



**Universiteit  
Leiden**  
The Netherlands

## **Development and validation of the first needs-based quality of life measure for individuals with manifest Huntington's disease**

Spray, I.; Grzeda, M.T.; Thorpe, J.; Johnstone, E.; Hannemann, W.; Squitieri, F.; ... ; Kuvarzina, O.

### **Citation**

Spray, I., Grzeda, M. T., Thorpe, J., Johnstone, E., Hannemann, W., Squitieri, F., ... Kuvarzina, O. (2026). Development and validation of the first needs-based quality of life measure for individuals with manifest Huntington's disease. *Journal Of Neurology*, 273(7). doi:10.1007/s00415-026-13898-8

Version: Publisher's Version  
License: [Creative Commons CC BY 4.0 license](#)  
Downloaded from: <https://hdl.handle.net/1887/4307095>

**Note:** To cite this publication please use the final published version (if applicable).



# Development and validation of the first needs-based quality of life measure for individuals with manifest Huntington's disease

Isobel Spray<sup>1</sup> · Mariusz T. Grzeda<sup>1</sup> · Jeanette Thorpe<sup>1</sup> · Ellie Johnstone<sup>1</sup> · Wiebke Hannemann<sup>2</sup> · Ferdinando Squitieri<sup>3</sup> · Giulia Giancaterino<sup>4</sup> · Paola Emilia Mina<sup>4</sup> · Barbara D'Alessio<sup>4</sup> · Astri Arnesen<sup>5</sup> · Jiří Klempíř<sup>6</sup> · Olga Ulmanová<sup>6</sup> · Pearl van Lonkhuizen<sup>7,8,9</sup> · Anne-Wil Heemskerk<sup>7,9</sup> · Jennifer Hoblyn<sup>10,11</sup> · Emer O'Malley<sup>10</sup> · Ramona Moldovan<sup>12,13,14</sup> · Georg Bernhard Landwehrmeyer<sup>2</sup> · Stephen P. McKenna<sup>1</sup> · Alzbeta Mühlbäck<sup>2,6,15</sup> on behalf of the HEALTHE-RND consortium

Received: 19 March 2026 / Revised: 14 May 2026 / Accepted: 19 May 2026  
© The Author(s) 2026

## Abstract

**Background** Huntington's disease (HD) is a progressive, hereditary neurodegenerative disorder currently without curative treatments, thus making quality of life (QoL) an important outcome for clinical care and therapeutic evaluation. Existing HD-specific patient-reported outcome measures (PROMs) do not adequately capture the experiences of individuals with manifest HD. This study presents the multinational development and validation of the Huntington's Disease Manifest Quality of Life measure (HD-mQoL), the first needs-based, disease-specific PROM for assessing QoL in individuals with manifest HD.

**Methods** Development followed three stages: (1) generation of measure content and translations; (2) testing of face and content validity; and (3) psychometric validation using data from a large international survey analysed with Rasch Measurement Theory (RMT) and Classical Test Theory (CTT).

**Results** The measure was completed by 238 individuals with manifest HD from the Czech Republic, Germany, Ireland, Italy, and the UK (59% male; age range 20–83 years). Rasch analysis reduced 49 items to a final set of 23, demonstrating good model fit (item–trait interaction  $\chi^2=0.391$ ), unidimensionality, no differential item functioning, no local dependency, and excellent reliability (Cronbach's  $\alpha=0.91$  at timepoint 1, 0.92 at timepoint 2; test–retest  $r=0.87$ ). The final measure showed minimal floor and ceiling effects and correlated moderately to strongly with relevant Nottingham Health Profile domains. HD-mQoL scores effectively differentiated subgroups by self-rated disease severity and general health ( $p<0.001$ ).

**Conclusions** The HD-mQoL is a robust, needs-based measure of QoL, suitable for international use in clinical practice and trials assessing treatment value from the patient's perspective.

**Keywords** Huntington's disease · Patient-reported outcome measures (PROM) · Quality of life (QoL) · Needs-based QoL · Patient value

## Introduction

Huntington's disease (HD) is a rare, progressive, autosomal dominant neurodegenerative disorder that affects an estimated 10 individuals per 100,000 in Western populations [1]. It typically manifests in adulthood and is characterised by a triad of motor, cognitive, and psychiatric symptoms [2]. Once manifest, the disease progresses inexorably, leading to increasing dependence on caregivers and premature death, typically occurring two decades after the onset of disease-specific motor symptoms [2]. The personal, social, and

economic consequences of HD are profound [3, 4], driven by the progressive loss of independence and increasing societal isolation experienced by patients and their families [5]. Currently, several disease-modifying strategies are being investigated, with recent findings from studies of AMT-130 (uniQure) suggesting that gene-silencing therapy may slow disease progression in HD. However, such approaches are not yet widely available to patients [6, 7]. Consequently, multidisciplinary management continues to represent the cornerstone of HD care, addressing the complex motor, cognitive, and psychiatric manifestations that define the disease [8]. In this context, suitable outcome measures are

Extended author information available on the last page of the article

essential for disease assessment, management, and treatment evaluation.

Patient-reported outcome measures (PROMs) provide a structured method for capturing the patient's perspective on disease burden and treatment outcomes and are increasingly incorporated into neurological research and clinical trials [9]. The importance of these measures is highlighted by the U.S. Food and Drug Administration (FDA) guidance, which emphasises the need to advance the collection of robust and meaningful patient input to inform medical product development and regulatory decision-making [10]. Symptom- and function-based PROMs are valuable for evaluating disease or health status, referring to an individual's functioning across physical, mental, and social domains, and can be reported either by the individual or by proxy perspective [9]. However, evidence suggests that such measures may not fully capture the lived experience of individuals with HD [11]. Anosognosia, a lack of awareness of deficits, has been shown to affect patients' insight into certain symptoms [11, 12]. In addition, research shows that ratings of symptom importance in HD vary from patient to patient [13]. Consequently, relying solely upon symptom and functioning measures to assess patient value of treatments risks making inaccurate assumptions about HD patients' experience of their disease [14].

Quality of life (QoL), in contrast, extends beyond the description of functional abilities to encompass an individual's subjective perception and evaluation of their well-being [15]. QoL assessment, therefore, can capture a more holistic understanding of how HD impacts everyday life. It is well-documented that HD can have a significant impact on QoL [4] with findings suggesting that QoL is exceptionally poor for HD patients, when compared to patients with other neurodegenerative conditions [16].

To date, three HD-specific, validated measures of health-related quality of life (HRQoL) have been developed. The first disease-specific patient-reported HRQoL instrument, the Huntington's Disease Health-related Quality of Life questionnaire (HDQoL), was developed and validated by Hocaoglu, Gaffan, and Ho (2012b), with a corresponding proxy version published in the same year [17]. Shortly thereafter, the Huntington Quality of Life Instrument (H-QoL-I) was developed and validated, based on findings from the European HD Burden Study [18]. More recently, HDQLIFE was introduced by Carozzi et al. [19], integrating the validated Neuro-QoL and PROMIS measurement systems with five additional HD-specific scales to provide a comprehensive assessment of HRQoL in this population. These instruments were developed using heterogeneous samples that included individuals at risk for HD, premanifest patients, and those with manifest disease, which included individuals at varying stages of disease progression. However, emerging evidence suggests that further stratification may improve

measurement precision, as patient experiences and perceptions of QoL differ across disease stages [4, 20]. Consequently, some items may be less relevant to particular subgroups, potentially limiting the sensitivity and applicability of these tools in specific clinical or research contexts [11]. A further limitation is that these measures are multidimensional and do not yield a single summary score, complicating interpretation and longitudinal measurement of outcomes. This poses challenges for routine clinical practice and clinical trial research settings, for which questionnaire length and ease of score aggregation are important factors [21, 22]. Long or complex tools also burden patients, particularly HD patients with cognitive impairments [11].

There is, therefore, a continuing need to develop additional measures to assess QoL in HD, ensuring that they accurately reflect stage-specific challenges and the lived experiences of affected individuals. QoL is a broader construct that extends beyond medical aspects to encompass social, relational, and occupational facets, including interpersonal relationships and career experiences [23]. This also aligns with the World Health Organization's (WHO) definition of QoL as "individuals' perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards, and concerns" [24].

The needs-based model of QoL, used in the current project, encompasses this definition by using a holistic theoretical framework. Developed in a study of patients diagnosed with depression, the model proposes that QoL is derived from the ability to meet basic human needs [25]. The model has since been successfully applied across several disease areas in the development of disease-specific QoL measures [26–28]. Whilst HRQoL is inherently multidimensional, an advantage of the needs-based model is that it enables the development of a unidimensional measure [29]. Although individual items reflect different areas of need, the overall QoL score represents a person's ability to fulfil these fundamental needs. A higher score on the measure indicates greater difficulty in meeting needs, and, therefore, poorer QoL. This approach allows for the calculation of a single composite score, which is valuable for monitoring change over time [30]. Needs-based measures have, therefore, been successfully applied in research and clinical trials across a range of disease groups [31–33].

