Toward the Definition of Patient-Reported Outcome Measurements in Hereditary Spastic Paraplegia

Matthias Amprosi, MD, Elisabetta Indelicato, MD, PhD, Andreas Eigentler, MD, PhD, Josef Fritz, PhD, Wolfgang Nachbauer, MD, PhD, and Sylvia Boesch, PD, MD, MSc

Correspondence
Dr. Boesch
sylvia.boesch@i-med.ac.at

Abstract

Background and Objectives
Hereditary spastic paraplegias (HSPs) are a heterogeneous group of rare neurodegenerative diseases, characterized by a progressive spastic paraparesis. Currently, there is a HSP-specific clinician-reported outcome measure (CROM) called Spastic Paraplegia Rating Scale (SPRS). There are, however, no specific patient-reported outcome measures (PROMs) for HSP. In the present cohort study, we prospectively follow up a well-examined Austrian HSP cohort using validated rating scales and compared PROM with disease-specific and non–disease-specific CROM.

Methods
Patients were recruited and followed up at the Center for Rare Movement Disorders, Innsbruck, Austria. CROM included the SPRS, Scale for the Assessment and Rating of Ataxia (SARA), Barthel Index (BI), and Mini-Mental State Examination (MMSE). PROM included the EQ-5D questionnaire and the Patient Health Questionnaire 9 (PHQ-9). Standardized response means (SRMs) were calculated for all scales at follow-up (FU) after 1 year.

Results
A total of 55 patients (36 males) with HSP were included in the study. FU was performed for 30 patients (21 males). Apart from females reporting more problems in the EQ-5D domain of anxiety and depression ($p = 0.008$), other clinician-reported outcomes (CROs) or patient-reported outcomes (PROs) did not differ significantly across sex. SPRS showed significant correlations with SARA ($p < 0.001$), mainly driven by the gait item, as well as the BI. Although SPRS did not correlate with EQ-5D visual analogue scale and PHQ-9 scores, several EQ-5D domains correlated significantly with SPRS. At FU, SPRS showed the highest responsiveness (SRM 1.11), followed by SARA (SRM 0.47). Neither MMSE nor PRO significantly increased at FU.

Discussion
In this study, we present an Austrian cohort of patients with HSP and a prospective study evaluating correlations of CRO and PRO as well as their progression. Demographics from our cohort are comparable with several other European cohort studies. Our data highlight the capabilities of the SPRS to show clinical progression and warrant consideration of ataxia rating scales such as SARA in HSP cohorts. We also show that the generic PROMs are not suitable to detect change in HSP, and thus, we propose to create a disease-specific PROM fully depicting the effect of HSP on the patients' lives.

Glossary

Hereditary spastic paraplegias (HSPs) are a heterogeneous group of rare neurodegenerative diseases. The prevalence is estimated to be between 1 and 10 per 100,000.[1] HSPs are characterized by a progressive spastic paraparesis, and based on the widely used definition by Harding[2] from 1983, they can be divided into pure HSP (pHSP) and complicated HSP (cHSP) forms. Although patients with pHSP solely exhibit spasticity and weakness of the lower extremities, possibly accompanied by impaired vibration sense and bladder disturbances, patients with cHSP can display a variety of additional signs and symptoms such as epilepsy, cognitive impairment, or ataxia.[2] Because of the additional symptoms in cHSP, some forms have a significant overlap with hereditary ataxias, and some genes, such as *SPG7*, are known to cause a spectrum of phenotypes ranging from predominantly ataxia to predominantly spasticity. It thus has been suggested to replace this classification and rather move toward a mechanistical classification system of a continuous ataxia-spasticity disease spectrum.[3]

The establishment of multicentric registry studies in the field led to the development of several clinician-reported outcome measures (CROMs) (the respective outcomes designated as clinician-reported outcomes [CROs]) such as the Spastic Paraplegia Rating Scale (SPRS) for HSP and the Scale for the Assessment and Rating of Ataxia (SARA) for ataxias.[4,5] Lately, the focus of clinical research shifted toward the development of patient-reported outcome measures (PROMs) (the respective outcomes designated as patient-reported outcomes [PROs]) as complementary instruments in the evaluation of patients with chronic progressive neurodegenerative disorders.[6-8] PROM can capture additional aspects of the multifaceted impact of neurologic disabilities and may even show a better sensitivity to change than CROM. For example, natural history data from the European Friedreich Ataxia Registry showed that the Activities of Daily Living part of the Friedreich Ataxia Rating Scale (FARS-ADLs), a PROM subscale, depict disease progression better than SARA.[9] To be prepared for possible upcoming clinical trials in HSP, established endpoints such as fluid biomarkers are required for phase 2 trials, whereas validated CROM and PROM are essential for later phases (e.g., phases 3 and 4) of clinical trials. Up to date, there are no validated HSP-specific PROMs. Currently, several generic quality of life measurements are available, which have been widely applied in natural history studies in spastic-ataxic disorders including HSP. However, their correlation with CROM and consequently their suitability for studies in HSP have been sparsely investigated.[10,11]

In this study, we aimed (1) to prospectively follow up (FU) a well-examined Austrian HSP cohort using validated rating scales and (2) to compare PROMs with disease-specific and non–disease-specific CROMs.

Methods

Patients

Patients were recruited at the Center for rare movement disorders, Innsbruck, Austria. An inclusion criterion was a clinical diagnosis of HSP[4] regardless of genetic assignment. Exclusion criteria were significant comorbidities that affected evaluations of both CROMs and PROMs. A total of 56 patients with a previous diagnosis of HSP were screened between April 2019 and December 2021. One patient was excluded because of relevant comorbidities, and 55 were included in the study. Follow-up with the same procedures as at baseline was performed after 1 year (FU) for 30 of these patients.

CROMs

SPRS

Severity of the disease was assessed with the SPRS, which encompasses a maximum total score of 52.[4]

SARA

To depict the full spectrum of signs on the ataxia-spasticity spectrum in HSP, we also performed the SARA to assess the severity of possible ataxia. The scale has a maximum score of 40 points.[5]

Mini-Mental State Examination

The Mini-Mental State Examination (MMSE) includes 11 items testing cognitive functions. A maximum score of 30 can be reached when all questions are answered correctly.[12] A recent review has evaluated several cutoff points for their sensitivity and specificity to distinguish between impaired and normal cognitive functions. A cutoff of 24 points has been shown to have a pooled diagnostic specificity of 0.90 and a sensitivity of 0.85.[13] We thus chose this cutoff for our study.

