

Psychometric Analysis of the Patient-Reported Hypoparathyroidism Symptom Diary Symptom Subscale Using Data from Two Clinical Trials

Lauren Nelson¹, Steven W Ing², Mishaela R Rubin³, Jia Ma¹, Susan Martin⁴, Rohini Sen⁵, Olulade Ayodele⁶

¹RTI Health Solutions, Research Triangle Park, NC, USA; ²Department of Endocrinology, Diabetes and Metabolism, Ohio State University Wexner Medical Center, Columbus, OH, USA; ³Vagelos College of Physicians and Surgeons, Columbia University, New York, NY, USA; ⁴RTI Health Solutions, Ann Arbor, MI, USA; ⁵Takeda Pharmaceuticals U.S.A., Inc., Lexington, MA, USA; ⁶Takeda Development Center Americas Inc., Lexington, MA, USA

Correspondence: Lauren Nelson, RTI Health Solutions, 3040 East Cornwallis Road, PO Box 12194, Research Triangle Park, NC, 27709-2194, USA, Tel +1 919-541-6000 ext. 1541, Email lnelson@rti.org

Purpose: The hypoparathyroidism symptom diary (HypoPT-SD) is a disease-specific patient-reported outcome (PRO) tool comprising a 7-item symptom subscale, a 4-item impact subscale and 1-item anxiety, and sadness or depression components. This analysis assessed the psychometric properties of the HypoPT-SD symptom subscale scores using data from two open-label, single arm, Phase 4 studies (Study 402 and Study 404).

Patients and Methods: Eligible patients were aged 18 years or older with a confirmed diagnosis of hypoparathyroidism. All patients received recombinant human parathyroid hormone (1–84) during the analysis period. Scores were recorded at baseline, and at months 6, 30 and 36 (end of treatment [EOT]) in Study 402, and at baseline and week 52 (EOT) in Study 404. The structure of the HypoPT-SD Symptom subscale was analyzed by measuring correlations between pairs of item scores; internal consistency and reliability were evaluated using Cronbach's coefficient α ; test-retest reliability was assessed using intraclass correlation; and construct validity was determined by performing correlational analyses between scores recorded using the HypoPT-SD and those for other conceptually similar PRO tools.

Results: A total of 60 patients were included in the analysis. Inter-item pairwise correlations were strong for all but 5 of the item pairs analysed. Cronbach's α values for the HypoPT-SD Symptom subscale were 0.88 using data from Study 402 and 0.92 using data from Study 404. In general, the HypoPT-SD Symptom subscale scores had moderate or strong correlations with scores recorded using PRO tools. Intraclass correlation coefficients exceeded 0.70 using test-retest data from all patients in Study 402 and from a subgroup of patients with stable disease from Study 404.

Conclusion: This analysis demonstrated the test-retest reliability, internal consistency and construct validity of the HypoPT-SD using data from longitudinal prospective studies and supports the use of the HypoPT-SD in future clinical studies.

Keywords: hypoparathyroidism, quality of life, patient-reported outcomes, psychometrics

Introduction

Chronic hypoparathyroidism (HypoPT) is a rare endocrine disorder in which the underproduction or total absence of parathyroid hormone leads to hypocalcemia.^{1,2} Patients with HypoPT can experience a wide range of symptoms, many of which are linked to low calcium levels, such as paresthesia, muscle cramps, and spasms.^{1,2} The treatment aims in HypoPT are to maintain calcium levels in the lower part of or slightly below the lower limit of the reference range and, in doing so, to relieve symptoms; conventional therapy, therefore, consists of oral calcium and vitamin D supplementation.³ The effects of HypoPT, however, extend beyond mineral homeostasis to the cardiovascular and renal systems, cognitive function, and psychological health of patients with this condition.^{4–7} Evidence suggests that supplementation alone does not control HypoPT adequately in many patients;⁸ a survey of patients with HypoPT treated with conventional therapy

found that 65% of responders rated their overall symptom severity as moderate or severe, over a third (34%) rated their physical fatigue as severe or very severe, and 19% reported severe or very severe slow/confused thinking.⁹ The same study also showed that increased symptom severity was associated with marked increases in the strain placed on caregivers.⁹ In addition, an online survey of patients with HypoPT in Germany, of whom the vast majority were being treated with conventional therapy, found that high symptom burden was the strongest predictor of clinically impaired functioning, across all domains measured.¹⁰