Needs-based measures also offer the unique advantage of assessing patient-perceived value across a variety of intervention types. This is particularly important for outcome assessment in HD, as symptoms, rate of progression, and age of onset can vary widely between individuals [2]. Consequently, treatment approaches are highly heterogeneous, encompassing pharmacological therapies as well as non-pharmacological interventions, such as physiotherapy, occupational therapy, and speech therapy. In this context,

healthcare professionals emphasise the importance of keeping the patient's voice central in the design and delivery of multidisciplinary care [8].

This paper outlines the development and validation of the first needs-based QoL measure for individuals with manifest Huntington's disease (HD-mQoL), utilising Rasch modelling to examine scaling characteristics and classical test theory to evaluate traditional psychometric properties of the measure.

## Methods

### Study population

Participants were recruited as part of the multinational HEALTHE-RND project, conducted within the framework of the EU Joint Programme-Neurodegenerative Disease (JPND) research initiative. Recruitment took place across six participating countries: the Czech Republic, Germany, Italy, Ireland, the Netherlands, and the United Kingdom.

Eligible participants were adults aged 18 years or older who demonstrated adequate comprehension of the primary language spoken in their country (Czech, Dutch, English, German, or Italian) and were able to provide informed consent. Inclusion criteria required a confirmed genetic diagnosis of HD, defined by the presence of a cytosine-adenine-guanine (CAG) repeat expansion of  $\geq 36$  in the Huntingtin (*HTT*) gene, together with clinical motor features consistent with manifest HD, as indicated by a Diagnostic Confidence Level (DCL) score of 4 on the Unified Huntington's Disease Rating Scale (UHDRS)[34].

Exclusion criteria included the presence of any serious psychiatric, neurological, sensory, or metabolic disorder, or any other comorbid condition that could influence QoL. Individuals residing in specialised care facilities were also excluded. The same eligibility criteria were applied consistently across all stages of the study. Participants were identified and recruited through local HD centres in each participating country, with recruitment coordinated by local research teams in accordance with standardised procedures. Ethical approval for the study was obtained from the Research Ethics Committee (REC) for London—West London and GTAC REC in the United Kingdom, and from the respective ethics committees in each participating European country.

### Study stages

The development and validation of the HD-mQoL measure were conducted in three sequential stages. The first stage (1) involved item content generation, aimed at identifying and developing a comprehensive pool of items reflecting the

lived experience of individuals with manifest HD. The second stage (2) focused on assessing face and content validity to ensure that items were conceptually appropriate, clearly worded, and representative of the intended construct. In the third stage (3), a draft version of the measure was administered to evaluate its psychometric properties, enabling refinement of the measure and confirmation of the final set of items. An overview of these stages is provided in Fig. 1.

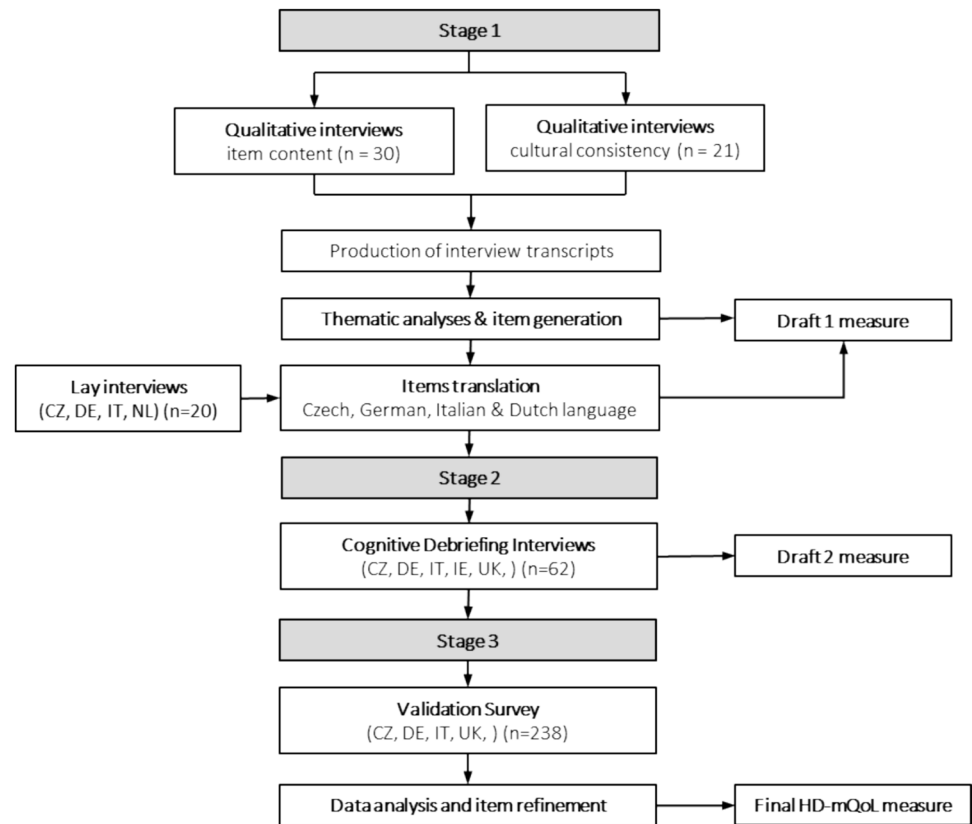
### Stage 1: Item content generation

Item content was generated in the United Kingdom and Ireland through semi-structured qualitative interviews conducted via secure video-conferencing platforms. Participants were invited to freely discuss the ways in which HD affected their daily lives. To ensure cross-cultural relevance, the resulting themes were verified against supplementary qualitative interviews conducted in the Czech Republic (CZ), Germany (GER), Italy (IT), and the Netherlands (NL). Interviews conducted in the United Kingdom (UK) and Ireland (IR) were audio-recorded and transcribed verbatim, with all identifying information removed to preserve participant confidentiality. Data were analysed using theoretical thematic analysis [35], guided by the needs-based model of QoL [25], to identify themes of need fulfilment deficit. Additional methodological details relating to Stage 1 have been described previously [36]. Based on the aforementioned thematic framework, an initial pool of items was developed to reflect the key need-deficit themes identified in the qualitative data.

### Translation and language adaptation

A dual-panel translation method was employed for the translation of draft items into four languages (Czech, Dutch, German, and Italian). This method is an established way of translating PROMs [37], confirming semantic equivalence and comprehensibility. Each translation of the draft measure required a translation call, conducted with a bilingual representative. The researcher discussed in detail each item and instruction to ensure the representative was familiar with the intended meanings. As the bilingual expert translated each item, providing a back translation, discussions took place to determine whether it was suitably equivalent in meaning, with alternatives given if a direct translation was inappropriate. Once translations, or translation options, were agreed upon by both parties, translations were taken to a lay panel ( $n = 5$  per country) to provide a more objective perspective on the comprehensibility. The lay sample included participants from a variety of professions and a range of ages. Lay interviewees had no diagnosis of HD, as item content relevance was not assessed at this stage. Where interviewees

**Fig. 1** Stages in the Development and Validation of the Huntington's Disease Manifest Quality of Life (HD-mQoL) measure. CZ (Czech Republic), DE (Germany), IT (Italy), IE (Ireland), NL (Netherlands), UK (United Kingdom), n (number)



could not decide upon a preferred translation, options would be taken for further testing in the next stage.

### Stage 2: Assessment of face and content validity

Cognitive debriefing interviews (CDIs) are used to refine measure items and instructions with the input of members of the relevant disease group, using a semi-structured interview assessing acceptability, comprehensibility and item content relevance [38, 39]. The interviewer guided HD patients through the draft measure, noting hesitations or comments. After completion of the measure, the researcher consulted an interview schedule containing questions relating to specific items that surfaced either a.) in lay interviews for countries in which a translation was required or b.) during the item generation process and/or in other CDI interviews. Participants were also encouraged to provide any additional feedback or suggestions for improvement.

### Stage 3: Validation survey

Eligible participants were sent a survey pack by their local HD centres in the participating sites, with recruitment guided by the local research teams in the respective centres. The survey pack contained a letter of invitation, a participant information sheet, a demographic questionnaire, the

draft version of the HD-mQoL and a copy of the Nottingham Health Profile (NHP) [40] as a comparator measure. The NHP is a 38-item measure of distress consisting of six scales: energy, pain, emotional reactions, sleep disturbance, social isolation and physical mobility. Patients who completed and returned the first pack were sent a second pack 2 weeks later, to assess reproducibility. This included the HD-mQoL and a shorter demographic questionnaire.

Participants were further examined using the standardised Unified Huntington's Disease Rating Scale (UHDRS) assessment protocol [34], which includes neuropsychological, motor, and functional evaluations. All tests and clinical examinations were performed by trained and experienced personnel (e.g., clinicians and psychologists) in the participants' local language. Motor impairment was assessed using the Total Motor Score (TMS), which evaluates the presence and severity of neurological motor symptoms associated with HD. The TMS ranges from 0 to 124 points, with higher scores indicating more severe motor impairment [41]. Functional status was assessed using three components of the UHDRS [34]: the Total Functional Capacity (TFC) [42], the Functional Assessment Scale (FAS), and the Independence Scale (IS). The TFC consists of five items—occupation, finances, domestic chores, activities of daily living, and level of care—and ranges from 0 to 13 [42]. The FAS includes 25 yes/no items assessing the ability to perform

common daily activities (range 0–25). The IS provides a single score reflecting the level of independence, ranging from 10 to 100. For all three measures, lower scores indicate greater functional impairment. The cognitive domain was assessed using the Symbol Digit Modalities Test (SDMT). Biological disease burden was estimated using CAPI00, calculated as the product of excess CAG repeat length and age, as described by Warner et al. [43]. The Prognostic Index for HD (PIN) was calculated as a composite measure incorporating age, CAG repeat length, motor signs, and cognitive performance to estimate disease progression risk [44].