Four-Stage Scale of Motor Disability

The 4-Stage Scale of Motor Disability (4SMD) was first introduced in 2009 and has subsequently been used in other HSP

studies.[11,14,15] It differentiates between (1) mild symptoms and signs at examination, ambulatory without an aid; (2) walking without an aid, unable to run; (3) walking with an aid; and (4) wheelchair bound.

Barthel Index

The Barthel Index (BI) has 8 items that specifically have been chosen to depict a patient's ability to care for themselves and therefore the ADLs. There is a maximum total score of 100 points, corresponding to complete independence in ADL, whereas 0 means complete dependence on help.[16] It was designed to be filled out by the caregiver or health care professional, but because it does represent the patients' ability in daily living and not a clinical examination, it may be considered an overlap between PROM and CROM.

PROMs

EQ-5D

The EQ-5D questionnaire is a non–disease-specific tool to measure quality of life (QoL). It consists of 5 dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) that are divided into 3 levels: 1 (no problems), 2 (some problems), and 3 (extreme problems). Furthermore, it contains a visual analogue scale (EQ-VAS) that records the self-rated health status on a scale between 0 (worst health state) and 100 (best health state).[17]

Patient Health Questionnaire 9

The Patient Health Questionnaire 9 (PHQ-9) is the depression module of the PHQ. It consists of 9 items representing the 9 *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* criteria for depressive disorders. Each item can be scored between 0 (not at all) and 3 (nearly every day). In this study, it was evaluated with a sum score and the categorical rating of depression severity with the cutoffs ≥5 points (mild depression), ≥10 points (moderate depression), ≥15 points (moderately severe depression), and ≥20 points (severe depression) according to Kroenke et al.[18]

Statistical Analysis

Statistical analysis was performed with SPSS version 26.0 (IBM 2019, Armonk, NY), and the threshold for statistical significance was set at $p < 0.05$ for all analyses. Data are reported as mean and SD or frequencies and percentages as appropriate. Normal distribution was tested with the Shapiro-Wilk test, and as all tested variables were not normally distributed, the Mann-Whitney U test, Kruskal-Wallis test, χ^2 test, and Spearman's rank correlation coefficient were used. To evaluate effects of sex, disease duration, age, BI, 4SMD, SPRS, and SARA on depression and QoL, a linear model with PHQ-9 and EQ-VAS with stepwise selection was performed. Another linear model with the SPRS sum score as a dependent variable and all SARA subitems as independent variables with stepwise inclusion was calculated. Furthermore, we calculated standardized response means (SRMs) to evaluate responsiveness of the different outcome measures. SRMs were calculated by dividing the mean

change in scores from baseline to FU by the SD of the change. The mean delta of each score at FU and baseline and the standard error were used to estimate annual progression rates.

Standard Protocol Approvals, Registrations, and Patient Consents

All investigations were performed in accordance with the Declaration of Helsinki. This study was approved by the institutional review board (Ethikkommission der Medizinischen Universität Innsbruck, Vote: 1255/2018), and all patients gave written informed consent before inclusion in the study.

Data Availability

Detailed genetic and clinical data are available on request.

Results

Demographics

For detailed information on the demographics, see Table 1. In total, 55 patients, 36 of whom were male (65.5%), with a clinical diagnosis of HSP were included. Thirty-one patients (56.4%) exhibited cHSP. The mean age at onset of the whole cohort was 30.3 years (SD ± 17.7). The average disease duration was 17.8 years (SD ± 12.1). Most cases (47.3%) were simplex, followed by autosomal dominant (29.1%) and autosomal recessive (23.6%). The diagnosis was genetically confirmed in 26 patients (47.3%), and *SPG4* was the most frequent genotype with 57.7% of genetically assigned cases and 27.3% of the total cohort (Figure 1). In genetically assigned cases, diagnosis was established after a mean period of 12.9 years (SD ± 10.0). Although 30.9% of the participants were prescribed neither antispastic drugs nor antidepressants, 49.1% took medication against spasticity, 5.5% took antidepressants, and 14.5% used both types of medications. Nineteen (34.5%) of the patients had spastic ataxia. There were no significant differences between genders concerning disease duration, age at onset, age at diagnosis, age at examination, disease form, or genetic assignment as analyzed by means of the Mann-Whitney U test. When comparing patients with pHSP and cHSP, patients with cHSP had a significantly lower age at onset ($p = 0.009$, 25.1 ± 18.2 vs 37.1 ± 14.7) and age at examination ($p = 0.006$, 43.6 ± 15.1 vs 54.1 ± 10.0).

Thirty of the 55 patients were followed up after a period of 1 year (mean in days: 392.7, SD ± 123.9). As this is an ongoing registry study at our center and patients are included continuously, 11 patients were included within the year before the analysis and thus did not return for a FU yet. In total, 14 patients were lost to FU (n = 7 for unspecified reasons and n = 7 postponed their visit because of the COVID-19 pandemic). There were no significant differences in demographics, as well as CRO or PRO results at baseline between the patients with and without FU.

