Validity, reliability, ability to detect change and meaningful within-patient change of the cUHDRS

Dylan Trundell (1)*, Giuseppe Palermo (2), Scott Schobel (3), Jeffery D. Long (4,5), Blair R. Leavitt (6), Sarah J. Tabrizi (7)

(1) Roche Products Ltd., Welwyn Garden City, UK; (2) Roche Pharma Research and Early Development, Roche Innovation Center Basel, Basel, Switzerland; (3) F. Hoffman-La Roche Ltd., Basel, Switzerland; (4) Department of Psychiatry, Carver College of Medicine, University of Iowa, Iowa City, USA; (5) Department of Biostatistics, College of Public Health, University of Iowa, Iowa City, USA; (6) Centre for Molecular Medicine and Therapeutics, Department of Medical Genetics, University of British Columbia, Vancouver, British Columbia, Canada; (7) Huntington's Disease Research Centre, UCL Institute of Neurology, London, UK

Summary

HD has a range of symptom domains that affect day-to-day life, including movement, cognitive function, behaviour and functional capacity. Existing assessment scales used in HD clinical studies typically measure one symptom domain. The cUHDRS scoring system combines existing measurement scales to holistically measure patient experience. Using registry data, this study found that using cUHDRS in clinical studies is a reliable way to measure meaningful changes in disease progression and the potential effects of investigational drugs.

Methods

Study Design/Population

- To assess the reliability, validity, and ability to detect change of the cUHDRS, data from ENROLL-HD ('trial-like': manifest HD, 25–65 years, IS ≥70, CAP ≥400) were used (except 'association with function', where all manifest HD patients were included).
- To estimate meaningful within-patient cUHDRS score change, data from ENROLL-HD (population as above) and data from REGISTRY (manifest HD, 25–65 years, TFC 5–13, CAP ≥400) were used.

Measures

- The cUHDRS is scored using the following formula:

$$cUHDRS = \frac{TFC - 8.8}{2.8} - \frac{TMS - 34.4}{17.4} + \frac{SDMT - 25.2}{12.4} + \frac{SWR - 58.0}{21.2} + 10$$

- The population reference means and standard deviations were calculated using baseline ENROLL-HD data for those patients Stage 1–3 (TFC 5–13;) and aged 20+ years.

Table 1: Measures used in present study

Measures							
Function/Independence			*Motor Function*	*Cognitive Function*		*Apathy*	*Global Severity*
TFC	IS	FA	TMS	SDMT	SWR	PBA-s	CGI-S

Analysis

Descriptive statistics

- Baseline descriptive statistics were reported for age and gender, for the 'trial-like' ENROLL-HD population.

Test-retest reliability

- ICCs were calculated, comparing cUHDRS scores at baseline and Month 12, in a subset of patients with no change in IS score at Month 12.
- An ICC ≥0.7 is considered acceptable.[1]

Convergent validity

- Spearman rank order correlation coefficients were calculated between the cUHDRS and the following measures at baseline: FA, IS, PBA-s: Apathy.
- Stronger correlations were expected with FA and IS than PBA-s: Apathy.

Known-groups validity

- ANCOVA was conducted at baseline, controlling for age and gender, for the following groups (≥median score vs <median score): FA, IS.
- Significant differences provide evidence of known-groups validity.

Ability to detect change

- Non-progressors' and progressors' LS mean change scores were compared (from baseline to 24 months) using ANCOVA, controlling for age, gender and baseline score.
- Non-progressors defined as improvement, no change, or <10-point decline (IS) or <3-point decline (FA). Progressors defined as ≥10-point decline (IS) or ≥3-point worsening (FA).
- Significant differences provide evidence of ability to detect change.

Association with Function

- Patients were split into deciles based on cUHDRS score. For each decile, the proportion of patients achieving each item on the FA was reported.

Estimation of Meaningful Within-Patient Change

Anchor based

- Linear regression models were calculated for change from baseline to 12 months with the cUHDRS change score as the dependent variable, and either IS change score (ENROLL-HD) or the CGI-S (REGISTRY) as an independent variable.
- The cUHDRS score change associated with a one category worsening on the CGI-S, and a 10-point worsening on the IS were considered to reflect a meaningful within-patient worsening.

Distribution based

- The Standard Error of Measurement (SD*sqrt[1-reliability]) was calculated for the cUHDRS, using baseline standard deviation (ENROLL-HD) and the test-retest reliability coefficient.

Results

Sociodemographic Descriptive Statistics

- N=2431. Mean age was 49.4 years (SD=9.6), and 51% were male.

Test-Retest Reliability

- ICC=0.92; demonstrating excellent test-retest reliability.

