A Scale to Assess Activities of Daily Living in Pantothenate Kinase-Associated Neurodegeneration

Randall D. Marshall, MD,[1,*] Abigail Collins, MD,[2] Maria L. Escolar, MD,[3] H.A. Jinnah, MD, PhD,[4] Thomas Klopstock, MD,[5] iD
Michael C. Kruer, MD,[6] iD Aleksandar Videnovic, MD,[7] Amy Robichaux-Viehoever, MD, PhD,[8] Laura Swett, MSW, PhD,[9]
Dennis A. Revicki, PhD,[9] iD Randall H. Bender, PhD,[10] and William R. Lenderking, PhD[9]

ABSTRACT: Objective: Pantothenate kinase-associated neurodegeneration (PKAN) is an autosomal-recessive, neurodegenerative disorder with a mixed-motor phenotype caused by a defective PanK2 enzyme, for which there are few adequate treatment options. Clinimetrically sound measures of patient-reported outcomes are necessary to facilitate therapeutic development for this debilitating disease. This study's objective was to develop such a scale and assess its clinimetric properties.

Methods: A conceptually driven, iterative, content development process incorporating input from experts, caregivers, and patients was used. Scale items were initially adapted from the Unified Parkinson's Disease Rating Scale (UPDRS) Part II resulting in the 12-item Pantothenate Kinase-Associated Neurodegeneration Activities of Daily Living (PKAN-ADL). The PKAN-ADL scale was administered to caregivers (n = 37) and patients (n = 2) twice over 2 weeks, along with selected Quality of Life in Neurological Disorders (Neuro-QoL) measures, selected attributes of the Health Utilities Index (HUI)-2/3, and the Stroke Aphasia Depression Questionnaire (SADQ-10) to assess construct validity.

Results: Internal consistency was 0.93, with excellent test-retest reliability (intraclass correlation coefficient = 0.99). Of the 12 items, 25% (n = 3) showed a ceiling effect >30% (range, 31–54) and 42% (n = 5) showed a floor effect >30% (range, 31–46), reflecting disease heterogeneity. Convergent validity was shown with Neuro-QoL measures (rs > 0.90) and HUI-2/3 attributes ($rs \geq$ 0.48); divergent validity was demonstrated with the SADQ-10 (r = 0.11). Participants reported a high level of comprehension (98%), and average item relevance ratings (0–10 scale) ranged from 7.0 to 9.9.

Conclusion: The PKAN-ADL scale demonstrated acceptable content validity, with evidence of construct validity and excellent reliability. Overall results support the use of the PKAN-ADL scale in clinical trials.

Pantothenate kinase–associated neurodegeneration (PKAN) is an autosomal–recessive disorder caused by mutations in the *PANK2* gene, with a variable clinical phenotype and a prevalence of 1 to 2 per million persons worldwide.[1] Clinical manifestations include focal and generalized dystonia, parkinsonism, dysarthria/anarthria,

and dysphagia.[1] The genotype/phenotype association is not well understood, in that key features such as rate of progression, age of onset, and signs and symptoms are highly variable, even among siblings and case clusters with identical mutations.[2] Currently, there are no approved disease-modifying therapies for PKAN,

[1]Formerly of Retrophin, Inc., Cambridge, Massachusetts, USA; [2]Departments of Pediatrics and Neurology, University of Colorado School of Medicine, Denver, Colorado, USA; [3]Department of Pediatrics, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA; [4]Departments of Neurology and Human Genetics, Emory University School of Medicine, Atlanta, Georgia, USA; [5]Department of Neurology, Friedrich-Baur-Institute, University of Munich, Munich, Germany and German Center for Neurodegenerative Diseases (DZNE), Munich, Germany; [6]Barrow Neurological Institute, Phoenix Children's Hospital, University of Arizona College of Medicine, Phoenix, Arizona, USA; [7]Department of Neurology, Massachusetts General Hospital/Harvard Medical School, Boston, Massachusetts, USA; [8]Departments of Neurology and Pediatrics, University of California-San Francisco, San Francisco, California, USA; [9]Evidera, Inc, Bethesda, Maryland, USA; [10]Formerly of Evidera, Inc, Bethesda, Maryland, USA

*Correspondence to: Dr. Randall D. Marshall, Neuroscience Therapeutic Area Lead, Retrophin, Inc., 3721 Valley Center Drive, Suite 200, San Diego, CA 92130, USA; E-mail: randallmarshall4@aol.com
Keywords: pantothenate kinase-associated neurodegeneration, activities of daily living, clinimetric.
Relevant disclosures and conflicts of interest are listed at the end of this article.
Received 19 September 2018; revised 9 November 2018; accepted 11 November 2018.
Published online 22 January 2019 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/mdc3.12716

and the benefits from symptom-targeted treatments, such as deep brain stimulation to decrease dystonia, are usually not sustained as PKAN progresses.[3] There is an urgent need for the development of disease-modifying treatments with sustained benefit for patients with PKAN.

Clinimetrically sound measures of patient-reported outcomes are necessary to facilitate therapeutics development. A comprehensive clinician rating scale (the PKAN-Disease Rating Scale [PKAN-DRS]) was recently developed to assess the multiple domains of PKAN.[4] The scale combines scores generated by a rater interview to assess domains of cognition, behavior, and disability with scores based on neurological assessment of dystonia, parkinsonism, and other neurological signs.

Regulatory agencies have increasingly promoted a central role for clinical outcomes assessments (COAs) that directly measure the effects of disease on health and functioning from the patient perspective[5,6] to supplement medical examination, laboratory, or biomarker measures that may require inferential assumptions about patient benefit. Thus, the PKAN-Activities of Daily Living (PKAN-ADL) scale was developed as a COA of PKAN-related patient functioning in everyday life, based on patient or caregiver self-report, for use in clinical trials.