PROMs and CROMs

PHQ-9, EQ-5D, SARA, and MMSE were available in all but one case (98.2%). Detailed information on all PROs and CROs can

Table 1 Demographics and Clinical Characteristics

	Total	pHSP	cHSP
n (%)	55 (100)	24 (43.6)	31 (56.4)
Female/male; n (%)	19 (34.6)/36 (65.5)	8 (33.3)/16 (66.7)	11 (35.5)/20 (64.5)
Age; mean (SD)	48.2 (±14.1)	54.1 (±10.0)	43.6 (±15.1)
Age at onset; mean (SD)	30.3 (±17.7)	37.1 (±14.7)	25.1 (±18.2)
Disease duration; mean (SD)	17.8 (±12.1)	17.0 (±12.5)	18.5 (±12.0)
Age at diagnosis; mean (SD)	45.0 (±13.1)	48.0 (±10.7)	39.2 (±15.8)
Inheritance, n (%)			
Autosomal dominant	16 (29.1)	13 (54.2)	3 (9.7)
Autosomal recessive	13 (23.6)	2 (8.3)	11 (35.5)
Simplex	27 (47.3)	9 (37.5)	17 (54.8)
Clinical characteristics; mean (SD)			
SPRS score	17.4 (±9.4)	15.0 (±7.3)	19.2 (±10.5)
SARA score	8.4 (±7.6)	4.1 (±2.2)	11.9 (±8.6)
MMSE score	28.7 (±3.9)	29.6 (±0.7)	27.9 (±5.0)
Barthel Index	92.9 (±15.7)	95.6 (±9.3)	90.8 (±19.2)
4SMD, n (%)			
Mild symptoms and signs, ambulatory without an aid	6 (10.9)	3 (12.5)	3 (9.7)
Walking without an aid, unable to run	24 (43.6)	12 (50.0)	12 (38.7)
Walking with an aid	20 (36.4)	8 (33.3)	12 (38.7)
Wheelchair bound	5 (9.1)	1 (4.2)	4 (12.9)
Medication, n (%)			
No medication	17 (30.9)	6 (25.0)	11 (35.5)
Antispastic drugs	27 (49.1)	13 (54.2)	14 (45.2)
Antidepressants	3 (5.5)	2 (8.3)	1 (3.2)
Both types	8 (14.5)	3 (12.5)	5 (16.1)

Abbreviations: 4SMD = 4-Stage Scale of Motor Disability; cHSP = complicated hereditary spastic paraplegia; HSP = hereditary spastic paraplegia; MMSE = Mini-Mental State Examination; pHSP = pure hereditary spastic paraplegia; SARA = Scale for the Assessment and Rating of Ataxia; SPRS = Spastic Paraplegia Rating Scale.
This table shows the demographics, clinical characteristics (including rating scales), and medication of all included patients and patients with pHSP and cHSP. Numbers are given with percentages in brackets and mean values with the SD in brackets.

be seen in Table 2. Scatter plots of correlation analyses can be found in eFigure 1, links.lww.com/NXG/A571. In the whole cohort, the mean SPRS score was 17.35 (SD ± 9.4) and the mean SARA score 8.40 (SD ± 7.6), mostly corresponding to advanced spasticity affecting the length of independent walking distance and the use of stairs in the respective items of the SPRS. Both SPRS (Spearman's rho 0.514, $p < 0.001$) and SARA (Spearman's rho 0.352, $p = 0.009$) scores significantly increased with disease duration. In addition, the SPRS inversely correlated with age at onset (Spearman's rho −0.328, $p = 0.015$). Furthermore, both scales showed a moderate correlation with each other (Spearman's rho 0.637, $p < 0.001$) (Figure 2A). A linear model with the SPRS score as a dependent and all SARA items

as independent variables revealed a significant effect of the gait, finger chase, and heel-shin slide items (adjusted R square 0.852; $p < 0.001$) (see eAppendix 1, links.lww.com/NXG/A571, for detailed information on the regression analyses). Furthermore, the gait item was the most dominant of those 3 items (R square of 0.796, $p < 0.001$). This can also be seen in the scatter plots in Figure 2. The SARA score was significantly higher in patients with cHSP than in those with pHSP ($p < 0.001$, 11.9 ± 8.6 vs 4.1 ± 2.2). However, the mean SPRS scores of patients with cHSP were not significantly higher than those of patients with pHSP.

The mean BI was 92.9 (SD ± 15.7). The most frequently affected domain was bladder control (21.8%), followed by

Figure 1 Genetic Assignment of the Cohort

The gene designations are given with the number of individuals in brackets.
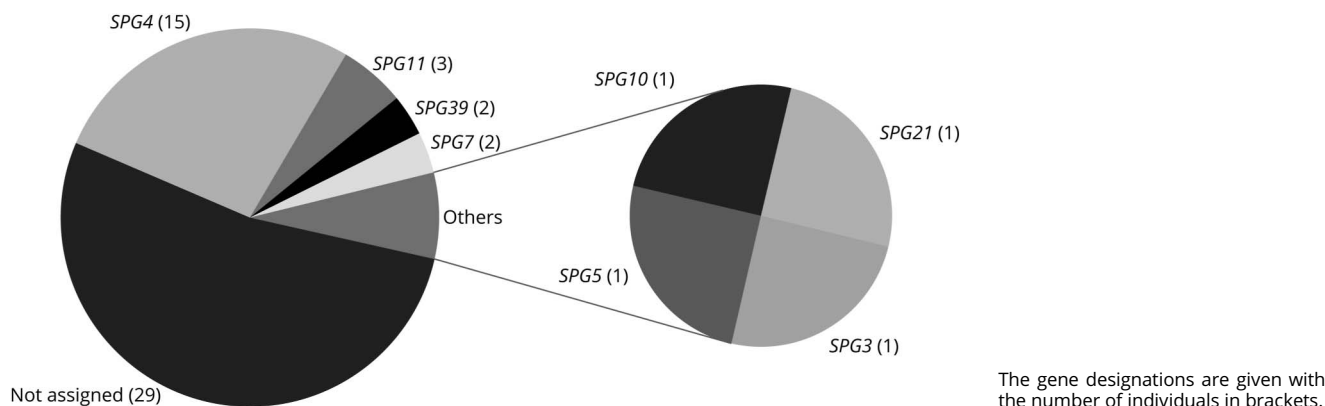
mobility on level surfaces (18.2%), bathing (10.9%), bowel control (10.9%), stairs, grooming (9.1% each), dressing, transfers and eating (7.3% each), and toilet use (5.5%). Impairment in ADL in this HSP cohort, as measured by the BI, correlated inversely with disease duration (Spearman's rho -0.401, $p = 0.002$) and positively with later age at onset (Spearman's rho 0.339, $p = 0.011$). The BI exhibited a significant inverse correlation with the SPRS (Spearman's rho -0.611, $p < 0.001$) and SARA (Spearman's rho -0.509, $p < 0.001$). The BI items mobility on level surfaces (Spearman's rho -0.617, $p < 0.001$), bathing (Spearman's rho -0.495, $p < 0.001$), stairs (Spearman's rho -0.459, $p < 0.001$), grooming

(Spearman's rho -0.433, $p < 0.001$), transfers (Spearman's rho -0.374, $p = 0.005$), bowel control (Spearman's rho -0.351, $p = 0.009$), dressing (Spearman's rho -0.337, $p = 0.012$), and bladder control (Spearman's rho -0.322, $p = 0.017$) correlated inversely with SPRS scores.

