

Patient-reported outcomes of pain and physical functioning in neurofibromatosis clinical trials

Pamela L. Wolters, PhD
Staci Martin, PhD
Vanessa L. Merker, BS
James H. Tonsgard, MD
Sondra E. Solomon,
PhD†
Andrea Baldwin, CRNP
Amanda L. Bergner, MS
Karin Walsh, PsyD
Heather L. Thompson,
PhD, CCC-SLP
Kathy L. Gardner, MD
Cynthia M. Hingtgen,
MD, PhD
Elizabeth Schorry, MD
William N. Dudley, PhD
Barbara Franklin, BS
For the REiNS
International
Collaboration

Correspondence to
Dr. Wolters:
woltersp@mail.nih.gov

ABSTRACT

Objective: Tumors and other disease complications of neurofibromatosis (NF) can cause pain and negatively affect physical functioning. To document the clinical benefit of treatment in NF trials targeting these manifestations, patient-reported outcomes (PROs) assessing pain and physical functioning should be included as study endpoints. Currently, there is no consensus on the selection and use of such measures in the NF population. This article presents the recommendations of the PRO group of the Response Evaluation in Neurofibromatosis and Schwannomatosis (REiNS) International Collaboration for assessing the domains of pain and physical functioning for NF clinical trials.

Methods: The REiNS PRO group reviewed and rated existing PRO measures assessing pain intensity, pain interference, and physical functioning using their systematic method. Final recommendations are based primarily on 4 main criteria: patient characteristics, item content, psychometric properties, and feasibility for clinical trials.

Results: The REiNS PRO group chose the Numeric Rating Scale-11 (≥ 8 years) to assess pain intensity, the Pain Interference Index (6-24 years) and the Patient-Reported Outcome Measurement Information System (PROMIS) Pain Interference Scale (≥ 18 years) to evaluate pain interference, and the PROMIS Physical Functioning Scale to measure upper extremity function and mobility (≥ 5 years) for NF clinical trials.

Conclusions: The REiNS Collaboration currently recommends these PRO measures to assess the domains of pain and physical functioning for NF clinical trials; however, further research is needed to evaluate their use in individuals with NF. A final consensus recommendation for the pain interference measure will be disseminated in a future publication based on findings from additional published research. *Neurology*® 2016;87 (Suppl 1):S4-S12

GLOSSARY

CAT = computerized adaptive test; **COA** = clinical outcome assessment; **FDA** = Food and Drug Administration; **IRT** = Item Response Theory; **NF** = neurofibromatosis; **NF1** = neurofibromatosis 1; **NF2** = neurofibromatosis 2; **NRS-11** = Numeric Rating Scale-11; **PII** = Pain Interference Index; **PRO** = patient-reported outcome; **PROMIS** = Patient-Reported Outcome Measurement Information System; **PROMIS-PF** = Patient-Reported Outcome Measurement Information System Physical Functioning scale; **PROMIS-PI** = Patient-Reported Outcome Measurement Information System Pain Interference scale; **QOL** = quality of life; **REiNS** = Response Evaluation in Neurofibromatosis and Schwannomatosis; **SF** = short forms; **SWN** = schwannomatosis.

Neurofibromatosis 1 (NF1), neurofibromatosis 2 (NF2), and schwannomatosis (SWN) are neurogenetic disorders that share a predisposition to develop multiple nerve sheath tumors. These tumors may cause significant complications such as visual impairments, hearing loss, facial nerve palsy, airway constriction, and spinal cord compression.¹ Generalized and focal orthopedic problems, including spinal and skeletal abnormalities such as scoliosis, congenital bowing, tibial

Supplemental data at Neurology.org

†Deceased.

From the Pediatric Oncology Branch (P.L.W., S.M., A.B.), National Cancer Institute, National Institutes of Health, Bethesda, MD; Department of Neurology and Cancer Center (V.L.M.), Massachusetts General Hospital, Boston; University of Chicago Pritzker School of Medicine (J.H.T.), IL; Department of Psychological Sciences (S.E.S.), University of Vermont, Burlington; Departments of Neurology and Genetics (A.L.B.), Johns Hopkins University, Baltimore, MD; Children's National Health System & The George Washington School of Medicine (K.W.), Washington, DC; Department of Speech Pathology & Audiology (H.L.T.), California State University, Sacramento; Veteran's Administration Pittsburgh Healthcare System and University of Pittsburgh (K.L.G.), PA; Department of Clinical Neurosciences (C.M.H.), Spectrum Health Medical Group and College of Human Medicine, Michigan State University, East Lansing; Division of Human Genetics (E.S.), Cincinnati Children's Hospital, OH; Department of Public Health Education (W.N.D.), School of Health and Human Sciences, University of North Carolina at Greensboro; and Advocate NF2 Inc. (B.F.), Los Angeles, CA.

REiNS International Collaboration members are listed on the *Neurology*® Web site at Neurology.org.