Convergent Validity

- FA: rho=0.75; IS: rho=0.69; PBA-s Apathy: -0.21
- Logical pattern and magnitude was observed, providing supportive evidence of convergent validity.

Known-Groups Validity

- Patients ≥median on FA and IS (i.e. less impacted) had significantly higher cUHDRS scores (i.e. less impacted) than those <median (both analyses: $P<0.001$), providing strong evidence of known-groups validity.

Ability to Detect Change

- Significant differences were identified ($P<0.001$), providing excellent evidence of the ability of the cUHDRS to detect change (Table 2).

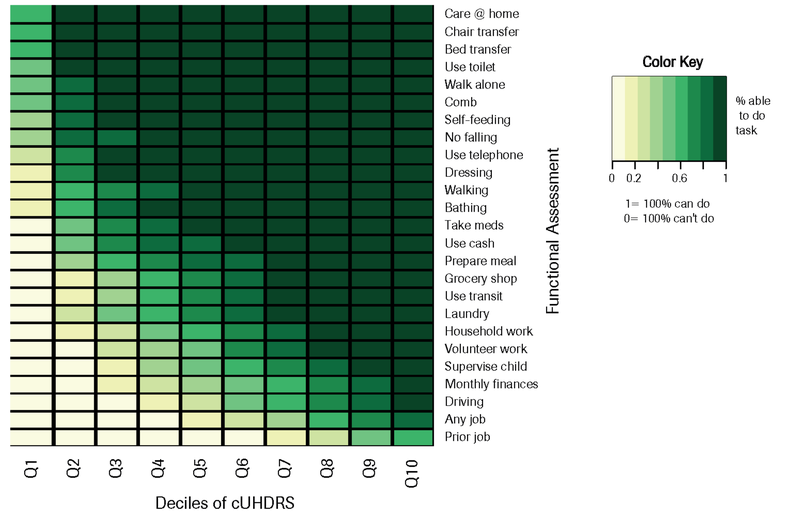
Table 2: Ability of the cUHDRS to detect change (Baseline to 24 months; IS and FA as anchors)

Anchor (progressor)	Progressor		Non-progressor		LS Mean Difference
	N	LS Mean Change	N	LS Mean Change	
IS (≥10-pt decline)	139	-1.76	309	-0.88	0.9 (P<0.001)
FA (≥3-pt decline)	126	-1.44	311	-0.02	1.4 (P<0.001)

Association with function

- **Figure 1** shows that as the cUHDRS score declines (from decile Q10–Q1), the proportion of patients achieving each item on the FA also declines, demonstrating an association between cUHDRS score and functional ability.

Figure 1: Proportion of patients achieving each FA item by cUHDRS decile



Meaningful within-patient change

- Anchor-based estimates were 1.22 (CGI-S) and 1.14 (IS); SEM=0.80.
- Based on these values, a decline of 1.2 points is recommended as an estimate of meaningful within-patient decline for use in a responder analysis endpoint in clinical trials in patients with manifest HD.

Conclusions

- The cUHDRS is valid, reliable and able to detect change in patients matching the intended clinical trial population for Roche's Phase 3 clinical study in manifest HD.
- Analyses anchored against CGI-S and IS support that a decline on the cUHDRS of 1.2 points is clinically meaningful, and that the cUHDRS tracks function decline.

Abbreviations

ANCOVA, analysis of covariance; CAP, CAG age-product; CGI-S, Clinical Global Impression of Severity; cUHDRS, composite Unified Huntington's Disease Rating Scale; FA, Functional Assessment; HD, Huntington's disease; ICC, Intraclass Correlation Coefficients; IS, Independence Scales; LS, least square; PBA-s, Problem Behaviors Assessment-Short form; SD, standard deviation; SDMT, Symbol Digit Modalities Test; SEM, standard error of the mean; SQRT, square root; SWR, Stroop Word Reading; TFC, Total Functional Capacity; TMS, Total Motor Score.

Reference

1. Nunnally JC and Bernstein I. *Psychometric Theory.* 1994; 3:248–292.

Acknowledgments

We thank all the patients who participate in our studies and their families. This study was funded by F. Hoffmann La-Roche. Writing and editorial assistance was provided by Matt Gooding of MediTech Media, UK, in accordance with the Good Publication Practice (GPP3) guidelines (http://www.ismpp.org/gpp3).

Please scan using your QR reader application to access this poster on your mobile device.
NB: there may be associated costs for downloading data. These costs may be high if you are using your smartphone abroad. Please check your mobile data tariff or contact your service provider for more details.
Alternatively this can be accessed at: https://bit.ly/2PusSIF

Presented at HSG 2018: Unlocking HD, Houston, TX, USA; November 8–10, 2018