Evaluation and Validation of the Hypoparathyroidism Daily Diary of Symptom Experience (HPT-DD-SE) and the Hypoparathyroidism Life Impact Questionnaire (HPT-LIQ)

Diane Whalley[1]; Rebecca Crawford[1]; Nadège Tremel[2]; Blandine Weiss[2]; Caroline Robo[2]; Gaëlle Charier[2]; Deborah Murphy[3]; Lynda Doward[1]; Soraya Allas[2]

[1]RTI Health Solutions, Manchester, United Kingdom; [2]Alexion Pharma France, Écully, France; [3]HypoPARAthyroidism Association, Inc., Lemoore, CA, United States

INTRODUCTION

- Hypoparathyroidism (HypoPT) is a rare disease of parathyroid hormone (PTH) deficiency that leads to broad functional consequences with a heavy symptom burden and impaired quality of life[1,2]
- The Hypoparathyroidism Daily Diary of Symptom Experience (HPT-DD-SE) and Hypoparathyroidism Life Impact Questionnaire (HPT-LIQ) are novel patient-reported outcome (PRO) measures assessing symptoms and health-related quality of life (HRQoL) in chronic HypoPT[3]
- The HPT-DD-SE and HPT-LIQ were developed in United States (US) English based on qualitative work conducted with patients with chronic HypoPT in the US and adapted into 18 new languages (16 countries)[3,4]

OBJECTIVES

- This study evaluated the psychometric properties and meaningful change thresholds for domains of the HPT-DD-SE and HPT-LIQ, which are novel PRO measures assessing symptoms and HRQoL in chronic HypoPT

CONCLUSIONS

- HPT-DD-SE and HPT-LIQ scale scores demonstrate strong reliability, construct validity, and responsiveness to assess chronic HypoPT symptoms and impacts in clinical trials
- These data reflect improvements in physical symptoms, cognitive symptoms, and physical functioning that would be meaningful for patients affected by chronic HypoPT

METHODS

Analysis sample

- Psychometric analyses were conducted in 2 stages using data collected in the randomized, placebo-controlled, phase 3 CALYPSO trial of eneboparatide in adults with chronic HypoPT (NCT05778071)[5]
- The stage 1 analysis sample included all randomized patients who had evaluable HPT-DD-SE or HPT-LIQ data at baseline and week 12
- The stage 2 analysis sample included randomized patients who had evaluable HPT-DD-SE or HPT-LIQ data at baseline and week 24

Psychometric evaluation

- HPT-DD-SE scale scores were computed as the weekly average of daily scores, with daily scores computed as the average of item scores; HPT-LIQ scale scores were computed as the average of item scores
- Evaluations included: test-retest reliability intraclass correlation coefficients (ICCs); Cronbach's alpha and McDonald's omega coefficients for internal consistency; construct validity and responsiveness correlations with the Patient Global Impression of Severity (PGIS), Patient Global Impression of Change (PGIC), Clinical Global Impression of Severity (CGIS), and Short Form Health Survey-36 version 2 (SF-36v2) scores, as well as with serum calcium (sCa) and urinary calcium (uCa) levels; convergent/divergent validity correlations; known-groups analysis; responsiveness correlations; change-groups analysis; and descriptive statistics
- Meaningful within-patient change (MWPC) and minimal important difference (MID) thresholds were estimated using descriptive statistics for change in HPT-DD-SE/HPT-LIQ scale scores at defined PGIS anchor levels

RESULTS

- The stage 1 and stage 2 analysis samples included 189 patients and 151 patients, respectively (Table 1)
- There were no problematic floor or ceiling effects in any of the HPT-DD-SE scale scores at baseline; specifically, fewer than 15% of participants obtained either the worst or best score at baseline
- For the HPT-LIQ Physical Functioning scale score, there was no floor effect, but there was a minor ceiling effect at baseline, with 19.8% of participants obtaining the best score of 0
- Scale scores demonstrated good test-retest reliability (ICCs > 0.7) and strong internal consistency (Cronbach's alpha and McDonald's omega > 0.9) (Tables 2 and 3)
- Most correlations with external measures were as hypothesized (Table 4)
- In the HPT-DD-SE known-groups analysis, higher mean scores were observed for subgroups of participants with severe symptoms based on the corresponding PGIS and CGIS items (all analysis of variance [ANOVA] *P*'s < 0.0001)
- In the HPT-DD-SE change-groups analysis, participants with the greatest improvements based on the PGIS, PGIC, and CGIS items showed the greatest improvement in scale scores (all ANOVA *P*'s < 0.01)
- In the HPT-LIQ known-groups analysis, higher mean scores were observed for subgroups of participants with greater impact or more severe symptoms on the corresponding PGIS and CGIS items for HPT-LIQ Physical Functioning (all ANOVA *P*'s < 0.0001)
- In the HPT-LIQ change-groups analysis, participants who had the greatest improvements based on the PGIS and PGIC items showed the greatest improvements in HPT-LIQ Physical Functioning (all ANOVA *P*'s < 0.01)
- The predefined primary PGIS 1-category improvement anchor was used to estimate MWPC thresholds, with the PGIS 1-category and 2-category improvement anchors used to derive an additional MWPC threshold estimate (Table 5)