Only 4 patients (7.4%) showed impairments according to the MMSE, but there were no correlations with disease severity as measured by the SPRS and SARA.

Regarding the 4SMD, 10.9% of the patients had mild symptoms and were ambulatory without an aid, 43.6% were

Table 2 Results of Clinician- and Patient-Reported Outcome Measures

	Total (n = 55)	Male (n = 36)	Female (n = 19)	*p* Value	pHSP (n = 24)	cHSP (n = 31)	*p* Value
SPRS score (mean, SD)	17.4 (9.4)	18.3 (9.6)	15.5 (9.1)	0.205	15.0 (7.3)	19.2 (10.5)	0.221
SARA score (mean, SD)	8.44 (7.6)	8.7 (7.0)	8.0 (8.7)	0.412	4.1 (2.2)	11.9 (8.6)	**<0.001**
4SMD (median, IQR)	2 (1)	3 (1)	2 (1)	0.064	2 (1)	3 (1)	0.629
Barthel Index (mean, SD)	92.9 (15.7)	93.1 (12.0)	92.6 (21.6)	0.648	95.6 (9.2)	90.8 (19.2)	0.366
EQ-VAS (mean, SD)	69.7 (16.9)	68.1 (17.3)	73.1 (15.9)	0.262	66.1 (13.8)	72.7 (18.7)	0.178
MMSE score (mean, SD)	28.7 (3.9)	28.9 (2.5)	28.2 (5.6)	0.360	29.6 (0.7)	27.9 (5.0)	0.525
PHQ sum score (mean, SD)	3.5 (4.4)	2.7 (3.0)	5.1 (6.1)	0.187	4.0 (5.2)	3.2 (3.7)	0.588
PHQ categories, n (%)							
None	35 (64.8)	24 (66.7)	11 (61.11)		14 (58.3)	21 (70.0)	
Mild depression	16 (29.6)	11 (30.6)	5 (27.78)		9 (37.5)	7 (23.3)	
Moderate	2 (3.7)	1 (2.8)	1 (5.56)		0 (0.0)	2 (6.7)	
Moderately severe depression	0 (0.0)	0 (0.0)	0 (0.0)		0 (0.0)	0 (0.0)	
Severe	1 (1.9)	0 (0.0)	1 (5.6)		1 (4.2)	0 (0.0)	

Abbreviations: cHSP = complicated hereditary spastic paraplegia; CRO = clinician-reported outcome; EQ-VAS = EuroQoL visual analogue scale; HSP = hereditary spastic paraplegia; MMSE = Mini-Mental State Examination; PHQ = Patient Health Questionnaire; pHSP = pure hereditary spastic paraplegia; PRO = patient-reported outcome; SARA = Scale for the Assessment and Rating of Ataxia; SPRS = Spastic Paraplegia Rating Scale.
This table shows detailed information on the different CROs and PROs for all patients, patients with pHSP and cHSP, as well as males and females. Scores are given as mean with SD in brackets and categorical variables as numbers with percentages in brackets (significant values in bold).

Figure 2 Scatter Plots and Regression Lines for the Correlation of the SPRS and SARA (A) and the SPRS and SARA Gait Item (B)
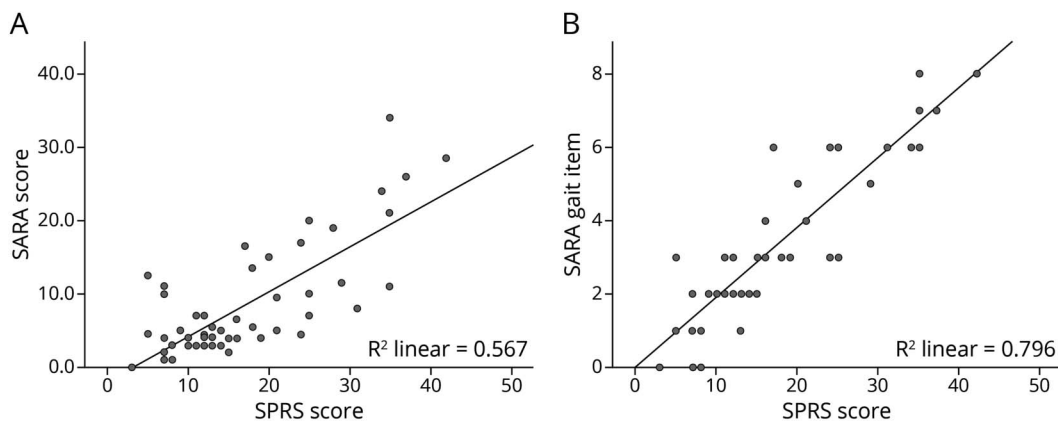
(A) The scatter plot and the regression line for the correlation of the SPRS and SARA scores. (B) The effect of the gait item alone on the correlation with SPRS scores. SARA = Scale for the Assessment and Rating of Ataxia; SPRS = Spastic Paraplegia Rating Scale.

ambulatory without an aid but not able to run anymore, 36.4% used a walking aid, and 9.1% were dependent on a wheelchair. The 4SMD score significantly correlated with disease duration (Spearman's rho 0.455, $p < 0.001$), with disability in ADL (Spearman's rho -0.584, $p < 0.001$), and with SARA (Spearman's rho 0.590, $p < 0.001$) and SPRS (Spearman's rho 0.830, $p < 0.001$) scores. There was, however, no correlation between 4SMD and cognition, QoL, or depression as measured by the MMSE, EQ-VAS, and PHQ-9, respectively.