This study had two objectives: (1) create an instrument measuring functional capacity and ADLs in PKAN patients; (2) determine its clinimetric properties using methods consistent with the U.S. Food and Drug Administration guidance on Patient-Reported Outcomes development[6] and COA qualification guidance,[7] and with advice received from the Committee for Medicinal Products for Human Use at the European Medicines Agency.

Patients and Methods

Motor manifestations of PKAN can be viewed as a mixture of dystonia and parkinsonism.[8] Ophthalmic and cognitive abnormalities are variable, and assessment of cognition can be confounded by communication impairment attributed to oromandibular dystonia. Thus, based on literature review and expert clinical input, it was decided to focus on functional consequences of motor symptoms as the major contributor to impairment in patients with PKAN. In addition, assessing functional impairment does not require parsing of symptom severity into dystonic or parkinsonian causes, which can be difficult and potentially unreliable across raters.[4,7]

Among the large number of functional disability scales for movement disorders,[9] Part II of the Unified Parkinson's Disease Rating Scale (UPDRS) was selected as a starting point for scale development because of its widespread use in clinical trials and its past use by physicians treating PKAN patients with fosmetpantotenate.[10,11]

Stage 1. Content Development

An iterative process was followed to adapt individual scale items from the UPDRS Part II to be PKAN specific[12] or to remove or add specific items. Scale items were initially revised based on literature review and author input. Systematic interviews were then conducted with an international group of experts, patient advocacy leaders, and caregivers to evaluate the relevance, clarity, and overall inclusiveness of content. Item language and content were revised, followed by a second round of interviews that led to the final draft version of the PKAN-ADL.

Key questions posed to participants throughout the process included: How relevant is the item/domain to patients' daily functioning? Would improvement by one level on an item be meaningful in patients' lives? Are there key aspects of disability in PKAN not covered by the items?

Stage 2. Assessment of Content Validity and Clinimetric Properties

A clinimetric evaluation of the draft PKAN-ADL was conducted using interviewer-administered study questionnaires and semi-structured interviews by telephone at baseline and approximately 2 weeks later. The interview included an open-ended section about symptoms and medical history and a more structured evaluation of the PKAN-ADL items.

Participants were recruited by the Neurodegeneration with Brain Iron Accumulation Disorders Association, clinicians treating PKAN patients, family networks, and social media sites until the prespecified study enrollment target was met (approximately n = 40). The sample size was based on the central limit theorem (n = 30), with an additional 30% to allow for possible study dropout. A sample of 30 to 40 is often considered sufficient for test-retest reliability evaluation. The ideal sample size, however, often depends on parameters such as the expected reliability, significance level, power, retest interval, etc., that are not always incorporated in retest studies.[13]

The study protocol received prior approval by Ethical and Independent Review Services, a central institutional review board (IRB). Once written informed consent and assent were obtained, the interviewer administered a patient and/or caregiver sociodemographic information form, a brief disease-specific history form, and a series of questionnaires (study measures are listed below) to the caregiver and/or patient. The interviewer recorded participant responses and the interviews were audio recorded.

Caregiver and Patient Inclusion and Exclusion Criteria

Caregiver inclusion criteria were: (1) caregiver of a patient age 6 years or older with a genetically confirmed diagnosis of PKAN; (2) caregiver age 18 years or older and willing and able to provide consent over the telephone; and (3) caregiver able to read and speak English. Caregivers were excluded if they had any clinically relevant physical or mental conditions that would interfere with participation in the study.

Patient inclusion criteria were: (1) easily understandable speech over the phone to a stranger (based on caregiver report and confirmed by interviewer); (2) age 16 years or older; (3) willing and able to provide written informed consent if over 18, or

if under the age of 18, willing to provide assent for study participation and have written informed consent provided by his or her caregiver; (4) able to read and speak English; (5) a genetically confirmed diagnosis of PKAN (based on caregiver report); and (6) willing and able to be interviewed in the presence of a caregiver who is 18 or older, if necessary. Patients were excluded if they had any signs or symptoms of cognitive impairment or other mental illness (based on caregiver assessment) or any other clinically relevant physical or mental conditions that would interfere with study participation.

Study Measures

Study measures included the PKAN-ADL and a semistructured interview designed to gauge impressions about the PKAN-ADL, collect information about patients and caregivers (sociodemographic questionnaire, PKAN medical history form), and administer measures to assess construct validity (selected Quality of Life in Neurological Disorders [Neuro-QoL] measures, selected attributes of the Health Utilities Index Mark 3 [HUI-3], and the Stroke Aphasic Depression Questionnaire [SADQ-10]).

Sociodemographic Questionnaires

Sociodemographic questionnaires included items on age, sex, race/ethnicity, and employment status.

PKAN-Specific Medical History Form

The PKAN-specific medical history form collected basic clinical information about the patient's experience with PKAN, including age of onset, illness duration, source of diagnosis, milestones of disease progression, and treatments received.

HUI-3: Proxy Version

The HUI-3 is a preference-based measure of health-related quality of life (HRQoL).[14] We selected four attributes of the HUI-2/3 (a proxy version of the HUI-3) for their relevance to PKAN: speech, ambulation, dexterity, and pain.

PKAN-ADL Scale

The final draft PKAN-ADL, adapted based on the UPDRS Part II, assessed 12 domains of ADLs: difficulty with speech; drooling; swallowing; writing; eating tasks; dressing; personal hygiene tasks; turning or changing position in bed; sitting; falling; walking; and discomfort or pain (Supporting Information Appendix S1). The 5-point scale response options for the PKAN-ADL ranged from 0 (indicating "normal [no problems]") to 4 (indicating "inability to perform the activity").