Table 1. Sample characteristics

	Stage 1 (N = 189)	Stage 2 (N = 151)
Age, mean (SD)	51.3 (12.38)	51.5 (12.05)
Sex, n (%)		
Female	145 (76.7)	120 (79.5)
Male	44 (23.3)	31 (20.5)
Race, n (%)		
Asian	2 (2.4)	1 (1.4)
Black or African American	1 (1.2)	1 (1.4)
White	81 (95.3)	70 (95.9)
Unknown	1 (1.2)	1 (1.4)
Missing	104 (55)	78 (51.7)
Ethnicity, n (%)		
Hispanic or Latino	8 (9.4)	6 (8.2)
Not Hispanic or Latino	74 (87.1)	64 (87.7)
Not reported	2 (2.4)	2 (2.7)
Unknown	1 (1.2)	1 (1.4)
Missing	104 (55.0)	78 (51.7)

SD, standard deviation.

Table 2. Test-retest reliability

	ICC (95% CI), n (participants with no change in PGIS*)		
	Screening to BL†	W8 to W12‡	W20 to W24‡
HPT-DD-SE Core Physical Symptoms	0.84 (0.78, 0.89), 104	0.94 (0.92, 0.96), 102	0.95 (0.92, 0.97), 83
HPT-DD-SE Overall Physical Symptoms	0.86 (0.80, 0.90), 104	0.96 (0.94, 0.97), 102	0.96 (0.94, 0.97), 83
HPT-DD-SE Cognitive Symptoms	0.79 (0.67, 0.87), 96	0.96 (0.94, 0.97), 101	0.96 (0.94, 0.98), 87
HPT-LIQ Physical Functioning	0.89 (0.84, 0.92), 104	0.91 (0.87, 0.94), 107	0.82 (0.74, 0.88), 96

*PGIS Physical Symptoms for HPT-DD-SE Core Physical Symptoms and HPT-DD-SE Overall Physical Symptoms; PGIS Cognitive Symptoms for HPT-DD-SE Cognitive Symptoms; PGIS Physical Functioning for HPT-LIQ Physical Functioning. †For HPT-DD-SE, the analysis was based on the single assessment day at screening and a single day (day 1) at BL. ‡For HPT-DD-SE, the analysis was based on a randomly selected day from the 7 days at each time point. BL, baseline; CI, confidence interval; W8, week 8; W12, week 12; W20, week 20; W24, week 24.

Table 3. Cronbach's alpha and McDonald's omega coefficients

	Cronbach's alpha, n			McDonald's omega, n		
	BL	W12	W24	BL	W12	W24
HPT-DD-SE Core Physical Symptoms	0.91, 189	0.90, 189	0.92, 150	0.92, 189	0.90, 189	0.92, 150
HPT-DD-SE Overall Physical Symptoms	0.95, 189	0.95, 189	0.95, 150	0.95, 189	0.95, 189	0.95, 150
HPT-DD-SE Cognitive Symptoms	0.97, 189	0.96, 189	0.98, 150	0.97, 189	0.96, 189	0.98, 150
HPT-LIQ Physical Functioning	0.89, 187	0.91, 187	0.92, 144	0.90, 187	0.92, 187	0.92, 144

BL, baseline; n, number of participants; W12, week 12; W24, week 24.