Go to Neurology.org for full disclosures. Funding information and disclosures deemed relevant by the authors, if any, are provided at the end of the article.

pseudoarthrosis, and chest wall malformations, also may contribute to morbidity.² Furthermore, chronic pain is common in children^{3,4} and adults^{5,6} with neurofibromatosis (NF) and may be related to tumors, physical impairments, and headaches.^{2,3,5,6} These complications of NF can significantly affect physical functioning, activities of daily living, and quality of life (QOL).^{4,6}

The number of clinical trials evaluating treatments for tumors and other manifestations of NF1, NF2, and SWN has increased over the last several years. To assess clinical benefit from the patients' perspective in drug trials, investigators have begun to incorporate patient-reported outcome (PRO) measures as secondary endpoints.⁷ Current behavioral intervention studies focusing on reducing pain and its interference with everyday functioning use PROs as primary endpoints.⁸ PRO measures also can further define baseline disease characteristics⁷ and document effects of drug toxicities.⁹ Regulatory agencies in the United States and Europe recommend including PRO tools in clinical trials to systematically evaluate symptoms of interest^{10,11} as the data can provide complementary information about the efficacy of a new treatment.¹⁰ To support medical product approval and labeling claims in the United States, PRO tools must be qualified by the Food and Drug Administration (FDA) as a clinical outcome assessment (COA)¹² and meet rigorous standards to be used as trial endpoints.¹¹ In meetings with Response Evaluation in Neurofibromatosis and Schwannomatosis (REiNS) members, the FDA has indicated that COAs of pain and physical functioning need to be incorporated into trial design as coprimary or secondary endpoints to document clinical benefit in addition to tumor reduction for NF trials seeking drug approval.

Despite the increased interest in PROs, challenges remain in using such measures of pain and physical functioning in NF clinical trials. First, the complications of NF are heterogeneous and vary widely in prevalence and severity.^{1,13} Individuals often have multiple tumors and complications, making it difficult to differentiate changes in pain or functioning related to one particular tumor type or complication being treated in a clinical trial. Pain may

be acute, episodic, or chronic, and vary in intensity and location. Thus, when designing an NF trial including pain PROs, researchers need to consider the frequency of evaluations, length of the response period (e.g., past 7 days), and how to track the locations of target tumors or complications in order to obtain reliable and valid assessments of pain throughout the study. Accurate recording of analgesic use also is required to help distinguish changes in pain related to tumor reduction due to the study drug from changes in pain due to analgesics.¹⁴ In addition, NF complications, such as plexiform neurofibromas^{3,5} and learning disabilities,¹⁵ are common in children through adults. Therefore, PRO measures used in NF trials must evaluate individuals across a wide range of ages and ability levels. To be feasible for multicenter and international studies, PRO tools need to be relatively brief, easy to administer, and available in different languages. Finally, no measures currently are validated to assess pain and physical functioning across the lifespan in NF, and there is a lack of consensus about how to assess PROs in NF clinical trials. The US FDA and European regulatory agencies have acknowledged the importance of developing a common core language and methodology for clinical trials evaluating new drugs that will permit comparisons of PROs across treatments.¹⁰

To respond to these challenges and promote consensus in the assessment of PROs in NF clinical trials, the PRO group of the REiNS International Collaboration is working to provide guidance about appropriate PRO measures and methodologies in 4 core domains (pain, physical functioning, disease-specific QOL, and general QOL) for use in NF clinical trials. The PRO group's first publication¹⁶ outlined their goals and described the systematic process they developed to review and choose PRO measures using pain intensity as an example. This article presents the current REiNS recommendations for the domains of pain and physical functioning, specifically focusing on self-report measures for ages 8 years and older and parent report questionnaires for 5 years and older.

METHODS The group's systematic method involves gathering information from extensive literature reviews and rating existing PRO measures on 6 criteria, as previously described.¹⁶ PRO measures may be re-reviewed to consider new published data, and when scores are close, the highest-rated measures are compared directly during the

same meeting. Of the 6 initial criteria, the group refined their focus to the 4 it deemed most important when making the final selection of a PRO measure¹⁷ for NF clinical trials. Thus, patient characteristics, item content, psychometric properties, and feasibility for multicenter trials were the primary criteria used to determine the recommendations in the domains of pain and physical functioning. All the PRO measures reviewed for these domains are listed in table e-1 on the *Neurology*[®] Web site at Neurology.org.

RESULTS Pain domain. Pain intensity recommendation.

The PRO group recommends the Numeric Rating Scale–11 (NRS-11) as a self-report measure of pain intensity (ages ≥ 8 years) for NF clinical trials, as described and compared in detail with other tools in our previous publication.¹⁶ On the NRS-11, respondents are asked to circle the number from 0 (no pain) to 10 (e.g., worst pain imaginable) that best represents their pain.¹⁸ The characteristics of the NRS-11 are shown in table 1.

Pain interference. In the domain of pain interference, our group initially identified 7 pain interference scales to consider. Of these, 4 met basic criteria to be suitable for use in NF clinical trials and were rated using our systematic process. The top-rated measures were the Pain Interference Index (PII),¹⁹ Patient-Reported Outcome Measurement Information System Pain Interference scale (PROMIS-PI),²⁰ and Brief Pain Inventory Pain Interference subscale.²¹ Our final side-by-side comparison resulted in the highest overall ratings, which were tied, for the PII and PROMIS-PI. Table 2 presents the characteristics of the PII and PROMIS-PI for each of the 4 main review criteria.

Pain Interference Index. The PII^{19,22} is a 6-item scale that assesses the extent to which pain has interfered with an individual's daily activities in the past 7 days. Items are rated on a 7-point Likert scale, and the total score is

the mean of the completed items. The PII is suitable for NF clinical trials because it covers a wide age range, assesses areas relevant for NF, and has parallel child and adult self-report and parent proxy forms. Importantly for use in pediatric NF trials, the child and parent forms demonstrate good reliability and validity in NF1,¹⁹ yield high correlations between parent and child scores,¹⁹ and show sensitivity to change after behavioral interventions for chronic pain in children.⁸ While good preliminary data exist in adults with NF1 (unpublished data) and other conditions involving pain,²³ more research is needed before recommending the adult version for NF trials.