Neuro-QoL Version 1.0

The Neuro-QoL is a validated set of self-reported measures developed by the National Institute of Neurologic Disorders and Stroke that assess the HRQoL of adults and children with neurological disorders, including symptoms, concerns, and problems with functioning.[12] Two item banks (adult versions) were selected as relevant to PKAN: the upper extremity function (20 items) and the lower extremity function (19 items).

SADQ-10

The short-form version of the SADQ-10[15] was used in this study. The SADQ-10 contains 10 items assessing observable behaviors indicative of depressed mood. The SADQ-10 was developed specifically for patients unable to communicate effectively, which describes a large proportion of patients with PKAN.[16] It was selected to assess divergent validity because the symptoms and functional limitations assessed with the PKAN-ADL are associated with ADLs and not emotional or behavioral disorders.

Measure of Disease Stability Between Interviews

For interpretation of test-retest reliability, disease stability was assessed at the second interview by asking "How is the patient's [your] condition now compared to the time of the first interview?" and included three response options: "the same," "worse," or "improved."

Qualitative Interview Guide

Participants were asked about their overall impression of the PKAN-ADL, the extent to which they understood the questions in the PKAN-ADL, if any ADL-related content was missing, and to rate the relevance of each item (cognitive debriefing) on a scale of 0 to 10, where 0 indicated "not at all relevant" and 10 indicated "extremely relevant."
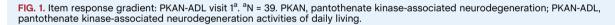
Data Analysis

The data analysis explored descriptive statistics (N, mean, standard deviation [SD], minimum, maximum, and floor and ceiling effects) and the reliability and validity of the PKAN-ADL. Internal consistency was evaluated using Cronbach's coefficient alpha. Analyses for test-retest reliability included random-effects analysis of variance (ANOVA) and intraclass correlations coefficients (ICCs) between the baseline mean and retest mean (approximately 2 weeks postbaseline) of the PKAN-ADL scores.

Construct validity was examined through correlations between the PKAN-ADL and Neuro-QoL and the selected HUI-3 item scores. Divergent validity was assessed by evaluating the correlation between the PKAN-ADL and SADQ-10. Criterion validity was assessed by regressing duration of illness on total PKAN-ADL score using ordinary least squares linear regression.

Item response theory analysis for evaluating whether PKAN-ADL items were ordered and well functioning was not possible

Participant ID	Salivation and drooling	Turning or changing position in bed	Discomfort or pain related to PKAN	Chewing, swallowing, and choking	Sitting	Eating tasks	Hygiene	Dressing	Speech	Problems walking independently	Writing	Falling	PKAN-ADL Total Score
1012	0	0	0	0	0	0	0	0	1	0	2	1	4
1005	0	0	0	1	0	0	0	0	0	1	2	1	5
1010	0	0	0	0	1	1	1	0	3	0	2	2	10
1026	0	0	0	0	0	1	1	0	3	1	2	3	11
1032	3	0	3	0	0	0	0	0	2	1	2	1	12
1008	0	0	0	1	0	2	0	2	3	1	2	2	13
1016	3	0	3	1	0	1	0	0	3	0	1	1	13
1024	0	0	0	1	1	2	1	2	1	2	2	2	14
1041	1	0	3	1	0	1	0	1	1	1	2	3	14
1015	2	0	1	1	1	1	0	0	3	2	2	2	15
1001	0	0	2	1	1	2	1	1	1	3	1	3	16
1023	0	0	0	1	0	1	1	2	2	3	3	3	16
1027	0	0	0	1	0	1	1	2	2	3	3	3	16
1034	0	0	1	1	2	3	3	2	0	2	0	2	16
1013	0	0	4	0	1	2	2	2	0	2	3	1	17
1011	2	1	1	1	1	1	2	3	3	1	1	1	18
1029	0	0	1	0	1	2	3	4	2	1	3	3	20
1036	0	0	4	2	2	1	1	2	3	2	3	2	22
1007	0	1	3	1	3	1	3	2	2	3	2	3	24
1040	1	0	2	2	2	3	2	1	3	3	3	3	25
1002	2	0	0	2	1	2	2	2	3	4	4	4	26
1020	1	1	0	1	1	2	4	3	2	3	3	3	26
1035	0	3	0	1	3	3	4	4	2	3	2	2	27
1014	4	1	1	1	2	1	3	3	4	3	3	3	29
1031	1	4	0	2	3	1	4	4	1	4	3	4	31
1017	3	0	2	2	3	3	4	4	3	3	3	3	33
1033	0	4	3	1	3	3	4	4	3	4	3	4	36
1037	0	4	1	1	4	4	4	4	2	4	4	4	36
1019	0	3	3	3	3	4	4	4	2	4	3	4	37
1022	3	1	3	3	4	4	4	4	4	4	4	4	38
1038	2	1	0	4	3	4	4	4	4	4	4	4	38
1003	4	1	3	4	3	4	4	4	4	4	1	4	40
1042	2	3	2	4	3	4	4	4	4	4	3	4	41
1004	2	3	2	4	3	4	4	4	4	4	4	4	42
1018	0	4	4	4	3	4	4	4	4	4	3	4	42
1006	3	3	2	4	4	4	4	4	4	4	4	4	44
1028	4	4	1	3	4	4	4	4	4	4	4	4	44
1030	2	4	3	4	4	4	4	4	4	4	4	4	45
1039	4	4	2	4	4	4	4	4	4	4	4	4	46
MEAN	1.1	1.3	1.6	1.7	1.9	2.3	2.4	2.5	2.6	2.7	2.7	2.9	25.8

FIG. 1. Item response gradient: PKAN-ADL visit 1[a]. [a]N = 39. PKAN, pantothenate kinase-associated neurodegeneration; PKAN-ADL, pantothenate kinase-associated neurodegeneration activities of daily living.

because of the small sample size. Therefore, an item response gradient was developed to evaluate item functioning (Fig. 1). Individual patient scores were ordered by total score, from lower to higher totals down the rows. Columns were also ordered from items with the lowest mean up to the highest mean, moving from left to right. The table visually represents how severity of impairment for individual items relates to overall severity of impairment of functioning (total score).