Symptom burden, along with many other disease-related factors, causes a profound negative impact on health-related quality of life (HRQoL) in patients with HypoPT. Findings from several studies suggest that despite treatment with conventional therapy, HRQoL is lower in patients with HypoPT than in the general population or in matched control populations who do not have HypoPT.^{9,11–13} a cross-sectional study in Norway found that Short-form 36 (SF-36) questionnaire scores (for all domains) and total Hospital Anxiety and Depression Scale (HADS) scores for patients with HypoPT were significantly worse than national average scores;¹¹ in a study carried out in Denmark, patients with HypoPT and hyperthyroidism had significantly lower SF-36 scores than matched healthy controls in all but one (“role emotional”) domain, and significantly lower “role physical” scores than matched patients with hypothyroidism who did not have HypoPT;¹⁴ and a US-based study showed that prior to treatment with recombinant parathyroid hormone 1–84 (rhPTH[1–84]), RAND 36-Item Short Form Health Survey (RAND-36) scores for patients with HypoPT were significantly lower (worse) than those recorded in the normal population across all domains.¹⁵

These studies, however, used generic questionnaires, which may be inadequate for determining the full impact of HypoPT on HRQoL and quantifying the burden of symptoms. There is an unmet need for a validated, disease-specific patient-reported outcomes (PROs) tool that could measure HRQoL over time and be used in clinical trials to assess the impact of treatments on the burden of HypoPT. To date, few such tools have been developed to address this unmet need.^{16–18} We have developed and tested a HypoPT-specific PRO tool, the HypoPT-Symptom Diary (SD), that conforms to the US Food and Drug Administration industry guidance.^{19–21} Previous analyses from a cross-sectional observational study indicated that our tool is “an appropriate measure of HypoPT-related symptoms.”²¹ However, the psychometric properties of the HypoPT-SD have not been assessed in longitudinal studies.

This study aimed to evaluate the test–retest reliability and construct validity of the HypoPT-SD symptom subscale using data from two prospective clinical trials.

Materials and Methods

PRO Tool Design

The HypoPT-SD was originally developed as an English language, daily diary with a 24-hour recall period and consists of a 7-item Symptom subscale, a 4-item Impact subscale, a single-item anxiety score, and a single-item Sadness or Depression score (Table 1). Full details of the HypoPT-SD have been published previously.¹⁹

Table 1 Hypoparathyroidism Symptom Diary Response Scales^a

Item	Response Scale	
	Study 402 (24-h Recall Period)	Study 404 (7-Day Recall Period)
I. How would you rate any muscle cramps you experienced?	0–10 numeric rating scale: 0 = No muscle cramps 10 = Worst possible muscle cramps This item was rescaled to 5 points for use in the subscale computation	5-point ordinal response scale: 0 = None 1 = Mild 2 = Moderate 3 = Severe 4 = Very severe

(Continued)

Table 1 (Continued).

Item	Response Scale	
	Study 402 (24-h Recall Period)	Study 404 (7-Day Recall Period)
2. How would you rate any tingling you experienced?	5-point ordinal response: 0 = None 1 = Mild 2 = Moderate 3 = Severe 4 = Very severe	5-point ordinal response: 0 = None 1 = Mild 2 = Moderate 3 = Severe 4 = Very severe
3. How would you rate any numbness you experienced?		
4. How would you rate any muscle spasms or twitching you experienced?		
5. How would you rate any feelings of heaviness you experienced in your arms or legs?		
6. How would you rate any physical fatigue you experienced?		
7. How would you rate any slowed or confused thinking, sometimes called brain fog, you experienced?		
8. How would you rate any anxiety you felt?		
9. How would you rate any sadness or depression you felt?		
10. How did your hypoparathyroidism impact your sleep last night?	3-point ordinal response: 0 = Not at all 1 = Somewhat 2 = Very much	3-point ordinal response: 0 = Not at all 1 = Somewhat 2 = Very much
11. How did your hypoparathyroidism impact your ability to exercise?		
12. How did your hypoparathyroidism impact your ability to complete your work (for example, at school, at home, or at a job)?		
13. How did your hypoparathyroidism impact your relationships with your family members?		

Notes: *The Hypoparathyroidism Symptom Diary Response Scales in Table 1 are reprinted with permission of Shire Human Genetic Therapies, Inc. © 2016 Shire Human Genetic Therapies, Inc. All rights reserved.