Patient-Reported Outcome Measurement Information System (PROMIS). PROMIS was created from an NIH-sponsored initiative to develop standardized PRO measures that can be compared across patient populations and studies in a variety of health domains (<http://www.nihpromis.org/>). It utilizes Item Response Theory (IRT), in which each item in a large item bank is a measure of a latent trait that falls along a continuum. A small subset of these items are administered as short forms (SF) of various lengths (e.g., 8, 10, or 20 items) or as a computerized adaptive test (CAT), which selects the most relevant questions from an item bank based on a person's answers to previous questions, to produce brief and precise outcome assessments.^{24,25} Items are rated on a 5-point Likert scale, and the measures yield standardized T scores. PROMIS scales are appropriate for NF clinical trials because they assess relevant domains in a wide age range, have good psychometric properties, and are translated into multiple languages.

PROMIS-PI. The PROMIS-PI²⁰ assesses the degree to which pain has interfered with a patient's ability to complete daily activities over the past 7 days. This

Table 1 Characteristics of the Numeric Rating Scale–11 (NRS-11) for the 4 main review criteria

Criteria	NRS-11 characteristics
Patient characteristics	• Self-report form for ≥ 8 years ^{18,40}
	• Preliminary validity data for children 6–7 years ⁴¹
	• No published data using the NRS-11 in neurofibromatosis to date ¹⁶
Domains/item content	• Assesses pain intensity
	• Pain intensity rated from 0 (no pain) to 10 (worst pain you can imagine) over the past week
	• Different instructions and pain descriptors used ⁴⁰
Psychometric data	• Good test-retest reliability for children (mean difference 0.2; 95% CI 0.0–0.4) ⁴² and adults (intraclass correlations 0.72–0.78) ⁴³
	• Good construct validity supported by correlations with other pain intensity measures (e.g., NRS-11 and FPS-R, $r = 0.78$) ¹⁸
	• Sensitive to change over time in children ⁴² and adults ⁴³
Feasibility	• Free, brief, simple to administer
	• Can be administered by paper/pencil, verbally, or electronically
	• Can modify descriptors and instructions
	• Available in multiple languages and can be translated

Abbreviations: CI = confidence interval; FPS-R = Faces Pain Scale–Revised.

Table 2 Characteristics of the Pain Interference Index (PII) and the Patient-Reported Outcome Measurement Information System Pain Interference scale (PROMIS-PI) for the 4 main review criteria

Criteria	PII characteristics	PROMIS-PI characteristics
Patient characteristics	• Self-report adult form ≥ 18 years	• Self-report adult form ≥ 18 years
	• Self-report pediatric form 6–24 years	• Self-report pediatric form 8–18 years
	• Parent proxy form 6–18 years	• Parent proxy form 5–18 years
	• Validation with children with chronic pain and with NF1	• Validation with general population and several pediatric/adult disease groups
Domains/item content	• Items assess appropriate content for NF	• Items assess appropriate content for NF
	• Child/parent/adult items: school/work, leisure time, mood, time with friends/family, physical activities, sleep	• Child/parent items: schoolwork, sleeping, standing, walking, running, having fun, paying attention, feeling angry
	• Parallel content and response format	• Adult items include day-to-day and social activities, work around home, family life, chores, fun things, and enjoyment of social activities and life
	• Developed using Classical Test Theory	• Not parallel content and response format between child/parent and adult forms
Psychometric data	• Internal consistency in child self-report (Cronbach $\alpha = 0.84$) and parent proxy (0.96) ¹⁹ very good in NF1/cancer	• Internal consistency of the PI bank very good in adults (Cronbach $\alpha = 0.99$) ²⁰ and children (0.87) ³⁵
	• No test-retest reliability data to date	• Good test-retest reliability in adults (intraclass correlations = 0.83–0.87) ²⁶ , lower in children (0.62) ³⁵
	• Good correlation between child self-report and parent proxy scores ($r = 0.62$, $p < 0.0001$) ¹⁹ in NF1/cancer	• Moderate agreement between child and parent PI scores ($r = 0.49$) ³⁰
	• Demonstrated construct validity based on correlation with established pain measures ($r = 0.81$) and differentiates patients with varying levels of NF1 disease severity ¹⁹	• Demonstrated construct validity with other PI scale (Spearman $\rho = 0.90$) and known groups validity in adults ²⁰
	• Responsive to change in children and adolescents after intervention ⁸	• Known groups validity by disease severity for child self-report PI scale ²⁷ ; no validity data for parent proxy to date
	• Good preliminary psychometric data in adult patients ²³	• Responsive to change in adults ²⁸ ; no change detected in one pediatric longitudinal study to date ²⁹
Feasibility	• Brief, easy to administer, free	• Brief, easy to administer, free
	• English and Swedish languages	• Multiple languages
	• Can modify items and format for NF	• Can add items to customize for NF
	• Can administer forms electronically	• Can administer SF electronically; CAT administered electronically

Abbreviations: CAT = computerized adaptive testing; NF = neurofibromatosis; NF1 = neurofibromatosis 1; SF = short form.

scale has demonstrated feasibility and validity in adults and youth with various medical conditions and chronic pain.^{20,26,27} The PROMIS-PI also has been shown to be responsive to changes in pain in a limited number of clinical trials with adults²⁸ but not yet with children.²⁹ Published studies using both the pediatric self-report and parent-proxy pain interference scales currently are scarce. The pediatric and parent forms are parallel and general parent-child agreement is moderate.³⁰ The item content and response options of the adult form are not consistent with the child/parent forms; thus, research is needed to evaluate whether these scales can provide a reliable assessment across age ranges for longitudinal studies involving children through adults. Studies investigating the self-report and parent-proxy versions in NF, including comparisons of CAT and SF administration in children, also are required before our group can recommend the PROMIS-PI for pediatric NF trials.