Qualitative Data Analysis

Qualitative analyses followed the principles of thematic narrative analysis, with the goal of identifying themes associated with participant feedback on the PKAN-ADL.[17] The qualitative data were also quantified by counting thematic categories of coded responses for each item such as "understood the question as intended" and "would recommend alternate wording."

Results

Stage 1. Content Development

The item development process resulted in retention of the five-level response scale structure of the UPDRS Part II. Through the iterative process, changes were made in formatting and language for clarity and to create a parallel structure across items, and clear examples for each item domain were generated. One item domain was added (getting into and out of a chair) and two were removed (tremor, freezing when walking). Tremor was removed because it was uncommon in PKAN patients. Although freezing is reported to be a common symptom in PKAN,[4] "freezing when walking" was removed because patient advocates and family members had difficulty separating falling attributed to freezing from falling attributed to other causes, or no obvious cause other than gait instability. It was decided that one item that assessed falling was sufficient to capture functional impairment related to falling, regardless of presumed cause. The

TABLE 1 *Sociodemographic characteristics*

Characteristic	Caregiver Interviewees (N = 37)	All Participants (Includes Patients) (N = 39)
Age, years		
Mean (SD)	46.8 (12.5)	20.2 (8.4)
Range [min, max]	[24.1–67.6]	[6.4–42.6]
Missing, n (%)	2 (5.4)	1 (2.6)
Relationship to the patient, n (%)		
Parent	33 (89.2)	N/A
Other relative	1 (2.7)	N/A
Professional caregiver	1 (2.7)	N/A
Other	2 (5.4)	N/A
Caregiver lives with patient, n (%)		
Yes	31 (83.8)	N/A
No	6 (16.2)	N/A
Sex, n (%)		
Male	11 (29.7)	24 (61.5)
Female	26 (70.3)	15 (38.5)
Racial background, n (%)		
White	33 (89.2)	35 (89.7)
Black or African American	0 (0.0)	0 (0.0)
Asian	3 (8.1)	2 (5.1)
Native Hawaiian or other Pacific Islander	1 (2.7)	1 (2.6)
Missing, n (%)	0 (0.0)	1 (2.6)
Employment status, n (%)		
Employed, full-time or part-time	28 (75.7)	4 (10.3)
Homemaker	4 (10.8)	0 (0.0)
Student	0 (0.0)	19 (48.7)
Unemployed	0 (0.0)	4 (10.3)
Retired	4 (10.8)	0 (0.0)
Disabled	1 (2.7)	12 (30.8)
Change in employment status attributed to caregiving, n (%)		
Yes	21 (56.8)	N/A
No	16 (43.2)	N/A

N/A, not applicable.

final PKAN-ADL assessed 12 aspects of motor functioning, one item per aspect.

Stage 2. Data Analysis

Demographic and Clinical Characteristics of the Study Participants

The sample included 37 caregivers and 2 patients residing in the United States (n = 35), Canada (n = 3), and Germany (n = 1). Participants' sociodemographic data is listed in Table 1.

Patients (n = 39) had a mean (SD) age of 20.2 years (8.4) and were mostly white (89.7%) and male (61.5%). Mean age (SD) at symptom onset was 8.0 years (5.8), but with a wide range of <1 to 20 years old. The most common first symptoms or functional limitations were difficulties with walking (69.2%), speech (30.8%), and writing (23.1%), followed by emotional and behavioral problems (15.4%). Dystonia at onset of PKAN symptoms was reported in 15.4% of patients, including involvement of hand (15.4% of patients), mouth/tongue (10.3%), back/trunk (7.7%), foot (5.1%), and neck (2.6%).

Qualitative Results: Content Validity

Most participants thought that the PKAN-ADL was a good, comprehensive instrument, and 97% could provide an accurate paraphrase of items, indicating they understood each item as it was intended. Caregivers and patients rated the relevance of each item on a scale of 0 to 10, where 0 indicated "not at all relevant" and 10 indicated "extremely relevant." Average relevance ratings for each item were high, with most items reflecting a range of 8.5 to 10 (Table 2). Lower relevance ratings were noted for Item 2 (Salivation/Drooling = 6.8) and Item 8 (Turning/Changing Position in Bed = 7.6); the highest relevance rating was for Item 10 (Falling = 10.0). The lower relevance rating for Salivation/Drooling may be attributed to effective control of symptoms through medication, as indicated by several participants in response to the key questions.

Whereas additional content for the instrument was recommended by 41% of the sample, most of the recommendations included symptoms of the disease (including fatigue/stamina, bladder and colon control, "ability to sit or stand straight," "biting your tongue," "neck issues from dystonia," and "tremors in hand") and were not specifically in the ADL domain. Suggestions for additions also included cognitive ("ability to concentrate or focus on a task") or emotional/behavioral categories of problems, such as emotional well-being, mood swings/emotional outbursts, anger, and compulsive behaviors such as "pulling fringe out of pillows," "pulling out hair," and "sexual drive" (referring to self-stimulation and viewing of pornography). Additions for social impacts were also suggested ("Does he have any friends now?" and "Are social interactions appropriate?"). These were rejected because the PKAN-ADL is intended to measure motor-related impairments.