Data Sources

Data used in the current psychometric evaluation were collected from two clinical studies:

SHP634-402 (Study 402), a 36-month, single-center, open-label, single-arm, phase 4 study (NCT02910466) designed to evaluate albumin-corrected serum calcium levels in adults with HypoPT who were receiving recombinant rhPTH(1–84),²² and SHP634-404 (Study 404), a 52-week, global, multicenter, open-label, phase 4 study (NCT03364738) designed to evaluate the safety and efficacy of once-daily subcutaneous rhPTH(1–84) in patients with HypoPT who had previously been exposed to rhPTH(1–84) in the PARALLAX (NCT02781844) study.²³

The protocols for both studies involved, all protocol amendments, the final approved informed consent document, relevant supporting information, and all types of subject recruitment information were submitted by the investigators to the following institutional review boards or independent ethics committees. Study 402: Columbia Research Human Research Protection Office Institutional Review Board; Study 404: Indiana University Review Board, Western Institutional Review Board, Advarra Institutional Review Board, Schulman Institutional Review Board, University of Kentucky Institutional Review Board, State University of New York Upstate Medical University Institutional Review Board, Comité d'éthique de la recherche du CHU de Québec - Université Laval, Videnskabssetiske Komitéer for Region Midtjylland, Egészségügyi Tudományos Tanács - Klinikai Farmakológiai Etikai.

Both studies from which data were used were conducted in accordance with current applicable regulations, International Conference on Harmonisation, European Union Directive 2001/20/EC and its updates, and local ethical and legal requirements.

Patients

Key criteria for inclusion in Study 402 were age of 18 to 85 years, a history of HypoPT for 12 months or longer before initiating rhPTH(1–84), and previous participation in the preceding HEXT study with uninterrupted rhPTH(1–84) treatment.²² Exclusion criteria included the administration of any other investigational product within 30 days of giving consent to enrollment and increased baseline risk of osteosarcoma.²²

Inclusion criteria for enrolment in Study 404 were: aged 18 or older and completion of the PARALLAX study. Exclusion criteria included use of any other investigational study drug within 3 months before the screening visit, treatment with PTH, PTH analog, or PTH fragment in the 30 days before the screening visit, and increased baseline risk for osteosarcoma.²³

Response Scale and Scoring

The original 24-hour recall version of the HypoPT-SD was administered in Study 402. An 11-point rating scale was used for item 1 (muscle cramps) and a 5-point rating scale for items 2–7 of the Symptom subscale (Table 1). Item 1 was rescaled from an 11-point response scale to a 5-point response scale. A modified 7-day recall period version of the HypoPT-SD was administered in Study 404. Items 1–7 of the Symptom subscale used a 5-point rating scale (Table 1). The HypoPT-SD Symptom subscale was computed as the mean scores of the symptom items (items 1–7) and ranged from 0 to 4 points. Patient-reported scores were recorded during visits at baseline, at months 6, 30 and 36 (end of treatment [EOT]) in Study 402, and at baseline and week 52 (EOT) in Study 404.

Response Distribution

Descriptive statistics of HypoPT-SD item and domain scores were produced, to evaluate non-response, floor and ceiling effects, and patterns of change. This was done at baseline and month 30 and EOT in Study 402 and at baseline and EOT in Study 404. Floor and ceiling effects refer to instances in which the proportion of patients recording the lowest or highest response at baseline are more than twice as high as would be expected if a uniform distribution was observed (eg, 40% for a 5-point scale).

Ability to Detect Change

The ability of the HypoPT-SD to detect change was assessed through comparing the change in symptom subscale scores from baseline to end of treatment in a subset of patients in Study 402 who had albumin-corrected serum calcium (ACSC) levels outside the reference range at baseline and within the reference level (7.5–10.6 mg/dL) at end of treatment, with the change in symptom subscale scores observed in patients with ACSC levels outside the reference range. The same comparison was carried out between a subset of patients in Study 404 who reported improved PGI-S scores over the observations period and patients whose PGI-S scores either got worse or did not change. The differences in change in scores were examined using analysis of variance (ANOVA) in both cases.

Inter-Item Correlation and Item-Total Correlations

The structure of the HypoPT-SD Symptom subscale was analyzed by comparing inter-item correlations between all pairs of the HypoPT-SD item scores at baseline and EOT, and individual item scores and total symptom subscale scores at baseline.