Pain interference recommendation. Since both of these pain interference tools have strengths and weaknesses for use in NF, the current REiNS recommendation is to utilize the PII in clinical trials involving children to young adults (6–24 years), the PROMIS-PI scale for adult trials (≥ 18 years), and when possible, both measures for trials enrolling children through adults. Since there are limited publications using the PROMIS-PI in longitudinal studies and none with the NF population to date, the PRO group recommends administering the longest PI SF available (8 items) or both the 8-item SF and CAT in NF clinical trials to compare findings. Our group will review newly published data in the future to determine the one pain interference tool most suitable for NF clinical trials.

Physical functioning domain. In the domain of physical functioning, our group identified 6 PRO measures

Table 3 Characteristics of the Patient-Reported Outcome Measurement Information System Physical Functioning scale (PROMIS-PF) for the 4 main review criteria

Criteria	PROMIS-PF characteristics
Patient characteristics	• Self-report adult form ≥ 18 years
	• Self-report pediatric form 8-17 years
	• Parent proxy form 5-18 years
	• Large validation samples including various disease groups
Domains/item content	• Items assess appropriate content (upper extremity and mobility) for NF
	• Measures both motor function and daily living activities
	• Child and parent mobility items include doing sports, playing with other children, getting up from the floor, standing up, standing on tiptoes, and walking upstairs; upper extremity items include buttoning shirt, opening a jar, opening school binder rings, pouring a drink, putting on shoes, and using a key to unlock doors
	• Adult items include how difficult it is to do chores, dress self including buttons/shoelaces, shampoo hair, wash/dry body, get on/off toilet, and run a short distance; how their health limits physical tasks like bending or kneeling, walking, climbing stairs, carrying groceries, and doing vigorous activities, like sports or lifting heavy objects
Psychometric data	• Adult forms not parallel with child/parent forms; adult scale is unidimensional (physical functioning) and child/parent scales have 2 dimensions (upper extremity and mobility) ³³
	• Internal consistency excellent in adults with cancer (PF scale Cronbach $\alpha = 0.92-0.96$) ³⁴ and moderate in children (SF upper extremity = 0.62; mobility = 0.73) ³⁵
	• Good test-retest reliability of PF scale in adults with osteoarthritis (intraclass correlations = 0.92 CAT; 0.95 SF) ²⁶ and moderate in children (SF upper extremity = 0.66; mobility = 0.77) ³⁵
	• Demonstrated construct validity with established measures (all $r_s > 0.68$) and known groups validity in adults with cancer ³⁴
	• Known groups validity by disease severity for child upper extremity and mobility PF scales ²⁷ ; no validity data for parent proxy to date
	• Good agreement between child self-report and parent proxy (upper extremity $r = 0.69$; mobility $r = 0.63$) ⁴⁴
	• Sensitive to change after spine augmentation in adults ²⁸ ; no change after surgery in children with cerebral palsy ²⁹
	• CAT version less accurate than the SF in children with cerebral palsy ³⁸ ; CAT was more precise than SF but CAT administered more items ³⁵
Feasibility	• Brief, free, easy to administer
	• Multiple languages for adult form; English and Spanish pediatric forms
	• Can add items to customize the PF scale to NF
	• Electronic administration of SF or CAT
	• Easy to integrate testing with other PROMIS domains

Abbreviations: CAT = computerized adaptive testing; NF = neurofibromatosis; SF = short form.

appropriate for consideration; 5 met basic criteria to be rated. Of these, the PROMIS Physical Functioning scale (PROMIS-PF),²⁵ Health Assessment Questionnaire,³¹ and Functional Disability Inventory³² were the top-rated measures. From our final side-by-side comparison of these 3 scales, the overall rating of the PROMIS-PF for use in NF clinical trials was the highest. Table 3 presents the characteristics of the PROMIS-PF for each of the 4 main review criteria.

PROMIS-PF. The PROMIS-PF scale measures the perception of one's capability of performing specific physical activities. It includes items assessing upper extremity function, mobility (lower extremity), and, for adults, axial or head and neck function.²⁵ The PROMIS-PF demonstrates good reliability and construct validity in adults and youth from the general population and various disease groups.^{25-27,33-35} Adult studies show that the IRT-based PROMIS-PF scales reduce floor and ceiling effects,³⁶ are more accurate

and efficient, decrease administration time, result in greater responsiveness, and require smaller study sample sizes than legacy PRO measures.^{36,37} Administration of the 10-item PROMIS-PF CAT to adults resulted in an efficient and precise measurement of physical functioning, which differentiated between age and disease groups.³⁶ However, in children with cerebral palsy, the PROMIS Mobility CAT did not correlate with performance measures or differentiate known mobility groups.³⁸ In clinical trials, the PROMIS-PF showed sensitivity to change after spine augmentation in adults²⁸ but not after surgery in children with cerebral palsy.²⁹

Physical functioning recommendation. Although published studies have not used the PROMIS-PF in NF to date, the evaluation of both upper and lower extremity function as well as the range, content, and wording of the items are well-suited to assess the spectrum of complications affecting physical function in NF. Therefore, the current REiNS recommendation is

Table 4 Response Evaluation in Neurofibromatosis and Schwannomatosis (REiNS) patient-reported outcome recommendations for the pain and physical functioning domains in neurofibromatosis clinical trials