TABLE 2 Content validation[1]

	Patient/caregiver	Frequency/Mean	%
Question 1 says: speech	Understood question as intended	39	100
	Relevance on scale from 0 to 10	9.5	N/A
	Would recommend alternate wording	1	3
Question 2 says: salivation and drooling	Understood question as intended	38	97
	Relevance on scale from 0 to 10	6.8	N/A
	Would recommend alternate wording	4	10
Question 3 says: chewing, swallowing, and choking	Understood question as intended	38	97
	Relevance on scale from 0 to 10	9.4	N/A
	Would recommend alternate wording	2	5
Question 4 says: writing	Understood question as intended	39	100
	Relevance on scale from 0 to 10	8.7	N/A
	Would recommend alternate wording	2	5
Question 5 says: eating tasks	Understood question as intended	39	100
	Relevance on scale from 0 to 10	8.7	N/A
	Would recommend alternate wording	5	13
Question 6 says: dressing	Understood question as intended	39	100
	Relevance on scale from 0 to 10	8.5	N/A
	Would recommend alternate wording	3	8
Question 7 says: hygiene	Understood question as intended	38	97
	Relevance on scale from 0 to 10	9.1	N/A
	Would recommend alternate wording	0	0
Question 8 says: turning or changing position in bed	Understood question as intended	39	100
	Relevance on scale from 0 to 10	7.6	N/A
	Would recommend alternate wording	4	10
Question 9 says: sitting	Understood question as intended	39	100
	Relevance on scale from 0 to 10	8.9	N/A
	Would recommend alternate wording	9	23
Question 10 says: falling	Understood question as intended	39	100
	Relevance on scale from 0 to 10	10.0	N/A
	Would recommend alternate wording	7	18
Question 11 says: walking	Understood question as intended	38	97
	Relevance on scale from 0 to 10	9.7	N/A
	Would recommend alternate wording	5	13
Question 12 says: discomfort or pain related to PKAN	Understood question as intended	39	100
	Relevance on scale from 0 to 10	9.2	N/A
	Would recommend alternate wording	8	21
Overall impression	Good/great/like it	11	28
	Good response options	1	3
	Normal/fine	2	5
	Very specific	3	8
	Helpful	1	3
	Covers everything/thorough/in depth/most comprehensive	8	21
	Too general/need to be more detailed	2	5
	Easy to answer	6	15
	Relevant to experience/hits home	6	15
	Straightforward with appropriate responses	1	3

[1] N = 39.
N/A, not applicable.

Several participants suggested that the response options were not capturing a fine enough level of detail regarding changes in functioning, given that even smaller changes between adjacent severity levels of an item would be clinically meaningful. For example, one respondent suggested that the difference between the ability to use a fork versus a spoon should be captured, given that using a fork is more difficult. However, these were infrequent suggestions, and the team thought that the scale should represent unambiguous improvement and be broadly applicable for the purpose of a clinical trial and subsequent review by regulatory authorities.

Descriptive Statistics of the PKAN-ADL: Total and Item Scores

The PKAN–ADL total score at baseline had a mean (SD) of 25.8 (12.7) with a follow-up mean of 25.8 (12.6). Floor effects were observed at baseline and follow-up (data at follow-up were similar and not shown) for five items: eating tasks (30.8%), dressing (43.6%), hygiene (43.6%), falling (38.5%), and problems walking independently (38.5%). Ceiling effects were observed at baseline and follow-up (similar data not shown) for three of the items:

TABLE 3 *PKAN-ADL interitem correlations (interview 1)**

PKAN-ADL Measure	Total	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10	Item 11
ADL Total Score	—											
Item 1. Speech	0.66**	—										
Item 2. Salivation and drooling	0.48*	0.65**	—									
Item 3. Chewing, swallowing, and choking	0.83**	0.67**	0.51**	—								
Item 4. Writing	0.72**	0.49*	0.30	0.56**	—							
Item 5. Eating tasks	0.85**	0.55**	0.28	0.74**	0.56**	—						
Item 6. Dressing	0.90**	0.48*	0.29	0.67**	0.61**	0.80**	—					
Item 7. Hygiene	0.91**	0.48*	0.30*	0.67**	0.60**	0.82**	0.93**	—				
Item 8. Turning or changing position in bed	0.83**	0.48*	0.28	0.61**	0.52**	0.64**	0.80**	0.83**	—			
Item 9. Sitting	0.91**	0.54**	0.36*	0.72**	0.60**	0.81**	0.83**	0.89**	0.82**	—		
Item 10. Falling	0.84**	0.56*	0.32*	0.75**	0.71**	0.75**	0.75**	0.78**	0.71**	0.73**	—	
Item 11. Problems walking independently	0.91**	0.51**	0.34*	0.81**	0.71**	0.78**	0.80**	0.84**	0.77**	0.82**	0.91**	—
Item 12. Discomfort or pain related to PKAN	0.38*	0.23	0.22	0.29	0.08	0.30	0.20	0.24	0.23	0.33*	0.17	0.22

[1] N = 39.
*P < 0.05;
**P < 0.001.
Spearman's correlation coefficients.

Clinimetric Evaluation Results

Reliability

In general, items were strongly intercorrelated (*r* > 0.50; Table 3), with lower correlations for Salivation and Drooling salivation and drooling (50%), turning or changing position in bed (50%), and discomfort or pain related to PKAN (30.8%).

and Pain and Discomfort items. Cronbach's alpha for PKAN–ADL scores were 0.93 and 0.94 for the first and second visits, respectively.

Test–retest reliability (ICCs) for the individual PKAN-ADL items ranged from 0.81 (discomfort or pain related to PKAN) to 0.97 (chewing, swallowing, and choking). Test-retest reliability (ICC) for the PKAN–ADL total score was 0.99. The disease stability rating between visits was "the same" for the majority of patients (90%) and "worse" for 10%.