The analysis of PHQ-9 scores revealed that 29.6% of the patients had a mild, 3.7% a moderate, and 1.9% a severe depression. 64.8% had values below the cutoff of 5 points. Only 7.4% of patients did not report problems in any of the EQ-5D dimensions, whereas most patients (31.5%) reported problems in 4 dimensions. Mobility was affected most frequently (87.0%), followed by problems in usual activities (64.8%), pain (51.9%), anxiety and depression (49.1%), and self-care (33.4%). The mean EQ-VAS was 69.7 (SD ± 16.9). The PHQ-9 score had a significant and inverse correlation with EQ-VAS (Spearman's rho -0.349, $p = 0.010$). Neither QoL nor depression, as measured by the EQ-VAS and PHQ-9, showed significant correlations with age at onset or disease duration.

When comparing genders, female patients reported significantly more problems in the EQ-5D domain of anxiety and depression compared with men ($p = 0.008$, 1.4 vs 1.8). No other significant differences between genders or between pHSP and cHSP were observed.

Comparing correlations between CRO and PRO, QoL correlated with independence in ADL (Spearman's rho 0.292, $p = 0.032$). Although correlations between overall perceived QoL and disease severity were not significant, they were significant between SPRS as well as SARA scores and several EQ-5D domains. The SPRS correlated with the number of affected EQ-5D domains (Spearman's rho 0.460, $p < 0.001$) and problems in the domains self-care (Spearman's rho 0.552, $p < 0.001$), usual activities (Spearman's rho 0.399, $p = 0.003$), and mobility (Spearman's rho 0.342, $p = 0.011$). SARA scores also had negative correlations with the domains self-care (Spearman's rho 0.460, $p = 0.001$) and usual activities (Spearman's rho 0.274, $p = 0.047$). Depression scores did not correlate with disease severity, cognitive functions, or independence in ADL.

The linear model with the EQ-VAS as a dependent variable confirmed the correlation between QoL and ADL ($p = 0.044$, adjusted R square = 0.059). The linear model with depression as a dependent variable revealed age, sex, and EQ-VAS ($p = 0.001$, adjusted R square = 0.281) as significant independent variables. None of the other demographic factors or disease severity was significantly influencing depression or QoL (see eAppendix1, links.lww.com/NXG/A571, for detailed information on the regression analyses). No significant differences concerning QoL or depression were observed across the 4SMD stages.

FU

Mean scores of the included PRO and CRO with estimated progression rates and SRM at baseline and FU are given in Table 3. In the total cohort with a FU available, the SPRS score significantly increased after 1 year ($p < 0.001$, 15.8 ± 8.2 vs 17.0 ± 8.5), exhibiting a mean progression of 1.17 (SD ± 1.05) points (see eFigure 2, links.lww.com/NXG/A571, for the box plots at baseline and FU for all scales). Progression was significant for both forms pHSP and cHSP. When compared directly, however, the rate of progression between cHSP and pHSP did not differ significantly ($p = 0.93$, 1.2 ± 1.1 vs 1.2 ± 1.0). The SARA score in the whole cohort also increased significantly ($p = 0.018$, 8.1 ± 6.7 vs 8.6 ± 7.1), with a mean progression of 0.57 (SD ± 1.2) points. Progression in

Table 3 Comparison of Scores at Baseline and Follow-up

	Patients (n)	Baseline Mean (SD)	1-y follow-up Mean (SD)	SRM	Annual progression rate (SE)	95% CI	p Value
SPRS score							
Total population	30	15.8 (8.2)	17.0 (8.5)	1.11	1.2 (0.2)	0.8 to 1.6	**<0.001**
pHSP	13	13.7 (5.3)	14.9 (5.7)	1.01	1.2 (0.3)	0.5 to 1.9	**0.011**
cHSP	17	17.5 (9.8)	18.7 (10.0)	1.16	1.2 (0.3)	0.7 to 1.7	**0.002**
SARA score							
Total population	30	8.1 (6.7)	8.6 (7.1)	0.47	0.6 (0.2)	0.1 to 1.0	**0.018**
pHSP	13	3.5 (1.0)	3.9 (1.5)	0.30	0.3 (0.3)	−0.3 to 0.9	0.317
cHSP	17	11.5 (7.2)	12.3 (7.6)	0.59	0.8 (0.3)	0.1 to 1.4	**0.028**
Barthel Index							
Total population	30	95.2 (9.7)	93.8 (13.3)	−0.26	−1.3 (0.9)	−3.2 to 0.6	0.131
pHSP	13	98.5 (3.2)	98.1 (3.8)	−0.12	−0.4 (0.9)	−2.3 to 1.6	0.655
cHSP	17	90.7 (12.4)	90.6 (13.9)	−0.34	−2.1 (1.5)	−5.2 to 1.1	0.102
EQ-VAS							
Total population	30	70.2 (17.5)	68.5 (19.9)	−0.11	−1.7 (2.8)	−7.5 to 4.1	0.298
pHSP	13	68.5 (12.6)	71.2 (11.9)	0.18	2.6 (4.1)	−6.4 to 11.6	0.765
cHSP	17	71.5 (20.8)	66.5 (24.5)	−0.32	−5.0 (3.7)	−12.9 to 2.9	0.154
PHQ sum score							
Total population	30	3.4 (3.2)	4.0 (4.4)	0.18	0.6 (0.6)	−0.7 to 2.0	0.305
pHSP	13	3.5 (2.8)	3.2 (4.0)	−0.07	−0.3 (1.2)	−2.9 to 2.3	0.838
cHSP	17	3.4 (3.5)	4.7 (4.7)	0.50	1.4 (0.8)	−0.0 to 2.7	0.062
MMSE score							
Total population	29	28.9 (2.7)	28.8 (2.8)	−0.11	−0.1 (0.2)	−0.5 to 0.3	0.782
pHSP	13	29.7 (0.6)	29.7 (0.5)	0.00	0.0 (0.2)	−0.4 to 0.4	1.000
cHSP	16	28.3 (3.6)	28.1 (3.6)	−0.17	−0.2 (0.3)	−0.8 to 0.4	0.705