Trial age range	Pain intensity	Pain interference	Physical functioning
Pediatric	NRS-11 (SR ≥ 8)	PII (SR 6-24; PP 6-18)	PROMIS-PF (SR 8-17; PP 5-17)
Adult	NRS-11 (SR ≥ 8)	PROMIS-PI (SR ≥ 18)	PROMIS-PF (SR ≥ 18)
Pediatric-adult	NRS-11 (SR ≥ 8)	PII (SR ≥ 6; PP 6-18) and PROMIS-PI (SR ≥ 8; PP 5-17) ^a	PROMIS-PF (SR ≥ 8; PP 5-17)

Abbreviations: NRS-11 = Numeric Rating Scale-11; PII = Pain Interference Index; PP = parent proxy; PROMIS-PF = Patient-Reported Outcome Measurement Information System Physical Functioning scale; PROMIS-PI = Patient-Reported Outcome Measurement Information System Pain Interference scale; SR = self-report.

^a Currently, REiNS is recommending the PII for pediatric-only trials and the PROMIS-PI scale for adult-only trials. In clinical trials enrolling both children and adults, it is recommended to give both the PII and PROMIS-PI until more data are published and REiNS can make a final recommendation.

to use the longest PF SF that is feasible for a particular study (e.g., 20-item SF was the most responsive in populations with physical impairments)³⁷ or the CAT in adult studies. The CAT alone is not recommended for pediatric studies at this time since it was less accurate than the SF in children with functional limitations.³⁸ When possible, giving both SF and CAT versions in pediatric NF studies will allow comparisons between the 2 versions to help refine these recommendations.

Recommendations for use of these measures for NF1, NF2, and SWN. The above recommendations on pain intensity, pain interference, and physical functioning PRO measures (listed in table 4) apply to clinical trials enrolling patients with NF1, NF2, or SWN. Each measure was evaluated for its ability to capture relevant changes in disease status regardless of the particular syndrome being studied (e.g., adequately measuring pain interference from plexiform neurofibromas in NF1 or from generalized chronic pain in SWN). Since the NRS-11, PII, PROMIS-PI, and PROMIS-PF assess generalized constructs rather than disease-specific manifestations of pain and disability, they are suited for a range of clinical trials in NF1, NF2, and SWN. However, the patient population and objectives of a particular clinical trial will dictate which of these PRO measures are most appropriate and whether additional measures may be needed to capture the full range of relevant PROs adequately. For example, a targeted measure on swallowing or breathing may be warranted for studies on plexiform tumors causing airway compression in NF1. Similarly, in trials for symptomatic vestibular schwannomas in NF2, targeted questionnaires on the effect of balance problems on physical activities or the effect of hearing decline on communicative ability may be warranted.

DISCUSSION PROs assessing pain and physical functioning are critical for documenting the clinical efficacy of treatments targeting manifestations of NF1, NF2, and SWN. The REiNS International Collaboration provides consensus recommendations

regarding PROs to the NF research community based on extensive review of existing tools in these domains.

To assess pain intensity, the REiNS International Collaboration recommends the NRS-11 as a brief, reliable, and valid self-report scale to assess pain intensity for ages 8 years and older, as previously described.¹⁶ However, for NF clinical trials targeting tumor reduction, research needs to explore modifications for assessing tumor-specific pain and indicating the location of the tumor pain being rated throughout a trial.

To assess pain interference, the current REiNS recommendation includes 2 measures until additional data are published to determine the most appropriate one for use in NF clinical trials across all age groups. At this time, the REiNS group recommends the PII for pediatric studies and the PROMIS-PI for adult trials. Both measures are recommended for NF studies including children and adults so the data can be compared. The PII is valid in children with NF1 and their parents¹⁹; however, there currently are no published data in adults with NF1. The PROMIS-PI²⁰ is an IRT-based measure that has been used in several published studies with adults, has excellent feasibility, and has robust psychometric properties. However, additional research is needed to examine the use of the PROMIS-PI in children and parents, individuals with NF, and CAT in longitudinal studies.

In the domain of physical functioning, the REiNS Collaboration recommends the PROMIS-PF scale.²⁵ Although relatively new, the PROMIS-PF was deemed the most appropriate for NF clinical trials because it evaluates both upper and lower extremity functioning, consisting of a range of motor functions and daily living activities that are important to assess in NF. This scale should provide an efficient and precise PRO assessment of physical functioning and require smaller sample sizes,³⁷ which is important for clinical trials of rare diseases. However, CAT administration was not as accurate as the SF in children with functional limitations so both administration options need to be examined in the NF population. Additional information also is needed on the use of the pediatric scales over time.²⁷

Limitations. The continued work of the REiNS Collaboration and the broader scientific community will be important in addressing the current limitations associated with these recommendations. First is the challenge of whether these PRO tools accurately assess the specific types of pain and physical functions being evaluated in clinical trials for individuals with NF when multiple factors may contribute to their pain experience or impact their daily functioning. Another issue is the lack of data in NF1, NF2, and SWN and whether the psychometric properties of the scale in these conditions will be consistent with those documented in the general population. Additionally, no PRO tools to date have shown the ability to measure pain and physical functioning adequately for all ages. In

particular, researchers need to address the evaluation of these domains in children under 8 years old since NF trials are enrolling these younger participants. This challenge is reflected in our recommendation of 2 pain interference scales, as currently there are insufficient published data for establishing the validity of either measure in children and adults with NF. Identifying the most appropriate methods to assess children with significant cognitive impairments also needs to be addressed. Our group will update our recommendations in these specific areas in future publications.