TABLE 4 *Construct and item discriminant validity: PKAN-ADL item correlations (interview 1)*[1]

PKAN-ADL Measure	Total	Neuro-QoL Upper	Neuro-QoL Lower	HUI Speech	HUI Dexterity	HUI Ambulation	HUI Pain	SADQ-10
ADL Total Score	**1.00**	**−0.94****	**−0.94****	**−0.75****	**−0.77****	**−0.87****	**−0.48***	**0.11**
Item 1. Speech	0.66**	−0.52**	−0.51*	−0.83**	−0.65**	−0.54**	−0.33	−0.04
Item 2. Salivation and drooling	0.48*	−0.37*	−0.32*	−0.60**	−0.39*	−0.36*	−0.15	−0.08
Item 3. Chewing, swallowing, and choking	0.83**	−0.74**	−0.74**	−0.63**	−0.70**	−0.74**	−0.39*	−0.00
Item 4. Writing	0.72**	−0.67**	−0.71**	−0.59**	−0.49*	−0.73**	−0.23	0.01
Item 5. Eating tasks	0.85**	−0.86**	−0.81**	−0.74**	−0.70**	−0.73**	−0.55**	0.22
Item 6. Dressing	0.90**	−0.91**	−0.86**	−0.63**	−0.74**	−0.75**	−0.37*	0.09
Item 7. Hygiene	0.91**	−0.92**	−0.91**	−0.68**	−0.68**	−0.79**	−0.36*	0.10
Item 8. Turning or changing position in bed	0.83**	−0.82**	−0.84**	−0.57**	−0.60**	−0.72**	−0.37*	−0.02
Item 9. Sitting	0.91**	−0.84**	−0.88**	−0.69**	−0.66**	−0.76**	−0.47*	0.13
Item 10. Falling	0.84**	−0.79**	−0.83**	−0.67**	−0.57**	−0.88**	−0.30	0.08
Item 11. Problems walking independently	0.91**	−0.87**	−0.92**	−0.61**	−0.68**	−0.96**	−0.34*	−0.02
Item 12. Discomfort or pain related to PKAN	0.38*	−0.26	−0.31	−0.19	−0.34*	−0.20	−0.82**	0.59**

[1] N = 39.
*P < 0.05;
**P < 0.001.
Pearson's correlation coefficients. The first row is bolded to set apart the Total Score of the PKAN-ADL correlations with other measures, vs the individual items from the PKAN-ADL that follow.

Convergent Validity

Significant correlations were observed between the PKAN-ADL total score and all domains of the Neuro-QoL and HUI-2/3 ($P < 0.001$; Table 4). Correlations between the Neuro-QoL upper extremities and lower extremities scores and the ADL total score were both $r = -0.94$ ($P < 0.0001$). Moderate-to-strong correlations were observed for HUI-2/3 attributes for speech, dexterity, and ambulation ($r = -0.75, -0.77, -0.87; P < 0.001$), and pain ($r = -0.48; P < 0.05$), indicating overall good convergent validity.[18] A similar pattern of convergent and divergent correlations was observed at visit 2.

Divergent Validity

A Spearman correlation coefficient of 0.11 ($P > 0.05$) between the PKAN-ADL total score and SADQ-10 indicated divergent validity, with a similar finding for visit 2.

Discriminant Validity

Correlations of PKAN-ADL individual items with the total score were moderate to high with r values ranging from 0.38 for Item 12 (discomfort or pain related to PKAN) to 0.91 for Items 7 (hygiene), 9 (sitting), and 11 (problems walking independently). Most PKAN-ADL items were equivalent with or more highly correlated with the PKAN-ADL total score than they were with other summary measures, with a few exceptions: the salivation and drooling item was more highly correlated with the HUI Speech variable than it was with the PKAN-ADL, and the pain item was more highly correlated with HUI Pain and the SADQ-10 than it was with the PKAN-ADL (Table 4). Similar findings were observed for visit 2 data.

Criterion Validity

Regression analyses to evaluate the relationship between duration of time since onset of clinical symptoms and severity of functional impairment as measured by the PKAN-ADL showed no significant relationships ($P > 0.05$). Regression analyses examining the association between severity of functional impairment and patient age at study enrollment ($r = -0.38; P < 0.05$), patient age at PKAN symptom onset ($r = -0.67; P < 0.0001$), and patient age at PKAN diagnosis ($r = -0.47; P < 0.05$) all indicated greater severity of functional impairment with younger age.

Item Functioning

The PKAN-ADL is a multidomain scale representing the overall concept of disease severity. As such, the item response gradient should reflect consistent and increasing individual item contributions to the overall score, which appears visually (Fig. 1) as gradually increasing shading in a diagonal pattern from upper left to lower right corners. High total scores (bottom rows of the figure) are associated with almost perfectly consistent use of the highest responses across individual items, whereas the

reverse is true for low overall severity scores. Near ideal item functioning is observed in the horizontal and vertical item response gradients, with two notable exceptions. Items that showed the most consistent response gradient as indicated by gradually darkening shading by individual disease severity, as well as use of the full set of response categories, were sitting, chewing, and eating tasks. The pain and speech items showed less uniform patterns, suggesting that severity in these domains is less closely related to overall severity of impairment in ADLs. A similar pattern of item responses was observed at visit 2 (data not shown).

Discussion

The PKAN-ADL quantitatively assesses domains of disability related to motor functioning in PKAN patients, improvement in which would indicate clinically meaningful change in clinical trials of therapeutics. In stage 1 of development, a systematic item-generation process developed a conceptual framework for domains of interest with one item per domain and included expert, patient advocacy, and family input in the refinement and selection of items.[6] Qualitative interviews with family members and patients in stage 2 supported the relevance and comprehensiveness of the items.