## Statistical analysis

Analyses using the data collected from Stage 3 were conducted within two complementary frameworks: first, Rasch Measurement Theory (RMT), to examine the measure's structural validity and measurement precision, followed by Classical Test Theory (CTT), to assess reliability and construct validity.

### Determining the final item set: Rasch measurement theory

RMT was selected as the principal analytic framework due to its capacity to construct unidimensional scales with interval-level measurement properties and its provision of detailed item performance diagnostics. This approach enables the identification of a parsimonious set of items that function as a unidimensional measure, demonstrating strong construct validity and robust measurement properties.

RMT provides a rigorous statistical method for transforming ordinal-level questionnaire responses into interval-level measurement, enabling the construction of measures that support objective comparisons across individuals and groups [45, 46]. Within this framework, both respondents and items are positioned along the same latent continuum representing the underlying theoretical construct, in this case QoL. This alignment allows for meaningful interpretation of the patient's level of the trait relative to the difficulty level of each item.

To validate the HD-mQoL measure, a series of diagnostic tests were applied to assess the extent to which the data conformed to the assumptions of the Rasch model. Each test targeted a specific psychometric property of the measure. The following were assessed in the evaluation process:

- (1) *Fit statistics*. Overall model fit was assessed using the item–trait interaction chi-square test, evaluating the invariance of item difficulty across different levels of the trait. These tests examined whether the observed item responses across groups of respondents with varying trait levels align with the expected responses under the Rasch model [47]. In addition, standardised item and person residuals were examined to evaluate the discrepancy between observed responses and model expectations. In this case, the chi-square and  $F$  tests were used to detect misfitting items. Fit residuals within the range  $\pm 2.5$  and non-significant  $p$  values from the chi-square and  $F$  tests were considered indicative of acceptable data-model fit. Furthermore, item characteristic curves (ICCs) and observed response patterns were inspected to assess the degree of agreement between the model expectations and the data collected [48].
- (2) *Person-separation index (PSI)*. The PSI was used to evaluate the ability of the measure to discriminate between individuals with varying levels of QoL. The threshold of 0.70 was set as the minimum acceptable value, indicating the measure can reliably distinguish between at least two strata of respondents [49]. This threshold is generally accepted as the minimum level required for group-level interpretation [50].
- (3) *Local independence*. Local independence checks were conducted by analysing residual correlations. The violation of this assumption suggests that item pairs share some variance beyond what is explained by the underlying latent trait, indicating potential redundancy or multidimensionality. Significant residual correlations were interpreted as evidence of local dependency in this study. A threshold of 0.20 above the average residual correlation across all item pairs was used to flag potential issues with local independence [51]. Item pairs exceeding this threshold required further action, with either one item being removed or both correlated items being grouped into a subset.
- (4) *Unidimensionality*. To evaluate the assumption of unidimensionality, a paired  $t$  test procedure was employed. This involves comparing person estimates derived from two subsets of items that exhibit the most divergent loadings obtained from a principal component analysis (PCA) of residuals [52]. If the measure is unidimensional, then individual person estimates calculated from different subsets of items (even from those with the most contrasting loadings) should remain consistent. The first subset contains items with the highest positive loadings, with the other consisting of items with the highest negative loadings on the first residual component. The person estimates on the latent trait are then derived separately for each subset. Paired  $t$  tests are conducted to compare these estimates for each individual. If fewer than 5% of the  $t$  tests yield statistically significant differences ( $p < 0.05$ ), unidimensionality is supported. This threshold is commonly used as the criterion to indicate that any observed differences are likely due to chance rather than a violation of unidimensionality. In addition, a 95% confidence interval for

a binomial test of the proportion is calculated based on the observed number of significant  $t$  tests. It is assumed that the lower bound of this confidence interval should fall below the 5% threshold to further support the assumption of unidimensionality.

- (5) *Differential item functioning (DIF)*. DIF analysis was performed to assess whether item parameters remained invariant across different groups of patients. In this study, the DIF was assessed in relation to selected respondent characteristics. Items exhibiting significant DIF were flagged as potentially biased, indicating that their inclusion in the measure could affect measure validity across diverse populations [53]. An ANOVA of standardised residuals was carried out to examine DIF by gender (males vs females), age group (below the median age vs the median age and above), country (UK, DE, CZ, IT), living situation of the patient (whether lives with a spouse/partner or other living situation), ethnicity (dominant majority of the given country vs other backgrounds), employment status (employed vs not employed) and care (care provided vs not). An ANOVA  $p$  value, adjusted using the Bonferroni method, was used to determine the presence of DIF.
- (6) *Targeting and item locations*. The alignment between item difficulty and person ability was visually inspected using person–item maps. This analysis assessed whether the items covered the full range of the latent trait and whether the item hierarchy matched theoretical expectations.

The analyses conducted within the Rasch modelling framework followed an iterative process aimed at refining the instrument to ensure both statistical and theoretical strength. At each stage of the analysis, items that demonstrated misfit, defined as exhibiting response patterns inconsistent with the Rasch model, were carefully examined. Depending on the nature and extent of the misfit, one of two corrective actions was taken. Either the problematic item was removed from the item pool, or a data transformation was applied to address the issue. Importantly, only one modification was implemented per iteration to isolate the effects of each change. Following each adjustment, the dataset was re-evaluated to verify that the assumptions of the Rasch model were met. This cycle of evaluation and refinement continued until the remaining set of items demonstrated acceptable fit statistics and aligned with theoretical expectations.

All statistical tests were Bonferroni-adjusted to control for the increased risk of Type I error due to multiple comparisons. This conservative correction ensured that the probability of falsely identifying a misfitting item remained within acceptable limits.

## Classical test theory

*Internal consistency*. The psychometric properties of the measure were further evaluated using CTT. Internal consistency reliability was assessed using Cronbach's alpha coefficient, which estimates the degree to which items within the measure are inter-related. A high alpha value indicates that the items function cohesively, supporting the reliability of the measure [54].

*Test–retest*. Time stability of the measure was assessed using Spearman's rank correlation coefficient, calculated from raw scores obtained from the same participants at baseline and follow-up. Participants who reported significant changes in their perceived general health or disease severity between timepoints were excluded from this analysis to ensure the stability of the underlying construct. Given the ordinal nature of the data, Spearman's rho was chosen as the appropriate non-parametric measure. A correlation coefficient of 0.85 or higher is generally considered acceptable for instruments intended for use in clinical trials or for monitoring individual patients over time [55, 56].

*Convergent validity*. This was evaluated by examining the relationship between scores and the subscales of the NHP. It was hypothesised that moderate to high correlations would be observed between the HD-mQoL and the NHP domains most relevant to HD, for example, impaired mobility and emotional well-being.

*Discriminative (known-groups) validity*. This was assessed by comparing HD-mQoL scores across groups of participants categorised by self-reported disease severity (mild, moderate, severe, very severe) and general health status (very good, good, fair, poor). To test for statistically significant differences between these independent groups, the Kruskal–Wallis one-way analysis of variance for comparisons involving three or more groups was used.

Psychometric testing was performed using the SPSS v29.0 statistical package. Rasch modelling was performed in RUMM2030 software v5.4.

## Results

### Item generation

A total of 30 qualitative interviews (UK  $n=25$ , IE  $n=5$ , aged 33–75, 56.7% male) were conducted to generate item content. A further 21 qualitative interviews took place in the Czech Republic ( $n=5$ ), Germany ( $n=7$ ), Italy ( $n=5$ ) and the Netherlands ( $n=4$ ) (aged 36–67, 33.3% male). Summary reports from these supporting interviews provided sufficient evidence that themes were consistent across countries and not culturally bound.

Analyses identified 597 statements reflecting the impact HD has on the lives of participants. A total of 11 themes related to how HD inhibited need fulfilment were established using these statements. Amongst the most frequently discussed were ‘autonomy’, ‘safety and security’, ‘role’, ‘interests and hobbies’, and ‘esteem’. Further details regarding the qualitative analysis results have been published previously [36].

Based on these themes and associated statements, 81 draft items were initially constructed. Items showing conceptual overlap or redundancy were removed after expert review, resulting in 50 draft items representing all major themes, which were taken forward to the next stage.

### Language adaptation

The 50 draft items were translated into Czech, Dutch, German, and Italian in preparation for cross-national testing using a dual-panel translation method. Initial translation panels revealed no major difficulties, though some linguistic nuances required adjustment. For example, the phrase ‘follow stories on television’ was problematic in the Czech Republic, Germany and the Netherlands: the verb “follow” was interpreted too literally in Czech and German, and “stories” translated ambiguously in Dutch. The term “storyline” was ultimately adopted as the preferred alternative.