Abbreviations: 4SMD = 4-Stage Scale of Motor Disability; cHSP = complicated hereditary spastic paraplegia; CRO = clinician-reported outcome; EQ-VAS = EuroQoL visual analogue scale; MMSE = Mini-Mental State Examination; PHQ = Patient Health Questionnaire; pHSP = pure hereditary spastic paraplegia; PRO = patient-reported outcome; SARA = Scale for the Assessment and Rating of Ataxia; SPRS = Spastic Paraplegia Rating Scale.
This table shows the corresponding mean scores with the SD of all PROs and CROs at baseline and 1-year follow-up for all patients and for patients with pHSP and cHSP individually. The mean delta of each score at follow-up and baseline was used to estimate annual progression rates. The paired Wilcoxon test was used to analyze whether progression from baseline to follow-up was significant, and p values are given (significant p values in bold). In addition, standard error (SE) and the 95% confidence interval (CI) for the mean progression are given. Standardized response means (SRMs) were calculated by dividing the mean change in scores from baseline to follow-up (FU) by the SD of the change and used to evaluate the responsiveness of different scores.

patients with cHSP was even higher (0.76) but was not statistically significant for patients with pHSP. SPRS showed the highest responsiveness (SRM 1.11), followed by SARA (SRM 0.47). BI, MMSE, nor any of the PROs significantly increased in the total cohort or the subtypes over the period of 1 year.

Limitations

Our study is limited by the small number of patients and the number of patients lost to FU. Furthermore, because of the explorative nature of the current study, we did not adjust for multiple testing to decrease the risk of overlooking relevant associations. This increases the likelihood of false-positive

findings. A confirmation of our exploratory analysis is warranted in multicentric cohorts.

Discussion

In this study, we (1) prospectively followed up an Austrian HSP cohort using validated rating scales and (2) compared PROMs with disease-specific and non–disease-specific CROMs.

To date, several HSP cohorts from other European countries, including Norway, Portugal, Estonia, Germany, and France,

Table 4 Comparison of the Different Scales

Items/categories	SPRS	SARA	4SMD	BI	EQ-5D	PHQ-9	MMSE
Ambulation	X	X	X	X	X	—	—
Activities of daily living	—	—	—	X	X	—	—
Overall quality of life	—	—	—	—	X	—	—
Pain	X	—	—	—	X	—	—
Urinary/bowel symptoms	X	—	—	X	X	—	—
Cognitive functions	—	—	—	—	—	—	X
Psychiatric symptoms	—	—	—	—	X	X	—

Abbreviations: 4SMD = 4-Stage Scale of Motor Disability; EQ-VAS = EuroQoL visual analogue scale; MMSE = Mini-Mental State Examination; PHQ = Patient Health Questionnaire; SARA = Scale for the Assessment and Rating of Ataxia; SPRS = Spastic Paraplegia Rating Scale.
This table shows relevant disease domains and the used clinician- and patient-reported outcome measures. If a domain is included in the scale, it is marked with X.

have been reported.[1,14,19-21] When comparing these studies, a high variability in demographics and clinical characteristics is noticeable. The percentage of patients with pHSP in these cohorts ranged from 35% of the families in Portugal to 81.4% of the patients in Estonia.[1,14,19-21] In the 2 larger published cohorts from Germany and France, 42.1% and 43.9% of patients, respectively, were diagnosed with pHSP.[1,21] In the present cohort, the rate of pHSP was comparable with German and French rates with 43.6%.[1,21] The age at onset in our patients was similar to the one reported in the German cohort with a mean of 30 years,[1] whereas the median in the French sample was 25 years,[21] and the age at onset was not given for the total cohorts in the remaining studies.[14,20,22] Furthermore, in the above studies, male patients accounted for 50%–61% of patients.[1,14,20-22] In our cohort, however, the male predominance was even more marked with 65.4% of male patients, possibly affected by the small number of patients included. In the cohort studies mentioned above, the SPRS was only used in the German cohort. [1,21] Mean SPRS scores were slightly higher in the German cohort with 18.2 points when compared with 17.35 points in our cohort.

Thirty patients (54.6% of the total cohort) returned for a FU, and demographics did not significantly differ from the total cohort. The mean FU time of 392.73 days was slightly longer than 1 year. In this small sample, SPRS scores significantly increased for the total cohort and both HSP subforms. The mean progression of the SPRS score was 1.17 points, and the SRM was 1.11 in the total cohort, depicting the ability of the SPRS to detect change over time in patients with HSP. Progression rates in 1 year did not differ significantly between pHSP and cHSP. Considering that HSP and hereditary ataxia are more increasingly regarded as a continuous ataxia-spasticity disease spectrum rather than different entities,[3] we also included the SARA in our study to be able to examine ataxia as a possible contributing factor to impairment and disease progression. We were able to show a significant progression for the whole cohort and the subgroup of cHSP. The SRM, however, was significantly lower than the one for the SPRS. When looking at the

linear model with the SPRS as a dependent variable, 3 SARA items (gait, finger chase, and heel-shin slide) significantly influence the SPRS score. The gait item alone, however, explains almost 80% of the model. Furthermore, items in the SARA score are weighted differently, and the gait and stance part constitutes 14 points, whereas all other items only make up 26 points.[5] Thus, it seems likely that the gait item significantly contributes to the high correlation (Figure 2, A and B). The remaining 2 items, however, represent ataxia of the extremities and are thus more often affected in cHSP. Considering that more than 34.5% of our patients in our cohort and almost 30% in the German cohort[1] had a phenotype of spastic ataxia, this further warrants consideration to implement the SARA score in genetically and clinically diverse HSP cohorts.