All items were clearly understood by participants and highly relevant to PKAN. Furthermore, respondents understood the items as intended; 97% to 100% of participants understood the content and response scales for all PKAN-ADL items. Participants sometimes indicated the need for additional items to cover the relevant areas of functioning. However, those additional concepts were mostly symptoms and not relevant for measuring motor-based functional limitations. These results overall indicate strong content validity of the PKAN-ADL.

Quantitative evaluation of reliability and validity occurred through coadministration of the PKAN-ADL with measures logically related or unrelated and a second assessment at a time interval over which the great majority of patients reported stable disease. The PKAN-ADL is a reliable measure with high internal consistency and test-retest reliability.

PKAN-ADL scores were strongly correlated with both Neuro-QoL measures and the HUI-2/3 scores as expected.

Very good item functioning was found overall, with items measuring across the full range of disease severity and response category use logically aligned with item and individual severity levels.

Given that there are few cross-sectional studies of symptom profiles in PKAN, several patterns in the item response gradient are potentially useful to the field. Speech item scores tended to be higher where overall severity was relatively mild. This is consistent with the clinical observation that oromandibular dystonia is a common and early symptom in PKAN.[4,19,20] Salivation and drooling were significantly correlated with dysphagia and dysarthria (problems chewing, swallowing, choking, and speech), but less so with other items, supporting the concept that these activities are related to motor impairment of

oromandibular, pharyngeal, and laryngeal muscles. Pain showed low association with the individual item and total scores, suggesting that causes of pain and discomfort may be intermittent (e.g., episodic dystonic spasms). Another explanation could be that because of dysarthria, pain is only communicated effectively to caregivers when it is more extreme. Thus, the correlations between pain and other aspects of ADLs that are more readily observable are attenuated.

Some items showed floor and ceiling effects, and this variability is consistent with the phenotypic heterogeneity in PKAN. However, the relationship between total score and the item response gradient suggests that the scale can be sensitive to clinically meaningful change across a wide range of functional disability. Individual item correlations were moderate overall, and none appeared redundant in content or interitem correlation.

Surprisingly, no significant relationship was found between duration of illness and severity of functional impairment, in contrast to a recent finding in a similar-sized cohort.[4] This may reflect cohort or scale differences. A possible explanation is that older patients with longer PKAN duration who participated in our study may have been relatively healthier patients. There were significant associations between greater severity of functional impairment and younger age at PKAN symptom onset, at PKAN diagnosis, and at study enrollment. Together with the item response gradient, our results support the conceptualization that PKAN is most parsimoniously viewed as a phenotypic spectrum rather than a disease with distinct subtypes.

Strengths of the PKAN-ADL include its ease of use, brevity, comprehensiveness, clinimetric properties, and conceptual approach. The PKAN-ADL avoids the difficult and potentially unreliable requirement for multiple raters to distinguish between dystonic and parkinsonian features by focusing on the functional consequences of both categories of symptoms. Such an approach prioritizes patient and family concerns, as well as those of regulators, over objective characterization of disease features. Although this may be viewed as a limitation, because a subjective viewpoint (that of the patient) is presented rather than an objective measurement by the physician, it should be noted that expert ratings also require a degree of subjective judgment, introducing rater reliability (e.g., Monbaliu and colleagues)[21] as a methodological problem to be addressed in multisite clinical trials.

Potential limitations to this study include its relatively small sample size, which may limit generalizability. However, we achieved sufficient power for fundamental assessments of reliability and validity with a cohort size similar to that in the development of the PKAN-DRS.[4] Our cohort was primarily drawn from North America, and there may be geographical differences in mutation distribution, although the only identified cluster to our knowledge is in the Dominican Republic.[22] We recruited mostly caregivers, which reflects the clinical reality of PKAN, given that verbal interviews of patients with dysarthria are difficult or impossible. Finally, assessment of responsiveness/sensitivity to clinical change can be performed in future studies, including the ongoing randomized, placebo-controlled trial of fosmetpantotenate (clinicaltrials.gov identifier: NCT03041116) in which the PKAN-ADL is the primary efficacy endpoint.

Conclusion

The PKAN-ADL scale has supportive evidence of validity and reliability as a clinical outcomes assessment endpoint for clinical trials. The instrument demonstrated good content validity, excellent reliability, and good construct validity, and may be useful for clinical and research applications. The PKAN-ADL scale provides a patient-centered assessment of the functional limitations specifically associated with PKAN. The PKAN-ADL scale may serve as a useful clinical and research tool in the evaluation of patient functional response to currently available therapies and PKAN treatments in development.

Author Roles

(1) Research Project: A. Conception, B. Organization, C. Execution; (2) Statistical Analysis: A. Design, B. Execution, C. Review and Critique; (3) Manuscript Preparation: A. Writing the First Draft, B. Review and Critique.