Lay testing (n = 5 per country, total n = 20) confirmed comprehensibility and that item meaning was maintained. For items and instructions in which the language was unclear, alternative translations were offered. One item, ‘I’m limited in the places I can go’, was removed at this stage, as it was difficult to translate and had significant conceptual overlap with the item ‘I can’t go to the places I want to go’. A total of 49 draft items remained after this stage.

### Assessment of content and face validity

A total of 62 CDIs took place to assess for face and content validity. The Netherlands was unable to continue participation from this point due to a lack of resources and, therefore, did not complete any CDIs. Across the remaining five countries (Czech Republic, Germany, Ireland, Italy, and the United Kingdom), 62 participants with manifest HD (54.8% male; aged 23–78 years, mean = 49.4) completed CDIs. The majority were unemployed (71.0%), reported good health (45.2%), and described their disease severity as mild (71.0%). Full sample characteristics by country are presented in the Supplementary Material Table 1.

After all interviews were complete and reports were reviewed, items that were consistently problematic across countries were flagged for removal if items were also identified as statistically misfitting in the next stage. Several items were adapted in response to CDI feedback. In the UK, 1

instruction and 11 items were adapted; in the Czech Republic, 1 instruction and 3 items were changed; in Germany, 1 instruction and 6 items were adapted; and in Italy, 8 items were changed. At this stage, one item was also removed, as it was perceived to be a truism (*I can’t control what is happening to me*), and one was added, as it was noted that this item more accurately reflected statements from the qualitative interviews (*I don’t answer the door to people I’m not expecting*) when compared to an existing item (*I don’t answer the door to anyone*). While ostensibly both items present with significant overlap in content, redundancy was purposefully built into these two draft items to facilitate the selection of the most appropriately worded item when examined in the next stage. A 49-item draft measure was taken to the final stage.

### Validation survey

The validation survey was completed by 238 participants with manifest HD (59% male; age range 20–83 years; see Table 1 for demographic characteristics). No major demographic differences were identified between participants recruited across the participating countries. Of these, 181 participants (76%) completed the second administration 2 weeks later to assess test–retest reliability. Ireland did not participate in this phase of the study due to a lack of resources.

The study cohort displayed clinical characteristics typical of manifest HD, with a median CAG repeat length of 43 (interquartile range IQR 42–45). Motor impairment was moderate, with a median UHDRS–TMS of 30 [21–41]. Functional impairment was evident, reflected by a median TFC score of 9 [7–12], a median FAS score of 20 [16–24], and a median IS of 80 (70–90). Cognitive processing speed was reduced, as indicated by a median SDMT score of 23 [17–31]. Disease burden was substantial, with CAP100 values showing a median of 110.02 (IQR 99.61–120.26) and a PIN score with a median of 3.02 (2.17–3.81), consistent with a cohort predominantly composed of individuals with manifest HD. Detailed clinical characteristics are provided in Supplementary Material Table 2.

### Results of statistical analyses.

The application of RMT to the initial pool of 49 candidate items resulted in the removal of 26 items to establish the final version of the HD-mQoL.

#### Item reduction process.

The Rasch analysis revealed that seven items violated assumptions of the theoretical model due to having item fit residuals outside the range  $\pm 2.5$ , and significant  $p$  values in

**Table 1** Demographic characteristics of manifest HD sample in stage 3 (Validation Survey)

Characteristics	UK n=74	Czech Republic n=26	Germany n=105	Italy n=33	Total N=238
<b>Sex n (%)</b>					
Female	33 (45.2)	7 (26.9)	43 (41.0)	13 (39.4)	96 (40.5)
Male	40 (54.8)	19 (73.1)	62 (59.1)	20 (60.6)	141 (59.5)
Age median (IQR)	58 (52–66)	50 (43–55)	56 (48–63)	51 (40–59)	55 (48–63)
<b>Living situation n (%)*</b>					
Live alone	9 (12.3)	1 (3.8)	18 (17.3)	1 (3.0)	29 (12.3)
Spouse partner	58 (79.5)	13 (50.0)	72 (69.2)	23 (69.7)	166 (70.3)
Children under 18	7 (9.6)	1 (3.8)	12 (11.5)	2 (6.1)	22 (9.3)
Children over 18	8 (11.0)	5 (19.2)	18 (17.3)	5 (15.2)	36 (15.3)
Parents	5 (6.8)	7 (26.9)	8 (7.7)	6 (18.2)	26 (11.0)
Other	1 (1.4)	4 (15.4)	7 (6.7)	3 (9.1)	15 (6.4)
Prefer not to say	0 (0.0)	2 (7.7)	0 (0.0)	0 (0.0)	2 (0.8)
<b>Employment n (%)</b>					
Full time	8 (11.0)	1 (3.8)	9 (8.7)	15 (45.5)	33 (14.0)
Part time	2 (2.7)	2 (7.7)	11 (10.6)	3 (9.1)	18 (7.6)
Self-employed	0 (0.0)	0 (0.0)	1 (1.0)	1 (3.0)	2 (0.8)
Not employed	63 (86.3)	23 (88.5)	84 (80.8)	14 (42.4)	184 (78.0)
<b>Care provided n (%)</b>					
No	20 (27.4)	4 (15.4)	13 (12.6)	20 (60.6)	57 (24.3)
Yes	53 (72.6)	22 (84.6)	90 (87.4)	13 (39.4)	178 (75.7)
<b>Care provider n (%) *</b>					
Other family member	15 (20.5)	8 (30.8)	37 (35.9)	0 (0)	60 (25.5)
Spouse/partner	43 (58.9)	11 (42.3)	65 (63.1)	13 (39.4)	132 (56.2)
Professional carer	3 (4.1)	7 (26.9)	19 (18.4)	0 (0)	29 (12.3)
Other	2 (2.7)	0 (0)	8 (7.8)	0 (0)	10 (4.3)
Not applicable	20 (27.4)	4 (15.4)	13 (12.6)	20 (60.6)	57 (24.3)
<b>Disease severity n (%)</b>					
Mild	29 (40.3)	8 (30.8)	37 (36.3)	24 (75.0)	98 (42.2)
Moderate	35 (48.6)	13 (50.0)	33 (32.4)	8 (25.0)	89 (38.4)
Severe	8 (11.1)	5 (19.2)	25 (24.5)	0 (0.0)	38 (16.4)
Very severe	0 (0.0)	0 (0.0)	7 (6.9)	0 (0.0)	7 (3.0)
<b>General health n (%)</b>					
Poor	10 (13.9)	2 (7.7)	19 (18.1)	2 (6.1)	33 (14.0)
Fair	22 (30.6)	12 (46.2)	27 (25.7)	11 (33.3)	72 (30.5)
Good	33 (45.8)	10 (38.5)	44 (41.9)	12 (36.4)	99 (41.9)
Very good	7 (9.7)	2 (7.7)	15 (14.3)	8 (24.2)	32 (13.6)

\*indicates that for these variables' participants were able to tick multiple answers. *mHD* manifest Huntington's Disease, *N* number of participants, *IQR* interquartile range

**Table 2** Summary of overall Rasch model fit statistics for the final 23-item HD-mQoL measure

	Item-trait interaction chi-square test	Person-separation index (PSI)	Item-fit residuals		Person-fit residuals		Unidimensionality <i>t</i> tests: proportion of significant paired <i>t</i> tests
			Mean	SD	Mean	SD	
HD-mQoL measure	p=0.391	0.853	0.405	0.982	-0.180	0.828	0.017
Target values	>0.05	>0.70	<0.50	<1.40	<0.50	<1.40	<0.05

the statistical tests applied. Thirteen items were removed because of high residual correlations with other items, thus violating the assumption of local independence. Four items demonstrated significant DIF and were removed. Three items demonstrated more than one type of issue, such as misfit, local dependence or DIF. Finally, four items were excluded from the measure, because they were found to be redundant, offering overlapping information with other items. A review of the remaining items confirmed that all essential content areas were still adequately represented, preserving face and content validity.

**Final data-model fit.**

The overall Rasch fit statistics for the final 23-item measure are shown in Table 2. These confirm that the measure fits the Rasch model and provides robust evidence of unidimensionality. Figure 2 shows that the items covered a wide range of severities and were well-matched to the severity level of the sample. Applied tests based on the results of PCA and paired *t* tests confirmed unidimensionality of the proposed measure. High scores on the HD-mQoL measure represent a more severe impact of HD on the QoL of patients. Could you please insert Table 2 above here, at the moment it is in wrong section.

**Classical psychometric analyses.**

The final set of 23 items demonstrated excellent reliability. Internal consistency, as assessed by Cronbach’s alpha, was exceptionally high reaching 0.91 at Time 1 and 0.92 at Time 2. These values indicate a strong degree of inter-relatedness amongst the items, suggesting that they consistently measure QoL across both timepoints. Test–retest reliability measured by the Spearman correlation coefficient was also high (0.871, *p* < 0.001; *N* = 156). Such a high reliability coefficient supports the measure’s stability over time.

Table 3 shows descriptive statistics for the measures used in the study. Some floor effects were found in the HD-mQoL measure (7% of patients scored the minimum possible score), with the number of patients scoring the maximum possible score also being very small (< 1%). In contrast, significant floor effects (36.5%–73.4%) were found for the NHP, showing the limitations of applying generic PROMs to rare diseases, such as HD.