Several studies in the past few years have focused on QoL in HSP, and all were able to show reduced QoL in HSP.[22-24] In this study, QoL was measured with the use of the EQ-5D questionnaire. As there is no Austrian control data set available, we compared it with the latest data set from 5 European countries (Germany, France, Spain, Italy, and the United Kingdom). We were able to confirm reduced QoL in our patients with HSP in all EQ-5D domains as well as the EQ-VAS (69.74 and 77.9).[25] Rates of depression were also higher in our sample when compared with normative data.[26]

PROMs, especially those measuring QoL and the impact of diseases on the ADLs, have become increasingly important for various diseases. In Friedreich Ataxia, for example, the disease-specific PROM FARS-ADL depicts impairments in the ADLs most affected by the disease. It has been shown to exhibit higher responsiveness to change over time than the SARA score and was thus considered a useful outcome measure for possible future trials in this disease.[9] Comparisons between PROM and CROM in HSP, however, are scarce. Only 3 studies evaluated clinical assessments and QoL in HSP. Two of these studies used the generic SF-36 questionnaire and found a significant inverse correlation of QoL and disease severity.[10,11] The third study used a questionnaire specifically developed for children with

cerebral palsy (CPCHILD), a disease similar to HSP,[27] in a cohort of patients with AP-4–associated HSP and was also able to demonstrate a significant correlation.[15] In this study, we were, however, only able to show a correlation between QoL and impairments in ADL. QoL did not correlate with disease severity, as measured by the SPRS and SARA, nor disease duration or dependence on walking aids. A linear model did confirm this result and was again not able to show any clear effect of disease severity or other demographic factors on QoL. Furthermore, CRO clearly showed clinical progression over the course of 1 year, but there were no significant impacts of this progression on the EQ-VAS or the EQ-5D domains. A study in spinocerebellar ataxias was able to detect change in the EQ-5D over a period of 8 years,[28] suggesting that the FU period in our study might be too short for the EQ-5D to detect significant changes. Despite the short FU, we still conclude that the EQ-5D questionnaire is useful to compare QoL of patients with HSP with population standards but does not seem to be specific enough for HSP symptoms. Although the SF-36 might be better suited for studies in HSP, it is still a generic tool not designed to specifically depict disease severity or progression in HSP. The CPCHILD questionnaire was developed for cerebral palsy and contains items relevant to HSP, but the average time to complete the questionnaire of around 35 minutes hampers practicability for study purposes.[15] In our opinion, a more compact and disease-specific PROM for HSP, which is able to show the full effect of the disease on patients, is required for possible future trials in HSP. To find suitable items for a possible future scale, we checked for relevant correlations between subdomains of the EQ-5D and BI with the SPRS. The subdomains self-care, usual activities, and mobility of the EQ-5D questionnaire and most items of the BI (all but eating and toilet use) showed significant correlations with the SPRS in our study. Symptoms of HSP, such as spasticity and weakness, clearly affect these areas of everyday life in our cohort. However, none of the used scales include all domains, and despite the EQ-5D covering most parts, in their current form the subdomains are not suitable to depict change in HSP (Table 4). Thus, we suggest including items covering all important domains in an adapted form suitable to detect change in the development of a HSP-specific PROM. Considering the properties of the FARS-ADL in Friedreich Ataxia and a substantial overlap in the experienced patients' disabilities between the 2 diseases, especially regarding mobility, personal hygiene, and bladder function, we endorse the evaluation of an adapted HSP-ADL score in future multicentric studies. Such a scale might constitute a suitable option as a disease-specific PROM in HSP and thus be applied as an SPRS subscale equivalent of the FARS-ADL.

In conclusion, we here present an Austrian cohort of patients with HSP and a prospective study evaluating the progression of PROs and CROs in a natural history study of patients with HSP. We show that, in line with the existing literature from other countries, the rate of pHSP in Austria is comparable with large studies from Germany and France.[1,21] Despite the limited sample size and the short observational period, we

highlight the suitability of the SPRS as a disease-specific tool to depict clinical progression in HSP. Our results furthermore warrant the consideration of additional scales such as SARA to portray the full clinical picture of HSPs apart from pure spastic paraplegia, especially considering the growing body of evidence of a spastic ataxia disease continuum. Of note, spastic ataxias account for 35% in this HSP cohort and 30% in a large German cohort.[1] Generic PROs clearly highlight the impact of HSPs on QoL, functional ability, and depression. Although some correlation of SF-36 with the SPRS has been shown,[10,11] multiple widely used PROMs did not correlate with neurologic severity in our study. Altogether, these findings advocate the introduction of disease-specific PROM fully showing effects of HSP on the patients' QoL, ADL, and mental well-being. In future studies on effects of therapeutic interventions, adequate and tailored HSP-PROM will be an integral part of functional and patient-oriented clinical outcome.

Study Funding

Disclosure

Publication History

Appendix Authors

Name	Location	Contribution
Matthias Amprosi, MD	Centre for Rare Neurological Diseases, Department of Neurology, Medical University of Innsbruck, Austria	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design; and analysis or interpretation of data
Elisabetta Indelicato, MD, PhD	Centre for Rare Neurological Diseases, Department of Neurology, Medical University of Innsbruck, Austria	Drafting/revision of the manuscript for content, including medical writing for content; study concept or design; and analysis or interpretation of data
Andreas Eigentler, MD, PhD	Centre for Rare Neurological Diseases, Department of Neurology, Medical University of Innsbruck, Austria	Drafting/revision of the manuscript for content, including medical writing for content; study concept or design; and analysis or interpretation of data
Josef Fritz, PhD	Department of Medical Statistics, Informatics and Health Economics, Medical University of Innsbruck, Austria	Drafting/revision of the manuscript for content, including medical writing for content; study concept or design; and analysis or interpretation of data

Appendix *(continued)*

Name	Location	Contribution
Wolfgang Nachbauer, MD, PhD	Centre for Rare Neurological Diseases, Department of Neurology, Medical University of Innsbruck, Austria	Drafting/revision of the manuscript for content, including medical writing for content; study concept or design; and analysis or interpretation of data
Sylvia Boesch, PD, MD, MSc	Centre for Rare Neurological Diseases, Department of Neurology, Medical University of Innsbruck, Austria	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design; and analysis or interpretation of data

