will determine whether any detected group difference is based upon items that display different statistical properties across groups. Having determined individual item location, we can then identify items which demonstrate a difference in location based solely on functional status or demography, i.e., differential item functioning (DIF), such as socioeconomic status. The most important indicator of DIF is not whether items systematically differentiate relevant subgroups, but whether they do so in an unmodeled (i.e. unpredicted) way.

To avoid measurement bias, we will ensure item parameter stability across gender and race/ethnicity. In particular we will take care to ensure the appropriateness of the instrument (i.e., culture fairness based on minimal DIF) for African American and US Hispanic populations. We will examine DIF for other variables (e.g., education levels) if sufficient sample is found for these characteristics. Items that do not hold stable psychometric properties across conditions (e.g., between male and female, or between Hispanic and non-Hispanic) are considered as demonstrating DIF. We will use at least two techniques (ordinal logistic regression and Rasch analysis) to identify DIF items. If sample size is sufficient within each sub-condition (e.g., >200), other techniques such as IRTLRDIF will be used as

well. Items that demonstrate DIF with all techniques will be removed from the final banks. Items that demonstrate DIF with some, but not all, techniques will be considered peripheral items. Inclusion/exclusion of peripheral DIF items will be determined via consensus.

Missing Data. We propose to treat missing data as "not presented" for calibration and will examine patterns of missing data initially to determine reasons for missing data. For example, we will look to see if missing data is due to response burden or lack of time in completing the survey (e.g., more prevalent later in the sequence of administered items) and whether missing data is more prevalent with sensitive item content.

A17. Reason Display of OMB Expiration Date is Inappropriate

Data collection will be conducted by internet-based interviews; the OMB expiration date will be displayed on the data collection screen.

A18. Exceptions to Certification for Paperwork Reduction Act Submissions

No exceptions are requested.