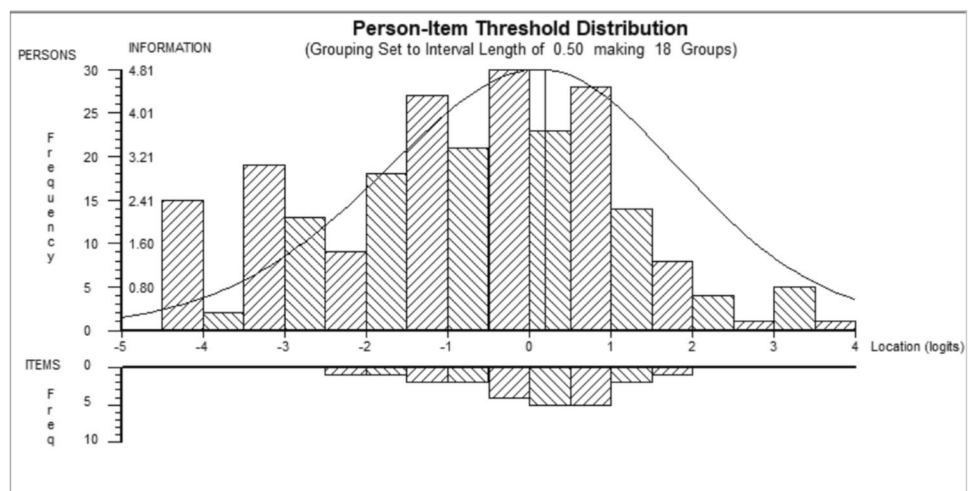
Huntington’s Disease Manifest Quality of Life (HD-mQoL) measure, Nottingham Health Profile (NHP) [40].

Table 4 shows the correlations between HD-mQoL scores and those on the NHP sections, across both physical and emotional impairments. All correlations with the HD-mQoL are positive with a moderate to high magnitude, suggesting that all these NHP domains influence the QoL of HD patients to some extent. However, stronger correlations were observed with the NHP domains of energy, emotional reactions, physical mobility and social isolation, whereas weaker

**Table 3** Descriptive and reliability statistics for the HD-mQoL and comparator measure NHP

	N	Median (IQR)	Range	% scoring minimum	% scoring maximum
HD-mQoL measure	215	8 (3–14)	0–23	7.0	0.5
<b>NHP subscales</b>					
Energy	224	33 (0–67)	0–100	39.7	17.4
Pain	222	0 (0–13)	0–100	73.4	1.4
Emotional reactions	217	11 (0–33)	0–100	37.3	1.4
Sleep	224	20 (0–40)	0–100	42.4	4.0
Social isolation	216	20 (0–40)	0–100	48.6	1.4
Physical mobility	219	13 (0–50)	0–100	36.5	0.9

**Fig. 2** Item-person map for HD-mQoL measure. The top section of the figure presents the distribution of patients on the interval measure with logit units. The bottom section shows the range of item locations



**Table 4** Spearman rank correlation coefficients were calculated between scores on HD-mQoL (time1) and NHP sections

	Correlation with HD-mQoL	p	N
<b>NHP subscale</b>			
Energy	0.669	<0.001	206
Pain	0.328	<0.001	203
Emotional reactions	0.670	<0.001	199
Sleep	0.327	<0.001	204
Social isolation	0.724	<0.001	199
Physical mobility	0.697	<0.001	201

levels of association were found with the pain and sleep scales. These results align with the typical patterns observed in people with HD, known to impact social isolation, mobility, emotional well-being and energy [57, 58].

All correlations are positive and statistically significant, supporting the convergent validity of the HD-mQoL. Huntington's Disease Manifest Quality of Life (HD-mQoL) measure, Nottingham Health Profile (NHP) (40).

Significant differences in HD-mQoL scores were observed between patients grouped by disease severity and general health ( $p < 0.001$ ) (see Fig. 3). This demonstrates that the HD-mQoL was able to distinguish successfully between the groups of known importance. Following the Kruskal–Wallis tests, pairwise Dunn's post hoc comparisons were applied and confirmed statistically significant differences between all groups. Specifically, the comparisons revealed significant differences between the 'mild' and 'moderate' groups (Bonferroni-adjusted  $p < 0.001$ ), between 'mild' and 'severe/very severe' ( $p < 0.001$ ), and between 'moderate' and 'severe/very severe' ( $p = 0.024$ ). Similar findings were observed in the post hoc comparisons based on

categories defined by the general health question. The results indicated statistically significant differences ( $p < 0.010$ ) across all pairs, with the exception of the 'fair' vs 'poor' comparison, which did not reach significance ( $p = 1.000$ ).

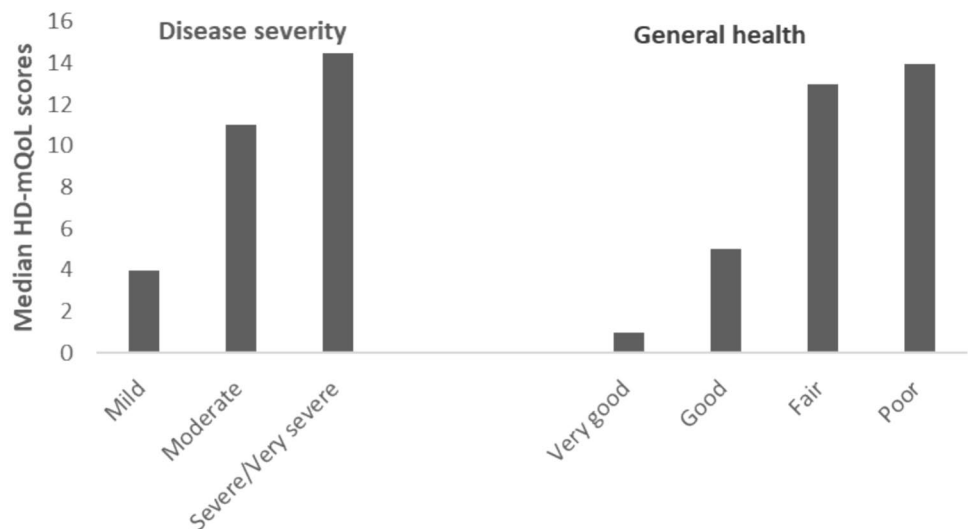
The final measure included items across a range of logit locations, with the content covering a variety of underlying needs (see Table 5).

## Discussion

PROMs are essential for evaluating the impact of disease and treatment from the patient's perspective, yet traditional symptom- and function-based instruments may not fully capture the complex, lived experience of individuals with HD, where emotional, cognitive, and social factors interact dynamically with physical decline [9]. Due to factors such as anosognosia and wide variability in how patients perceive the importance of specific symptoms, relying solely on functional measures risks misrepresenting the patient's QoL [11]. As clinical trials, aimed at slowing HD progression, continue to advance, there is a growing need for a sensitive, patient-centric tool capable of detecting subtle

**Table 5** Examples of final HD-mQoL items with corresponding underlying needs and Rasch item locations

Item	Underlying Need	Location (logit)
I'm unable to do things spontaneously	Autonomy	-0.526
I neglect my friends	Socialising	0.278
I'm unable to do my hobbies	Pleasure	-0.428
I feel vulnerable when I'm on my own	Safety and Security	0.467

**Fig. 3** Median HD-mQoL scores by disease severity and self-reported general health

yet clinically meaningful changes in QoL [59]. The present study addresses this gap through the development and validation of the first needs-based, disease-specific PROM designed specifically for individuals with manifest HD.

The HD-mQoL, a concise, 23-item, unidimensional measure of QoL, captures the extent to which fundamental human needs are met or unmet as a consequence of HD. The items were derived directly from patient interviews and refined through cross-national cognitive debriefing, ensuring strong conceptual relevance and patient-centredness. The structure and scaling properties of the HD-mQoL are supported by the application of Rasch analysis, thereby confirming unidimensionality and enabling interval-level measurement of QoL. The Rasch framework enabled a detailed evaluation of item performance, allowing for the removal of 26 draft items through an iterative process that combined statistical evidence with qualitative feedback. This integrated approach ensured that the final scale retained both psychometric robustness and conceptual relevance to the patient experience. This feature is particularly valuable for clinical and research applications, allowing for meaningful comparisons across patients, timepoints, and interventions [60]. The utilisation of dichotomous response options has also been demonstrated to enhance the ease of completion for patients with cognitive challenges and facilitate straightforward scoring and interpretation by clinicians [61].

In addition, this work was strengthened by the collaborative, international development process. Through CDIs conducted in the Czech Republic, Germany and Italy, the relevance and clarity of item content were confirmed across cultural contexts. Crucially, Rasch analysis demonstrated no evidence of DIF by country in the final measure, supporting the HD-mQoL's cross-cultural validity and its potential for use in future multinational trials. Furthermore, the successful translation process demonstrated that future linguistic adaptations can be readily implemented using established dual-panel translation procedures.

The HD-mQoL is, therefore, recommended as a robust measure for assessing QoL in patients with manifest HD. It is considered suitable for use in both clinical practice and research settings to monitor QoL changes over time.