R.D.M.: 1A, 1B, 1C, 2A, 2B, 2C, 3A, 3B
A.C.: 1A, 1B, 2C, 3A, 3B
M.L.E.: 1A, 1B, 2C, 3A, 3B
H.A.J.: 1A, 1B, 2C, 3A, 3B
T.K.: 1A, 1B, 2C, 3A, 3B
M.C.K.: 1A, 1B, 2C, 3A, 3B
A.V.: 1A, 1B, 2C, 3A, 3B
A.R.-V.: 1A, 1B, 2C, 3A, 3B
LS: 1A, 1B, 1C, 2A, 2B, 2C, 3A, 3B
D.A.R.: 1A, 1B, 1C, 2A, 2B, 2C, 3A, 3B
R.H.B.: 1A, 1B, 1C, 2A, 2B, 2C, 3A, 3B
W.R.L.: 1A, 1B, 1C, 2A, 2B, 2C, 3A, 3B

Acknowledgments

Disclosures

Ethical Compliance Statement: The study protocol received prior approval by Ethical and Independent Review Services, a central IRB. Interested potential participants were mailed a study packet that included two copies of the informed consent and assent forms. Evidera staff explained the study by telephone and obtained written and verbal informed consent for participants prior to the first data collection interview. Participants returned the signed informed

References

1. Orphanet. Pantothenate kinase-associated neurodegeneration. 2010. http://www.orpha.net/consor/cgi-bin/Disease_Search.php?lng=EN&data_id=17156&Disease_Disease_Search_diseaseGroup=PKAN&Disease_Disease_Search_diseaseType=Pat&Disease(s)/group%20of%20diseases=Pantothenate-kinase-associated-neurodegeneration&title=Pantothenate-kinase-associated-neurodegeneration&search=Disease_Search_Simple. Accessed 27 November 2017.

2. Marelli C, Piacentini S, Garavaglia B, Girotti F, Albanese A. Clinical and neuropsychological correlates in two brothers with pantothenate kinase-associated neurodegeneration. *Mov Disord* 2005;20:208–212.

3. Hogarth P, Kurian MA, Gregory A, et al. Consensus clinical management guideline for pantothenate kinase-associated neurodegeneration (PKAN). *Mol Genet Metab* 2017;120:278–287.

4. Darling A, Tello C, Marti MJ, et al. Clinical rating scale for pantothenate kinase-associated neurodegeneration: a pilot study. *Mov Disord* 2017;32:1620–1630.

5. Gregory A, Hayflick SJ. Pantothenate kinase-associated neurodegeneration. In: Adam MP, Ardinger HH, Pagon RA, et al., eds. *GeneReviews*® [Internet]. Seattle, WA: University of Washington, Seattle; 2017.

6. U.S. Food and Drug Administration, U.S. Department of Health and Human Services, 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. 2009. http://www.fda.gov/downloads/Drugs/Guidances/UCM193282.pdf. Accessed 18 September 2018.

7. U.S. Food and Drug Administration, U.S. Department of Health and Human Services, U.S. Center for Drug Evaluation and Research. Guidance for industry and FDA staff: qualification process for drug development tools. 2014. https://www.fda.gov/downloads/drugs/guidances/ucm230597.pdf. Accessed 27 November 2017.

8. Schneider SA, Bhatia KP, Hardy J. Complicated recessive dystonia parkinsonism syndromes. *Mov Disord* 2009;24:490–499.

9. Shulman LM, Armstrong M, Ellis T, et al. Disability rating scales in Parkinson's disease: critique and recommendations. *Mov Disord* 2016;31:1455–1465.

10. Roa P, Stoeter P, Perez-Then E, Santana M, Marshall RD. A pilot study of a potential phosphopantothenate replacement therapy in 2 patients with pantothenate kinase-associated neurodegeneration. *Int J Rare Dis Orphan Drugs* 2017;2:1006.

11. Christou YP, Tanteles GA, Kkolou E, et al. Open-label fosmetpantotenate, a phosphopantothenate replacement therapy in a single patient with atypical PKAN. *Case Rep Neurol Med* 2017;2017:3247034.

12. Fahn S, Elton R. *Recent Developments in Parkinson's Disease*, Vol. 2. Florham Park, NJ: Macmillan Health Care Information; 1987.

13. Polit DF. Getting serious about test-retest reliability: a critique of retest research and some recommendations. *Qual Life Res* 2014;23:1713–1720.

14. Furlong W, Feeny D, Torrance G. 2016 *Health Utilities Index*® *mark 2 and mark 3 (HUI2/3) 40-item questionnaire for interviewer-administered,*

proxy-assessed "usual" health status assessment. Dundas, Ontario, Canada: Health Utilities; 2016.

15. Sutcliffe LM, Lincoln NB. The assessment of depression in aphasic stroke patients: the development of the Stroke Aphasic Depression Questionnaire. *Clin Rehabil* 1998;12:506–513.

16. Perez L, Huang J, Jansky L, et al. Using focus groups to inform the Neuro-QOL measurement tool: exploring patient-centered, health-related quality of life concepts across neurological conditions. *J Neurosci Nurs* 2007;39:342–353.

17. Lincoln N, Sutcliffe L, Unsworth G. Validation of the Stroke Aphasic Depression Questionnaire (SADQ) for use with patients in hospital. *Clin Neuropsychol* 2000;1:88–96.

18. Riessman C. *Narrative Methods for the Human Sciences*. Thousand Oaks, CA: SAGE; 2008.

19. Hayflick SJ, Westaway SK, Levinson B, et al. Genetic, clinical, and radiographic delineation of Hallervorden-Spatz syndrome. *N Engl J Med* 2003;348:33–40.

20. Sachin S, Goyal V, Singh S, et al. Clinical spectrum of Hallervorden-Spatz syndrome in India. *J Clin Neurosci* 2009;16:253–258.

21. Monbaliu E, Ortibus E, Roelens F, et al. Rating scales for dystonia in cerebral palsy: reliability and validity. *Dev Med Child Neurol* 2010;52:570–575.

22. Delgado RF, Sanchez PR, Speckter H, et al. Missense PANK2 mutation without "eye of the tiger" sign: MR findings in a large group of patients with pantothenate kinase-associated neurodegeneration (PKAN). *J Magn Reson Imaging* 2012;35:788–794.

Supporting Information

Supporting information may be found in the online version of this article.

Appendix S1. PKAN–ADL scale