Table 4. Correlations with external supporting measures

	Correlation coefficient											
	HPT-DD-SE Core Physical Symptoms			HPT-DD-SE Overall Physical Symptoms			HPT-DD-SE Cognitive Symptoms			HPT-LIQ Physical Functioning		
	BL	W12	W24	BL	W12	W24	BL	W12	W24	BL	W12	W24
PGIS*	0.70	0.77	0.69	0.70	0.77	0.70	0.80	0.85	0.88	0.79	0.78	0.75
PGIS change*	-	0.45	0.51	-	0.49	0.52	-	0.51	0.67	-	0.49	0.46
PGIC†	-	0.59	0.53	-	0.62	0.58	-	0.43	0.52	-	0.35	0.31
CGIS‡	0.48	0.54	0.58	0.47	0.52	0.56	0.50	0.62	0.65	0.34	0.35	0.45
CGIS change‡	-	0.32	0.34	-	0.33	0.36	-	0.34	0.32	-	0.20	0.20
SF-36v2 PCS	−0.71	−0.65	−0.66	−0.71	−0.68	−0.67	−0.59	−0.49	−0.41	−0.78	−0.73	−0.71
SF-36v2 PCS change	-	−0.48	−0.61	-	−0.52	−0.60	-	−0.41	−0.42	-	−0.58	−0.53
SF-36v2 MCS	−0.44	−0.51	−0.32	−0.49	−0.51	−0.36	−0.51	−0.58	−0.55	−0.41	−0.40	−0.31
SF-36v2 MCS change	-	−0.29	−0.43	-	−0.33	−0.49	-	−0.38	−0.45	-	−0.28	−0.21
Albumin-adjusted sCa	0.26	0.09	−0.02	0.24	0.08	−0.01	0.22	0.07	0.04	0.19	0.05	−0.01
uCa	−0.02	0.04	0.01	−0.01	0.05	0.04	−0.05	0.03	0.17	−0.01	0.02	−0.04

*Correlations are PGIS Physical Symptoms with both HPT-DD-SE Core Physical Symptoms and HPT-DD-SE Overall Physical Symptoms, PGIS Cognitive Symptoms with HPT-DD-SE Cognitive Symptoms, and PGIS Physical Functioning with HPT-LIQ Physical Functioning. †Correlations are PGIC Physical Symptoms with both HPT-DD-SE Core Physical Symptoms and HPT-DD-SE Overall Physical Symptoms, PGIC Cognitive Symptoms with HPT-DD-SE Cognitive Symptoms, and PGIC Physical Functioning with HPT-LIQ Physical Functioning. ‡Correlations are CGIS Physical Symptoms with both HPT-DD-SE Core Physical Symptoms and HPT-DD-SE Overall Physical Symptoms, CGIS Cognitive Symptoms with HPT-DD-SE Cognitive Symptoms, and CGIS Overall Symptoms with HPT-LIQ Physical Functioning. BL, baseline; MCS, Mental Component Summary; PCS, Physical Component Summary; W12, week 12; W24, week 24.

Table 5. Meaningful change threshold estimates

	MWPC (primary)	MWPC (supportive)	MID
HPT-DD-SE Core Physical Symptoms	−1.2	−1.9	0.8
HPT-DD-SE Overall Physical Symptoms	−1.1	−1.8	0.8
HPT-DD-SE Cognitive Symptoms	−1.4	−2.9	0.7
HPT-LIQ Physical Functioning	−2.1	−3.4	2.0

For each scale, MWPC was estimated from the mean change score from baseline to week 24 for participants with a PGIS 1-category improvement (the predefined primary anchor). Additionally, to reflect the potential underestimation of meaningful change associated with a PGIS 1-category improvement and overestimation associated with a PGIS 2-category improvement, the midpoint between the mean change scores for these anchor levels was used to derive a second MWPC threshold estimate. MID estimates were computed as the difference between PGIS 1-category improvement and no change.

Acknowledgements

The authors express their gratitude to the patients who contributed to the research. Medical writing support was provided by Sarah Gonzalez, PhD, of Arbor Scientia Group (Carlsbad, CA, USA) and funded by Alexion, AstraZeneca Rare Disease (Boston, MA, USA). Publication management was provided by Simone E. Auteri, MSc, EMS, PhD, of Alexion, AstraZeneca Rare Disease (Barcelona, Spain).

Disclosures

This work was funded by Alexion, AstraZeneca Rare Disease, Boston, MA, USA. **NT**, **BW**, **CR**, **GC**, and **SA** are employees of Alexion Pharma France, Écully, France, and some held stock options. **DW**, **RC**, and **LD** are employees of RTI Health Solutions, Manchester, UK.

References

1. Gafni RI, et al. *N Engl J Med*. 2019;380(18):1738-1747.
2. Mannstadt M, et al. *Nat Rev Dis Primers*. 2017;3:17055.
3. Crawford R, et al. *J Bone Miner Res*. 2023;38(S1):S184.
4. Morrison R, et al. *Value Health*. 2022;25(12):S418.
5. ClinicalTrials.gov. Updated September 8, 2025. Accessed September 15, 2025. https://clinicaltrials.gov/study/NCT05778071

Poster

Poster presented at ISPOR Europe 2025; 9-12 November 2025; Glasgow, Scotland, UK

Corresponding author email address: soraya.allas@alexion.com