To address the limitations noted above, PRO group members are using these recommended pain and physical functioning measures in descriptive studies and clinical trials with a wide age range of individuals who have all types of NF. Several of these PROs have been accepted into the FDA's Drug Development Tool COA Qualification Program¹² and are being further studied with both qualitative and quantitative research methods¹¹ to examine possible modifications and their use as NF clinical trial endpoints. The PRO working group also plans to form a subgroup that will focus on identifying measures to assess young children and individuals with significant cognitive impairments. Finally, researchers need to investigate the use of new methodologies in NF clinical trials to improve PRO data, such as using electronic administration to increase the accuracy and speed of data collection.³⁹

The REiNS PRO group is working to identify PRO measures that can be used consistently as endpoints across trials to assess clinical benefit of treatments to reduce tumors or other disease manifestations in individuals with NF1, NF2, and SWN. The REiNS recommendations for PROs in the domains of pain and physical functioning are presented with the understanding that further evaluation of the measures is needed in the NF population. Our group encourages researchers to use these PRO measures in upcoming NF studies. Final consensus recommendations regarding the one specific pain interference measure and format of the PROMIS-PF scales to use in NF clinical trials will be disseminated in a future publication based on results of additional published research.

AUTHOR CONTRIBUTIONS

Pamela Wolters: design and conceptualization of the study, collection and interpretation of the data, drafting and revising the manuscript. Staci Martin: design and conceptualization of the study, collection and interpretation of the data, drafting and revising the manuscript. Vanessa Merker: study concept, collection and interpretation of the data, drafting and revising the manuscript. Jim Tongsgard: study concept, collection and interpretation of the data, drafting and revising the manuscript. Sondra Solomon: study concept, collection and interpretation of the data, drafting and revising the manuscript. Andrea Baldwin: study concept, collection and interpretation of the data, drafting and revising the manuscript. Amanda Bergner: study concept, collection and interpretation of the data, drafting and revising the manuscript. Karin Walsh: study concept, collection and interpretation of the data, drafting and revising the manuscript.

Comment: Clinical outcome assessments in transition

The Response Evaluation in Neurofibromatosis and Schwannomatosis (REiNS) International Collaboration systematically evaluated existing patient-reported outcome (PRO) measures of pain intensity, pain interference, and physical functioning as potential primary and secondary outcomes for neurofibromatosis trials. On the basis of 4 criteria—patient characteristics, item content, psychometric properties, and feasibility—specific measures were recommended.¹ These recommendations are valuable for investigators designing clinical trials and can serve as a basis for developing improved outcome assessments for patients with these disorders.

In 2006, the US Food and Drug Administration (FDA) released a draft guidance on PRO measures that initiated a broader consideration of clinical outcome assessments (COAs) of “a patient’s symptoms, overall mental state, or the effects of a disease or condition on how the patient functions.”² COAs are used in evaluating efficacy and safety, and include patient-reported, clinician-reported, observer-reported, and performance outcome measures. Such measures should be used within a carefully defined context of use for which available data provide support.² The FDA recently released a compendium of COAs that have either been used as a basis for regulatory decisions or are being developed in accord with its qualification program.³

In evaluating COAs, the FDA examines their content validity (e.g., as assessed by patient focus groups and cognitive interviewing), reliability, construct validity, and longitudinal measurement properties (e.g., ability to detect change, responder criteria).² The REiNS recommendations were based on some of these characteristics, but most existing outcome measures fall short of the rigorous standards now being applied by the FDA. Indeed, as the REiNS investigators acknowledge, specification of the appropriate context of use (e.g., patient age, specific disorders) for pain and physical functioning measures in neurofibromatosis trials requires additional data. Systematic efforts to improve the validity, reliability, and responsiveness to change of COAs will ensure that clinical trial outcomes provide meaningful information about the treatment benefits and risks that patients consider most important.

1. Wolters PL, Martin S, Merker VL, et al. Patient-reported outcomes of pain and physical functioning in neurofibromatosis clinical trials. *Neurology* 2016;87(suppl 1):S4–S12.
2. US Food and Drug Administration. Clinical outcome assessment qualification program. Available at: <http://www.fda.gov/Drugs/DevelopmentApprovalProcess/DrugDevelopmentToolsQualificationProgram/ucm284077.htm>. Accessed June 16, 2016.
3. US Food and Drug Administration. Clinical outcome assessment compendium (COA compendium): version 1. Available at: <http://www.fda.gov/downloads/Drugs/DevelopmentApprovalProcess/DevelopmentResources/UCM481225.pdf>. Accessed June 16, 2016.

Robert H. Dworkin, PhD

From the Department of Anesthesiology, University of Rochester School of Medicine and Dentistry, NY. Study funding: No targeted funding reported.

Disclosure: Dr. Dworkin has received, in the past 12 months, research grants and contracts from the US Food and Drug Administration and US NIH, and compensation for activities involving clinical trial research methods from Abide, Adynxx, Aptinyx, Astellas, Biogen, Dong-A, Glenmark, Hope, Hydra, Immune, Medavante, Novartis, NsGene, Olatec, Periphagen, Phosphagenics, Reckitt Benckiser, Relmada, Semnur, Spinifex, Syntrix, Teva, Trevena, and Vertex. Go to Neurology.org for full disclosures.

Heather Thompson: study concept, collection and interpretation of the data, drafting and revising the manuscript. Kathy Gardner: study concept, collection and interpretation of the data, revising the manuscript. Cynthia Hingtgen: study concept, collection and interpretation of the data, revising the manuscript. Elizabeth Schorry: study concept, collection and interpretation of the data, revising the manuscript. William Dudley: study concept, collection and interpretation of the data, revising the manuscript. Barbara Franklin: study concept, collection and interpretation of the data, revising the manuscript.