## Limitations

First, during the postal survey phase, some missing data were observed. This was likely due to suboptimal placement of instructions and layout, as well as the length of the questionnaire, which may have caused confusion for participants with cognitive impairments or motor difficulties. Similar issues have been reported in the development of other HD-specific QoL instruments, such as the H-QoL-I [17], emphasising the need to consider cognitive processing limitations at all stages

of PROM development. To address this, the final version of the HD-mQoL incorporates clearer instructions, simplified formatting, and a substantially reduced item count following the removal of 26 items. These revisions are expected to enhance usability and completion accuracy, particularly among individuals with cognitive or fine motor challenges.

Second, although the HD-mQoL reflects the experiences of the largest HD subgroup [1], patients with mild to moderate manifest disease, it was challenging to include individuals with severe HD during item generation. Recruiting and interviewing those with advanced disease presents substantial ethical and practical challenges, including difficulties with informed consent and communication [62]. While proxy interviews could have been employed, prior evidence suggests that proxies tend to underestimate patients' QoL [63, 64], which could have compromised the patient-centred nature of the measure. This complex methodological issue warrants further discussion within the HD research community and reflects the broader challenge of incorporating the voices of individuals with severe cognitive and psychiatric impairments in research. While the generalisability of the findings to this population is, therefore, reduced, the current study did, however, integrate the development of a proxy version of the HD-mQoL into the HEALTHE-RND project [65]; the patient-derived item content but is designed to be completed by a proxy. This measure will provide a reasonable reflection of patient-centric QoL for patients with advanced HD unable to complete the questionnaire themselves.

An additional limitation regarding the study inclusion criteria should be acknowledged. At the conception of the study, the most widely accepted clinical definition of manifest HD was a DCL of 4. However, when the study commenced in 2020, the HD-ISS staging system [66] had not yet been fully developed and, therefore, could not be incorporated into the study design. While the clinical characteristics presented in Supplementary Table 2 indicate how the sample aligns with the HD-ISS stages, the use of HD-ISS staging during recruitment would have been preferable. Furthermore, a DCL of 4 excludes individuals experiencing cognitive symptoms in the absence of motor manifestations. It is hoped that the needs-based QoL measure currently under development for the premanifest population will capture both asymptomatic individuals and those in the early stages of HD presenting with cognitive and/or psychiatric symptoms only.

## Conclusion

The HD-mQoL is a psychometrically sound, patient-centred measure of QoL for individuals with manifest HD. Its needs-based content ensures aspects of QoL beyond

symptoms and functioning are captured, offering a meaningful unidimensional outcome metric for a notoriously complex, multi-faceted condition. With demonstrated cross-cultural validity, and applicability for clinical and research use, the HD-mQoL represents a significant advancement in outcome measurement for HD. The HD-mQoL thus provides clinicians and researchers with a reliable, easy-to-use tool for assessing QoL from the patient's perspective in manifest HD.

**Acknowledgements** We sincerely thank all the patients with Huntington's disease and their families who generously contributed their time and experiences to this study. We are deeply grateful to the European Huntington's Association (EHA) and the network of HD family associations in Italy, Ireland, the United Kingdom, the Netherlands, the Czech Republic, and Germany for their invaluable support throughout this project. In Italy, we extend special thanks to the patient organisations affiliated with the LIRH Foundation, including LIRH Apulia, LIRH Tuscany, LIRH Friuli, LIRH Sardinia, LIRH Umbria, Life for Huntington, and Noi Huntington, for their dedicated collaboration. We are especially grateful to all study participants for their invaluable input and active involvement throughout the study. We also wish to acknowledge the researchers, coordinators, and clinical staff at all participating sites for their sustained effort and professionalism. We would like to acknowledge the multinational HEALTHE-RND consortium for their valuable contributions to this project. HEALTHE-RND Consortium: United Kingdom: Stephen McKenna, Ian McKenna, Jeanette Thorpe, Ellie Johnstone, Isobel Spray, Mariusz Grzeda, Ramona Moldovan, Peter Foley, Jacqueline Kerr; The Netherlands: Niels H. Chavannes, Eline Meijer, Anne-Wil Heemskerk, Erik van Duijn, Susanne T. de Bot, Pearl J. C. van Lonkhuizen, Niko Vegt, Leanne Slutter, Stephanie Feleus, Esther C. Arendts, Amy Putman; Germany: Georg Bernhard Landwehrmeyer, Alžbeta Mühlböck, Peter Brieger, Rainer Hoffmann, Wiebke Frank, Rasa Fuchs, Jakob Neumüller, Sarah von Grebmer zu Wolfsthurn, Manolya Lauko, Ursula Speidel, Tanya Blinder, Florian Ciolek, Elhasna Mekhelfi, Michael Bachmeier, Inge Schram, Franziska Steck; Czechia: Jiří Klempíř, Romana Konvalinková, Eva Bezuchová, Kristýna Dolečková, Olga Klempířová, Jan Roth, Olga Ulmanová; Italy: Ferdinando Squitieri, Sabrina Maffi, Giulia Giancaterino, Chiara Di Giorgio, Melissa Casella; Ireland: Jennifer Hoblyn, Muthukumaran Thangaramanujam, Tom Burke, Emer O'Malley; Patient and Family Panel Representatives: Astri Arnesen, Michaela Winkelmann, Barbara D'Alessio, Rob Haselberg, Zdeňka Vondráčková, Patricia Towey. For licensing enquiries regarding the HD-mQoL, please contact Galen Research, Koba, 100 Barbirolli Square, Manchester, M2 3BD, . Tel: +44 161 226 4446; Email: gr@galen-research.com.

**Author contributions** S.P. McKenna, M.T. Grzeda, I. Spray, J. Thorpe, G.B. Landwehrmeyer, and A. Mühlböck conceptualised and designed the project, oversaw the execution of the research, interpreted the data, drafted the initial manuscript, and revised it for intellectual content. W. Hannemann, F. Squitieri, G. Giancaterino, P.E. Mina, B. D'Alessio, A. Arnesen, J. Klempir, O. Ulmanová, P. van Lonkhuizen, A.-W. Heemskerk, J. Hoblyn, E. O'Malley, and R. Moldovan conducted the study at their respective sites, contributed to participant recruitment, and critically reviewed the manuscript. M.T. Grzeda, I. Spray, J. Thorpe, and E. Johnstone also performed the statistical analyses. All authors read and approved the final version of the manuscript.

**Funding** Open Access funding enabled and organized by Projekt DEAL. This review was conducted as part of the HEALTHE-RND project, which was funded by the EU Joint Programme Neurodegenerative Disease Research (JPND) (grant number: 01ED1903).

**Data availability** The data that support the findings of this study are available from the corresponding author on reasonable request. The data are not publicly available due to privacy or ethical restrictions.

## Declarations

**Conflict of interest** The authors have no competing interests to declare that are relevant to the content of this article.

**Ethical approval** Ethical approval for the study was obtained from the Research Ethics Committee (REC) for London – West London and GTAC REC in the United Kingdom, and from the respective ethics committees in each participating European country. This project was conducted in accordance with the ethical standards laid down in the Declaration of Helsinki of 1964 and its later amendments.

**Consent to participate** All individual participants included in the study provided written informed consent for their participation.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.





















## References

1. Crowell V, Houghton R, Tomar A, Fernandes T, Squitieri F (2021) Modeling manifest Huntington's Disease prevalence using diagnosed incidence and survival time. *Neuroepidemiology* 55(5):361–368. <https://doi.org/10.1159/000516767>
2. Bates GP, Dorsey R, Gusella JF, Hayden MR, Kay C, Leavitt BR et al (2015) Huntington disease. *Nat Rev Dis Primers* 1(1):15005. <https://doi.org/10.1038/nrdp.2015.5>
3. Gokhale P, Villa Zapata L (2025) Economic burden of Huntington's disease: a systematic review. *J Huntington Dis*. <https://doi.org/10.1177/18796397251319209>
4. van Lonkhuizen PJC, Frank W, Heemskerk AW, van Duijn E, de Bot ST, Mühlböck A et al (2023) Quality of life, health-related quality of life, and associated factors in Huntington's disease: a systematic review. *J Neurol*. <https://doi.org/10.1007/s00415-022-11551-8>
5. Ellison M (2017) The impact of Huntington disease on young people. *Handb Clin Neurol* 144:179–182. <https://doi.org/10.1016/b978-0-12-801893-4.00015-8>
6. Tabrizi SJ, Estevez-Fraga C, van Roon-Mom WMC, Flower MD, Scahill RI, Wild EJ et al (2022) Potential disease-modifying therapies for Huntington's disease: lessons learned and future opportunities. *Lancet Neurol* 21(7):645–658. [https://doi.org/10.1016/s1474-4422\(22\)00121-1](https://doi.org/10.1016/s1474-4422(22)00121-1)
7. Mwape C, Qureshi AA, Saeed MZ, Fatima A, Jamil A, Mahmood H et al (2026) AMT-130 gene therapy: a promising disease-modifying approach for Huntington's disease. *Ann Med Surg Lond* 88(1):1144–1145. <https://doi.org/10.1097/ms9.00000000000004574>