References

1. Schüle R, Wiethoff S, Martus P, et al. Hereditary spastic paraplegia: clinicogenetic lessons from 608 patients. *Ann Neurol.* 2016;79(4):646-658. doi: 10.1002/ana.24611
2. Harding AE. Classification of the hereditary ataxias and paraplegias. *Lancet (London, England).* 1983;321(8334):1151-1155. doi: 10.1016/s0140-6736(83)92879-9
3. Synofzik M, Schüle R. Overcoming the divide between ataxias and spastic paraplegias: shared phenotypes, genes, and pathways. *Mov Disord.* 2017;32(3):332-345. doi: 10.1002/mds.26944
4. Schule R, Holland-Letz T, Klimpe S, et al. The Spastic Paraplegia Rating Scale (SPRS): a reliable and valid measure of disease severity. *Neurology.* 2006;67(3):430-434. doi: 10.1212/01.wnl.0000228242.53336.90
5. Schmitz-Hubsch T, du Montcel ST, Baliko L, et al. Scale for the assessment and rating of ataxia: development of a new clinical scale. *Neurology.* 2006;66(11):1717-1720. doi: 10.1212/01.wnl.0000219042.60538.92
6. Schmahmann JD, Pierce S, MacMore J, L'Italien GJ. Development and validation of a patient-reported outcome measure of ataxia. *Mov Disord.* 2021;36(10):2367-2377. doi: 10.1002/mds.28670
7. Peto V, Jenkinson C, Fitzpatrick R. PDQ-39: a review of the development, validation and application of a Parkinson's disease quality of life questionnaire and its associated measures. *J Neurol.* 1998;245(S1):S10-S14. doi: 10.1007/pl00007730
8. Subramony SH, May W, Lynch D, et al, for the Cooperative Ataxia Group. Measuring Friedreich ataxia: interrater reliability of a neurologic rating scale. *Neurology.* 2005;64(7):1261-1262. doi: 10.1212/01.wnl.0000156802.15466.79
9. Reetz K, Dogan I, Hilgers R-D, et al. Progression characteristics of the European Friedreich's ataxia consortium for translational studies (EFACTS): a 2 year cohort study. *Lancet Neurol.* 2016;15(13):1346-1354. doi: 10.1016/s1474-4422(16)30287-3
10. Klimpe S, Schüle R, Kassubek J, et al. Disease severity affects quality of life of hereditary spastic paraplegia patients. *Eur J Neurol.* 2012;19(1):168-171. doi: 10.1111/j.1468-1331.2011.03443.x
11. Orsucci D, Petrucci L, Ienco EC, et al. Hereditary spastic paraparesis in adults. A clinical and genetic perspective from Tuscany. *Clin Neurol Neurosurg.* 2014;120:14-19. doi: 10.1016/j.clineuro.2014.02.002
12. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state." *J Psychiatr Res.* 1975;12(3):189-198. doi: 10.1016/0022-3956(75)90026-6
13. Creavin ST, Wisniewski S, Noel-Storr AH, et al. Mini-Mental State Examination (MMSE) for the detection of dementia in clinically unevaluated people aged 65 and over in community and primary care populations. *Cochrane Database Syst Rev.* 2016;2016(1):CD011145. doi: 10.1002/14651858.cd011145.pub2
14. Erichsen AK, Koht J, Stray-Pedersen A, Abdelnoor M, Tallaksen CME. Prevalence of hereditary ataxia and spastic paraplegia in southeast Norway: a population-based study. *Brain.* 2009;132(6):1577-1588. doi: 10.1093/brain/awp056
15. Jordan C, Geisel G, Alecu JE, Zhang B, Sahin M, Ebrahimi-Fakhari D. Disease severity and motor impairment correlate with health-related quality of life in AP-4-associated hereditary spastic paraplegia. *Neurol Genet.* 2021;7(4):e605. doi: 10.1212/nxg.0000000000000605
16. Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. *MD State Med J.* 1965;14:61-65.
17. EuroQol Group. EuroQol–a new facility for the measurement of health-related quality of life. *Health Policy.* 1990;16(3):199-208. doi: 10.1016/0168-8510(90)90421-9
18. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med.* 2001;16(9):606-613. doi: 10.1046/j.1525-1497.2001.016009606.x
19. Braschinsky M, Luus SM, Gross-Paju K, Haldre S. The prevalence of hereditary spastic paraplegia and the occurrence of SPG4 mutations in Estonia. *Neuroepidemiology.* 2009;32(2):89-93. doi: 10.1159/000177033
20. Coutinho P, Ruano L, Loureiro JL, et al. Hereditary ataxia and spastic paraplegia in Portugal: a population-based prevalence study. *JAMA Neurol.* 2013;70(6):746-755. doi: 10.1001/jamaneurol.2013.1707
21. Méreaux JL, Banneau G, Papin M, et al, the French SPATAX clinical network. Clinical and genetic spectra of 1550 index patients with hereditary spastic paraplegia. *Brain.* 2022;145(3):1029-1037. doi: 10.1093/brain/awab386
22. Braschinsky M, Rannikmäe K, Krikmann U, et al. Health-related quality of life in patients with hereditary spastic paraplegia in Estonia. *Spinal Cord.* 2011;49(2):175-181. doi: 10.1038/sc.2010.61
23. Servelhere KR, Faber I, Saute JAM, et al. Non-motor symptoms in patients with hereditary spastic paraplegia caused by SPG4 mutations. *Eur J Neurol.* 2016;23(2):408-411. doi: 10.1111/ene.12839
24. Rattay TW, Boldt A, Völker M, et al. Non-motor symptoms are relevant and possibly treatable in hereditary spastic paraplegia type 4 (SPG4). *J Neurol.* 2020;267(2):369-379. doi: 10.1007/s00415-019-09573-w
25. Janssen MF, Pickard AS, Shaw JW. General population normative data for the EQ-5D-3L in the five largest European economies. *Eur J Health Econ.* 2021;22(9):1467-1475. doi: 10.1007/s10198-021-01326-9
26. Kocalevent RD, Hinz A, Brähler E. Standardization of the depression screener patient health questionnaire (PHQ-9) in the general population. *Gen Hosp Psychiatry.* 2013;35(5):551-555. doi: 10.1016/j.genhosppsych.2013.04.006
27. Hedera P. Hereditary spastic paraplegia overview. In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens K, et al, eds. *GeneReviews(R):* University of Washington, Seattle; 1993.
28. Jacobi H, du Montcel ST, Bauer P, et al. Long-term evolution of patient-reported outcome measures in spinocerebellar ataxias. *J Neurol.* 2018;265(9):2040-2051. doi: 10.1007/s00415-018-8954-0