ACKNOWLEDGMENT

The authors thank Ana-Maria Vranceanu, PhD, Department of Psychiatry, Harvard Medical School, a new member of the REINS PRO working group, for participation in selected conference calls to discuss reviewed PRO measures.

STUDY FUNDING

This research is supported by the Intramural Research Program of the National Institutes of Health, National Cancer Institute.

DISCLOSURE

P. Wolters received funding from the Neurofibromatosis Therapeutics Acceleration Program and holds stock in Bristol-Meyers-Squibb. S. Martin received funding from the Neurofibromatosis Therapeutics Acceleration Program. She also received funded travel for speaking engagements from the University of Chicago and the Children's Tumor Foundation. V. Merker reports no disclosures relevant to the manuscript. J. Tonsgard receives funding from the Department of Defense Army Research Command as part of the NF Consortium and from Midwest NF, Inc. S. Solomon is deceased; she reported no disclosures relevant to the manuscript. A. Baldwin and A. Bergner report no disclosures relevant to the manuscript. K. Walsh received funding from the Children's Tumor Foundation, the Department of Defense, and the Gilbert Neurofibromatosis Institute. H. Thompson and K. Gardner report no disclosures relevant to the manuscript. C. Hingtgen holds stock in Merck. B. Schorry receives funding from the Department of Defense as a site PI for the NF Consortium and also as a coinvestigator for a clinical trial of vitamin D for adults with NF1. W. Dudley and B. Franklin report no disclosures relevant to the manuscript. Go to Neurology.org for full disclosures.

Received August 19, 2015. Accepted in final form April 12, 2016.

REFERENCES

1. Korf BR. Neurofibromatosis. *Handb Clin Neurol* 2013; 111:333–340.
2. Eleftheriou F, Kolanczyk M, Schindeler A, et al. Skeletal abnormalities in neurofibromatosis type 1: approaches to therapeutic options. *Am J Med Genet A* 2009;149A:2327–2338.
3. Nguyen R, Kluwe L, Fuensterer C, Kentsch M, Friedrich RE, Mautner VF. Plexiform neurofibromas in children with neurofibromatosis type 1: frequency and associated clinical deficits. *J Pediatr* 2011;159:652–655.
4. Wolters PL, Burns KM, Martin S, et al. Pain interference in youth with neurofibromatosis type 1 and plexiform neurofibromas and relation to disease severity, social-emotional functioning, and quality of life. *Am J Med Genet A* 2015;167A:2103–2113.
5. Creange A, Zeller J, Rostaing-Rigattieri S, et al. Neurological complications of neurofibromatosis type 1 in adulthood. *Brain* 1999;122:473–481.
6. Merker VL, Bredella MA, Cai W, et al. Relationship between whole-body tumor burden, clinical phenotype, and quality of life in patients with neurofibromatosis. *Am J Med Genet A* 2014;164A:1431–1437.
7. Widemann BC, Babovic-Vuksanovic D, Dombi E, et al. Phase II trial of pirfenidone in children and young adults with neurofibromatosis type 1 and progressive plexiform neurofibromas. *Pediatr Blood Cancer* 2014;61:1598–1602.
8. Martin S, Wolters P, Toledo-Tamula MA, et al. Acceptance and commitment therapy in youth with neurofibromatosis type 1 (NF1) and chronic pain and their parents: a pilot study of feasibility and preliminary efficacy. *Am J Med Genet* 2016;170:1462–1470.
9. Kim A, Dombi E, Tepas K, et al. Phase I trial and pharmacokinetic study of sorafenib in children with neurofibromatosis type 1 and plexiform neurofibromas. *Pediatr Blood Cancer* 2013;60:396–401.
10. Acquadro C, Berzon R, Dubois D, et al. Incorporating the patient's perspective into drug development and communication: an ad hoc task force report of the Patient-Reported Outcomes (PRO) Harmonization Group meeting at the Food and Drug Administration, February 16, 2001. *Value Health* 2003;6:522–531.
11. US Department of Health and Human Services FDA Center for Drug Evaluation and Research; US Department of Health and Human Services FDA Center for Biologics Evaluation and Research; US Department of Health and Human Services FDA Center for Devices and Radiological Health. Guidance for Industry: patient-reported outcome measures: use in medical product development to support labeling claims. *Health Qual Life Outcomes* 2006;4:79.
12. US Food and Drug Administration. Clinical Outcome Assessment Compendium. Mary Land: US Food and Drug Administration; 2016.
13. Hoa M, Slattery WH III. Neurofibromatosis 2. *Otolaryngol Clin North Am* 2012;45:315–332.
14. Cleeland CS, O'Mara A, Zagari M, Baas C. Integrating pain metrics into oncology clinical trials. *Clin Cancer Res* 2011;17:6646–6650.
15. Hyman SL, Shores A, North KN. The nature and frequency of cognitive deficits in children with neurofibromatosis type 1. *Neurology* 2005;65:1037–1044.
16. Wolters PL, Martin S, Merker VL, et al. Patient-reported outcomes in neurofibromatosis and schwannomatosis clinical trials. *Neurology* 2013;81:S6–S14.
17. McClimans LM, Browne J. Choosing a patient-reported outcome measure. *Theor Med Bioeth* 2011;32:47–60.
18. Miro J, Castarlenas E, Huguet A. Evidence for the use of a numerical rating scale to assess the intensity of pediatric pain. *Eur J Pain* 2009;13:1089–1095.
19. Martin S, Nelson Schmitt S, Wolters PL, et al. Development and validation of the English pain interference index and pain interference index-parent report. *Pain Med* 2015; 16:367–373.
20. Amtmann D, Cook KF, Jensen MP, et al. Development of a PROMIS item bank to measure pain interference. *Pain* 2010;150:173–182.
21. Cleeland CS, Ryan KM. Pain assessment: global use of the brief pain inventory. *Ann Acad Med Singapore* 1994;23: 129–138.
22. Holmstrom L, Kemani MK, Kanstrup M, Wicksell RK. Evaluating the statistical properties of the pain interference index in children and adolescents with chronic pain. *J Dev Behav Pediatr* 2015;36:450–454.
23. Kemani M, Zetterqvist V, Kanstrup M, Wicksell R, Holmström L. A validation of the pain interference index in adults with longstanding pain. *Acta Anaesthesiol Scand* 2016;60:250–258.
24. Bevans M, Ross A, Cella D. Patient-Reported Outcomes Measurement Information System (PROMIS): efficient, standardized tools to measure self-reported health and quality of life. *Nurs Outlook* 2014;62:339–345.