8. Mühlbäck A, van Walsem M, Nance M, Arnesen A, Page K, Fisher A et al (2023) What we don't need to prove but need to do in multidisciplinary treatment and care in Huntington's disease: a position paper. *Orphanet J Rare Dis* 18(1):19. <https://doi.org/10.1186/s13023-023-02622-8>
9. Churruca K, Pomare C, Ellis LA, Long JC, Henderson SB, Murphy LED et al (2021) Patient-reported outcome measures (PROMs): a review of generic and condition-specific measures and a discussion of trends and issues. *Health Expect* 24(4):1015–1024. <https://doi.org/10.1111/hex.13254>
10. FDA. FDA patient-focused drug development guidance series for enhancing the incorporation of the patient's voice in medical product development and regulatory decision making.: U.S. Food and Drug Administration; 2025 [Available from: <https://www.fda.gov/drugs/development-approval-process-drugs/fda-patient-focused-drug-development-guidance-series-enhancing-incorporation-patients-voice-medical>].
11. Mestre TA, Carlozzi NE, Ho AK, Burgunder J-M, Walker F, Davis AM et al (2018) Quality of life in Huntington's Disease: critique and recommendations for measures assessing patient health-related quality of life and caregiver quality of life. *Mov Disord* 33(5):742–749. <https://doi.org/10.1002/mds.27317>
12. Isaacs D, Gibson JS, Stovall J, Claassen DO (2020) The impact of anosognosia on clinical and patient-reported assessments of psychiatric symptoms in Huntington's Disease. *J Huntingtons Dis* 9(3):291–302. <https://doi.org/10.3233/jhd-200410>
13. Glidden AM, Luebke EA, Elson MJ, Goldenthal SB, Snyder CW, Zizzi CE et al (2020) Patient-reported impact of symptoms in Huntington disease: PRISM-HD. *Neurology* 94(19):e2045–e2053. <https://doi.org/10.1212/wnl.00000000000008906>
14. McKenna SP, Wilburn J (2018) Patient value: its nature, measurement, and role in real world evidence studies and outcomes-based reimbursement. *J Med Econ* 21(5):474–480. <https://doi.org/10.1080/13696998.2018.1450260>
15. Karimi M, Brazier J (2016) Health, health-related quality of life, and quality of life: what is the difference? *Pharmacoeconomics* 34(7):645–649. <https://doi.org/10.1007/s40273-016-0389-9>
16. McCabe MP, Firth L, O'Connor E (2009) A comparison of mood and quality of life among people with progressive neurological illnesses and their caregivers. *J Clin Psychol Med Settings* 16(4):355–362. <https://doi.org/10.1007/s10880-009-9168-5>
17. Hocaoglu MB, Gaffan EA, Ho AK (2012) Health-related quality of life in Huntington's disease patients: a comparison of proxy assessment and patient self-rating using the disease-specific Huntington's disease health-related quality of life questionnaire (HDQoL). *J Neurol* 259(9):1793–1800. <https://doi.org/10.1007/s00415-011-6405-2>
18. Clay E, De Nicola A, Dorey J, Squitieri F, Aballéa S, Martino T et al (2012) Validation of the first quality-of-life measurement for patients with Huntington's disease: the Huntington Quality of Life Instrument. *Int Clin Psychopharmacol* 27(4):208–214. <https://doi.org/10.1097/YIC.0b013e3283534fa9>
19. Carlozzi NE, Schilling SG, Lai JS, Paulsen JS, Hahn EA, Perlmutter JS et al (2016) HDQLIFE: development and assessment of health-related quality of life in Huntington disease (HD). *Qual Life Res* 25(10):2441–2455. <https://doi.org/10.1007/s11136-016-1386-3>
20. Read J, Jones R, Owen G, Leavitt BR, Coleman A, Roos RA et al (2013) Quality of life in Huntington's disease: a comparative study investigating the impact for those with pre-manifest and early manifest disease, and their partners. *J Huntington's Dis* 2(2):159–175. <https://doi.org/10.3233/jhd-130051>
21. Al Sayah F, Jin X, Johnson JA (2021) Selection of patient-reported outcome measures (PROMs) for use in health systems. *J Patient Rep Outcomes* 5(Suppl 2):99. <https://doi.org/10.1186/s41687-021-00374-2>
22. Crossnohere NL, Brundage M, Calvert MJ, King M, Reeve BB, Thorner E et al (2021) International guidance on the selection of patient-reported outcome measures in clinical trials: a review. *Qual Life Res* 30(1):21–40. <https://doi.org/10.1007/s11136-020-02625-z>
23. Gill TM, Feinstein AR (1994) A critical appraisal of the quality of quality-of-life measurements. *JAMA* 272(8):619–626
24. WHO (1995) The World Health Organization Quality of Life assessment (WHOQOL): position paper from the World Health Organization. *Soc Sci Med* 41(10):1403–1409. [https://doi.org/10.1016/0277-9536\(95\)00112-k](https://doi.org/10.1016/0277-9536(95)00112-k)
25. Hunt SM, McKenna SP (1992) The QLDS: a scale for the measurement of quality of life in depression. *Health Policy* 22(3):307–319. [https://doi.org/10.1016/0168-8510\(92\)90004-u](https://doi.org/10.1016/0168-8510(92)90004-u)
26. Doward LC, McKenna SP, Whalley D, Tennant A, Griffiths B, Emery P et al (2009) The development of the L-QoL: a quality-of-life instrument specific to Systemic Lupus Erythematosus. *Ann Rheum Dis* 68(2):196–200. <https://doi.org/10.1136/ard.2007.086009>
27. Heaney A, Wilburn J, Rouse M, Langmead S, Blakeley JO, Huson S et al (2020) The development of the PlexiQoL: a patient-reported outcome measure for adults with Neurofibromatosis Type 1-associated plexiform neurofibromas. *Mol Genet Genomic Med* 8(12):e1530. <https://doi.org/10.1002/mgg3.1530>
28. Wilburn J, McKenna SP, Heaney A, Rouse M, Taylor M, Culkin A et al (2018) Development and validation of the Parenteral Nutrition Impact Questionnaire (PNIQ), a patient-centric outcome measure for home parenteral nutrition. *Clin Nutr* 37(3):978–983. <https://doi.org/10.1016/j.clnu.2017.04.004>
29. McKenna SP, Heaney A, Langley PC (2021) Fundamental outcome measurement: selecting patient reported outcome instruments and interpreting the data they produce. *Innov Pharm.* <https://doi.org/10.24926/ipp.v1i2i2.3911>
30. McKenna SP, Heaney A. The contribution of need fulfilment to quality of life: A reflection on the relation between the needs-based model of quality of life and Max Neef's theory of Human Scale Development. *Archives of Psychology*. 2020;4(2).
31. Jones D, Lal S, French C, Sowerbutts AM, Gittins M, Gabe S et al (2023) Investigating the relationship between home parenteral support and needs-based quality of life in patients with chronic intestinal failure: a national multi-centre longitudinal cohort study. *Nutrients* 15(3):622
32. Rozjabeck H, Li N, Hartmann H, Fu DJ, Canuso C, Jamieson C (2022) Assessing the meaningful change threshold of Quality of Life in Depression Scale using data from two phase 3 studies of Esketamine nasal spray. *J Patient Report Outcomes* 6(1):74. <https://doi.org/10.1186/s41687-022-00453-y>
33. Strand V, de Vlam K, Covarrubias-Cobos JA, Mease PJ, Gladman DD, Graham D et al (2019) Tofacitinib or Adalimumab versus placebo: patient-reported outcomes from OPAL Broaden-a phase III study of active psoriatic arthritis in patients with an inadequate response to conventional synthetic disease-modifying antirheumatic drugs. *RMD Open* 5(1):e000806. <https://doi.org/10.1136/rmdopen-2018-000806>
34. HSG (1996) Unified Huntington's disease rating scale: reliability and consistency. *Mov Disord* 11:136–142. <https://doi.org/10.1002/mds.870110204>
35. Braun V, Clarke V (2006) Using thematic analysis in psychology. *Qual Res Psychol* 3(2):77–101. <https://doi.org/10.1191/1478088706qp0630a>
36. Spray I, Johnstone E, Thorpe J, Frank W, Squitieri F, Giancaterino G, et al. F104 Quality of life in Huntington's disease: a qualitative exploration using the needs-based model. *J Neurol Neurosurg Psychiatry*. 2024;95(1):115. <https://doi.org/10.1136/jnnp-2024-EHDN.222>.

37. Hunt SM, Alonso J, Bucquet D, Niero M, Wiklund I, McKenna S (1991) Cross-cultural adaptation of health measures. *Health Policy* 19(1):33–44. [https://doi.org/10.1016/0168-8510\(91\)90072-6](https://doi.org/10.1016/0168-8510(91)90072-6)
38. Hughes K. Comparing Pretesting Methods: Cognitive Interviews, Respondent Debriefing, and Behavior Coding. Annual Meeting of the Federal Committee on Statistical Methodology November 17–19, 2003 Arlington, VA September 2004: Statistical Research Division U.S. Bureau of the Census Washington D.C. ; 2004. p. 2, 1–20.
39. Scott K, Ummer O, LeFevre AE (2021) The devil is in the detail: reflections on the value and application of cognitive interviewing to strengthen quantitative surveys in global health. *Health Policy Plan* 36(6):982–995. <https://doi.org/10.1093/heapol/czab048>
40. Hunt SM, McKenna SP, McEwen J, Williams J, Papp E (1981) The Nottingham Health Profile: subjective health status and medical consultations. *Soc Sci Med A* 15(3 Pt 1):221–229. [https://doi.org/10.1016/0271-7123\(81\)90005-5](https://doi.org/10.1016/0271-7123(81)90005-5)
41. Long JD, Paulsen JS, Marder K, Zhang Y, Kim J-I, Mills JA et al (2014) Tracking motor impairments in the progression of Huntington's disease. *Mov Disord* 29(3):311–319. <https://doi.org/10.1002/mds.25657>
42. Shoulson I, Fahn S (1979) Huntington disease: clinical care and evaluation. *Neurology* 29(1):1–3. <https://doi.org/10.1212/wnl.29.1.1>
43. Warner JH, Long JD, Mills JA, Langbehn DR, Ware J, Mohan A et al (2022) Standardizing the CAP Score in Huntington's Disease by predicting age-at-onset. *J Huntingtons Dis* 11(2):153–171. <https://doi.org/10.3233/jhd-210475>
44. Langbehn DR, Fine EM, Meier A, Hersch S (2022) Utility of the Huntington's Disease Prognostic Index Score for a perimanifest clinical trial. *Mov Disord* 37(5):1040–1046. <https://doi.org/10.1002/mds.28944>
45. Rasch G. Studies in mathematical psychology: I. Probabilistic models for some intelligence and attainment tests. Oxford, England: Nielsen & Lydiche; 1960. xiii, 184–xiii, p.
46. Andrich D (1988) Rasch models for measurement. Sage Publications, Inc, Thousand Oaks, CA, US, p 95
47. Wright BS, M. Best test design 1979.
48. Hobart J, Cano S (2009) Improving the evaluation of therapeutic interventions in multiple sclerosis: the role of new psychometric methods. *Health Technol Assess* 13(12):1–177. <https://doi.org/10.3310/hta13120>
49. Fisher JW (1992) Reliability statistics. *Rasch Meas Trans* 6:238
50. Tennant A, Conaghan PG (2007) The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Rheum* 57(8):1358–1362. <https://doi.org/10.1002/art.23108>
51. Christensen KB, Makransky G, Horton M (2017) Critical values for Yen's Q(3): identification of local dependence in the Rasch model using residual correlations. *Appl Psychol Meas* 41(3):178–194. <https://doi.org/10.1177/0146621616677520>
52. Smith EV Jr. (2002) Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *J Appl Meas* 3(2):205–231
53. Holland PW, Wainer H (2012) Differential item functioning. Routledge
54. Nunnally JC, Bernstein IH (1994) Psychometric theory. McGraw-Hill Companies, Incorporated
55. Weiner EA, Stewart BJ (1984) Assessing individuals: psychological and educational tests and measurements. Little, Brown
56. Netemeyer R, Bearden W, Sharma S. Scaling Procedures. Thousand Oaks, California: SAGE Publications, Inc.; 2003. Available from: <https://methods.sagepub.com/book/mono/scaling-procedures/toc>.
57. Cruickshank T, Reyes A, Peñailillo L, Ziman M, Thompson J (2014) Factors that contribute to balance and mobility impairments in individuals with Huntington's Disease. *Basal Ganglia*. <https://doi.org/10.1016/j.baga.2014.04.002>
58. Chisholm LZ, Flavin KT, Paulsen JS, Ready R (2013) Psychological well-being in persons affected by Huntington's disease: a comparison of at-risk, prodromal, and symptomatic groups. *J Health Psychol* 18(3):408–418. <https://doi.org/10.1177/1359105312444646>
59. Tabrizi SJ, Reilmann R, Roos RAC, Durr A, Leavitt B, Owen G et al (2012) Potential endpoints for clinical trials in premanifest and early Huntington's disease in the TRACK-HD study: analysis of 24 month observational data. *Lancet Neurol* 11(1):42–53. [https://doi.org/10.1016/S1474-4422\(11\)70263-0](https://doi.org/10.1016/S1474-4422(11)70263-0)
60. Weemering DN, Beelen A, Kliet T, van Leeuwen LAG, van den Berg LH, van Eijk RPA (2024) Trial participation in neurodegenerative diseases: barriers and facilitators. *Neurology* 103(1):e209503. <https://doi.org/10.1212/WNL.0000000000209503>
61. Snyder C, Crossnohere N, King M, Reeve BB, Bottomley A, Calvert M et al (2022) The PROTEUS-Trials Consortium: optimizing the use of patient-reported outcomes in clinical trials. *Clin Trials* 19(3):277–284. <https://doi.org/10.1177/17407745221077691>
62. Jones A, Morgan-Jones P, Busse M, Shepherd V, Wood F (2021) Conducting focus groups in neurodegenerative disease populations: ethical and methodological considerations. *BMJ Open* 11(1):e041869. <https://doi.org/10.1136/bmjopen-2020-041869>
63. Crocker TF, Smith JK, Skevington SM (2015) Family and professionals underestimate quality of life across diverse cultures and health conditions: systematic review. *J Clin Epidemiol* 68(5):584–595. <https://doi.org/10.1016/j.jclinepi.2014.12.007>
64. Hutchinson C, Khadka J, Crocker M, Lay K, Milte R, Whitehurst DG et al (2024) Examining interrater agreement between self-report and proxy-report responses for the quality of life-aged care consumers (QOL-ACC) instrument. *J Patient Rep Outcomes* 8(1):28. <https://doi.org/10.1186/s41687-024-00705-z>
65. Grzeda MT, Thorpe J, Johnstone E, Spray I, Hannemann W, Squitieri F et al (2025) PCR155 Measuring Quality of Life in Huntington's Disease Using a Proxy Approach: Validation of the HD-mQoL-prx Measure. *Value Health* 28(12):S578. <https://doi.org/10.1016/j.jval.2025.09.2628>
66. Tabrizi SJ, Schobel S, Gantman EC, Mansbach A, Borowsky B, Konstantinova P et al (2022) A biological classification of Huntington's disease: the Integrated Staging System. *Lancet Neurol* 21(7):632–644. [https://doi.org/10.1016/s1474-4422\(22\)00120-x](https://doi.org/10.1016/s1474-4422(22)00120-x)

## Authors and Affiliations

Isobel Spray<sup>1</sup>  · Mariusz T. Grzeda<sup>1</sup>  · Jeanette Thorpe<sup>1</sup>  · Ellie Johnstone<sup>1</sup>  · Wiebke Hannemann<sup>2</sup>  ·  
 Ferdinando Squitieri<sup>3</sup>  · Giulia Giancaterino<sup>4</sup>  · Paola Emilia Mina<sup>4</sup>  · Barbara D'Alessio<sup>4</sup>  · Astri Arnesen<sup>5</sup>  ·  
 Jiří Klempíř<sup>6</sup>  · Olga Ulmanová<sup>6</sup>  · Pearl van Lonkhuizen<sup>7,8,9</sup>  · Anne-Wil Heemskerk<sup>7,9</sup>  · Jennifer Hoblyn<sup>10,11</sup>  ·  
 Emer O'Malley<sup>10</sup>  · Ramona Moldovan<sup>12,13,14</sup>  · Georg Bernhard Landwehrmeyer<sup>2</sup>  · Stephen P. McKenna<sup>1</sup>  ·  
 Alzbeta Mühlbäck<sup>2,6,15</sup>  on behalf of the HEALTHE-RND consortium

- ✉ Alzbeta Mühlbäck  
alzbeta.muehlbaeck@uni-ulm.de
- 1 Galen Research, Manchester, UK
  - 2 Department of Neurology, University Hospital Ulm, University Ulm, Ulm, Germany
  - 3 Unità Huntington E Malattie Rare Fondazione IRCCS Casa Sollievo Della Sofferenza Hospital, San Giovanni Rotondo, Italy
  - 4 Qrare Medical Centre, Fondazione Lega Italiana Ricerca Huntington (LIRH), Rome, Italy
  - 5 European Huntington Association, Kristiansand, Norway
  - 6 Department of Neurology and Centre of Clinical Neuroscience, General University Hospital and First Faculty of Medicine, Charles University, Prague, Czech Republic
  - 7 Department of Public Health and Primary Care, Leiden University Medical Center, Leiden, Netherlands
  - 8 National eHealth Living Lab, Leiden University Medical Center, Leiden, Netherlands
  - 9 Huntington Center Topaz Overduin, Katwijk, Netherlands
  - 10 Trinity College Dublin, Dublin, Ireland
  - 11 John of God University Hospital, Dublin, Ireland
  - 12 Division of Evolution, Infection and Genomics, School of Biological Sciences, University of Manchester, Manchester, UK
  - 13 Manchester Centre for Genomic Medicine, St. Mary's Hospital, Manchester University NHS Foundation Trust, Manchester, UK
  - 14 Department of Psychology, Babeş-Bolyai University, Cluj-Napoca-Napoca, Romania
  - 15 Huntington-Zentrum Süd, kbo Isar-Amper-Klinikum, Taufkirchen, Germany