25. Rose M, Bjorner JB, Becker J, Fries JF, Ware JE. Evaluation of a preliminary physical function item bank supported the expected advantages of the Patient-Reported Outcomes Measurement Information System (PROMIS). *J Clin Epidemiol* 2008;61:17–33.
26. Broderick JE, Schneider S, Junghaenel DU, Schwartz JE, Stone AA. Validity and reliability of patient-reported outcomes measurement information system instruments in osteoarthritis. *Arthritis Care Res* 2013;65:1625–1633.
27. DeWalt DA, Gross HE, Gipson DS, et al. PROMIS pediatric self-report scales distinguish subgroups of children within and across six common pediatric chronic health conditions. *Qual Life Res* 2015;24:2195–2208.
28. Shahgholi L, Yost KJ, Kallmes DF. Correlation of the National Institutes of Health patient reported outcomes measurement information system scales and standard pain and functional outcomes in spine augmentation. *AJNR Am J Neuroradiol* 2012;33:2186–2190.
29. Mulcahey MJ, Haley SM, Slavin MD, et al. Ability of PROMIS pediatric measures to detect change in children with cerebral palsy undergoing musculoskeletal surgery. *J Pediatr Orthop Epub* 2015 Jun 5.
30. Varni JW, Thissen D, Stucky BD, et al. PROMIS(R) Parent Proxy Report Scales: an item response theory analysis of the parent proxy report item banks. *Qual Life Res* 2012;21:1223–1240.
31. Bruce B, Fries JF. The Health Assessment Questionnaire (HAQ). *Clin Exp Rheumatol* 2005;23:S14–S18.
32. Walker LS, Greene JW. The Functional Disability Inventory: measuring a neglected dimension of child health status. *J Pediatr Psychol* 1991;16:39–58.
33. DeWitt EM, Stucky BD, Thissen D, et al. Construction of the eight-item patient-reported outcomes measurement information system pediatric physical function scales: built using item response theory. *J Clin Epidemiol* 2011;64:794–804.
34. Jensen RE, Potosky AL, Reeve BB, et al. Validation of the PROMIS physical function measures in a diverse US population-based cohort of cancer patients. *Qual Life Res* 2015;24:2333–2344.
35. Varni JW, Magnus B, Stucky BD, et al. Psychometric properties of the PROMIS (R) pediatric scales: precision, stability, and comparison of different scoring and administration options. *Qual Life Res* 2014;23:1233–1243.
36. Rose M, Bjorner JB, Gandek B, Bruce B, Fries JF, Ware JE Jr. The PROMIS Physical Function item bank was calibrated to a standardized metric and shown to improve measurement efficiency. *J Clin Epidemiol* 2014; 67:516–526.
37. Fries JF, Krishnan E, Rose M, Lingala B, Bruce B. Improved responsiveness and reduced sample size requirements of PROMIS physical function scales with item response theory. *Arthritis Res Ther* 2011;13:R147.
38. Kratz AL, Slavin MD, Mulcahey MJ, Jette AM, Tulskey DS, Haley SM. An examination of the PROMIS(R) pediatric instruments to assess mobility in children with cerebral palsy. *Qual Life Res* 2013;22: 2865–2876.
39. Palermo TM, Valenzuela D, Stork PP. A randomized trial of electronic versus paper pain diaries in children: impact on compliance, accuracy, and acceptability. *Pain* 2004; 107:213–219.
40. von Baeyer CL. Numerical rating scale for self-report of pain intensity in children and adolescents: recent progress and further questions. *Eur J Pain* 2009;13:1005–1007.
41. Castarlenas E, Miro J, Sanchez-Rodriguez E. Is the verbal numerical rating scale a valid tool for assessing pain intensity in children below 8 years of age? *J Pain* 2013;14:297–304.
42. Bailey B, Daoust R, Doyon-Trottier E, Dauphin-Pierre S, Gravel J. Validation and properties of the verbal numeric scale in children with acute pain. *Pain* 2010;149:216–221.
43. Good M, Stiller C, Zauszniewski JA, Anderson GC, Stanton-Hicks M, Grass JA. Sensation and distress of pain scales: reliability, validity, and sensitivity. *J Nurs Meas* 2001;9:219–238.
44. Varni JW, Thissen D, Stucky BD, et al. Item-level informant discrepancies between children and their parents on the PROMIS pediatric scales. *Qual Life Res* 2015;24: 1921–1937.