Internal Consistency and Reliability

Reliability was evaluated using Cronbach's coefficient α to determine the internal consistency of the items included in the HypoPT-SD at baseline. Cronbach's α values of 0.7–0.9 indicate that an item is strongly associated with other items in a scale, but is not redundant.^{24,25}

Test–Retest Reliability

To assess the test–retest reliability of HypoPT-SD measurements when applied on two separate occasions that were 6 months apart in Study 402 (month 30 and EOT [month 36]) and 52 weeks apart in Study 404 (baseline and EOT [week 52]), intraclass correlation coefficients (ICCs) were computed and compared using a two-way (participant and time) mixed effects ANOVA. Patients in Study 404 were essentially rhPTH(1-84)-naïve; to ensure that any improvement in symptoms owing to PTH replacement did not affect the measurement of test–retest reliability, the same process was used on a subset of patients with stable disease, as determined by the Patient Global Impression of Severity (PGI-S) tool, measured at baseline and EOT. ICCs that exceed 0.70 for multi-item scales indicate that the measurements taken at different time points are stable and that test-retest reliability is acceptable.²⁶

Construct Validity

Construct validity was determined by performing correlational analyses between the scores recorded for the HypoPT-SD and those measured using the Functional Assessment of Cancer Therapy – Cognitive Function (FACT-Cog) and RAND-36 PRO tools. In addition, an ANOVA model was used to compare HypoPT-SD scores among severity groups defined by existing PRO tools to confirm the known-groups validity of the HypoPT-SD Symptom subscale at baseline and EOT.

Results

Patient Demographics

Psychometric analyses included data from 38 patients from Study 402 and 22 patients from Study 404 (Table 2). In Study 402, all patients were enrolled in the USA, most participants were women (78.9%), and most were White (97.4%). Participants had

Table 2 Patient Demographics and Characteristics at Baseline: Study 402 and Study 404

	Study 402 (n = 38)	Study 404 (n = 22)
Age, years		
n (%)	38 (100.0)	22 (100.0)
Mean (SD)	52.0 (12.4)	50.0 (11.4)
Median (range)	51.5 (33.0–79.0)	47.5 (28.0–71.0)
Sex, n (%)		
Male	8 (21.1)	4 (18.2)
Female	30 (78.9)	18 (81.8)
Race, n (%)		
White	37 (97.4)	20 (90.9)
Asian	1 (2.6)	–
Black or African American	–	1 (4.5)
Other	–	1 (4.5)
Ethnicity, n (%)		
Hispanic or Latino	1 (2.6)	–
Not Hispanic or Latino	33 (86.8)	21 (95.5)
Not reported	1 (2.6)	–
Unknown	3 (7.9)	1 (4.5)

(Continued)

Table 2 (Continued).

	Study 402 (n = 38)	Study 404 (n = 22)
Duration of HypoPT, years		
n (%)	38 (100.0)	21 (95.5)
Mean (SD)	18.9 (12.0)	10.3 (10.2)
Median (range)	15.8 (4.9–51.9)	7.5 (1.7–39.0)

Abbreviations: HypoPT, hypoparathyroidism; SD, standard deviation.

a mean age of 52.0 years (standard deviation [SD], 12.4; range, 33–79) and a mean HypoPT duration of 18.9 years (SD, 12.0; range, 4.9–51.9). In Study 404, most participants were women (81.8%) and most were White (90.9%). Participants had a mean age of 50.0 years (SD, 11.4; range, 28–71) and a mean HypoPT duration of 10.3 years (SD, 10.2; range, 1.7–39.0). Patients were enrolled in Hungary, the USA, Canada, and Denmark (41%, 27%, 23%, and 9%, respectively).

Response Distributions

In both studies, mean and median HypoPT-SD Symptom subscale scores were 1.2 or less at baseline and EOT (scores range from 0 [no symptoms] to 4 very severe) (Figure 1A and B). Score distributions for the supportive measures (eg, Functional Assessment of Chronic Illness Therapy – Fatigue) also indicated that patients had mild and stable disease (data not shown). In both studies, the highest single (most severe) item symptom scores were reported for physical fatigue whereas the lowest scores (least severe) were reported for heaviness in arms and legs (Figure 1C and D).

In Study 402 ceiling effects were observed at baseline for muscle spasms, and at baseline and EOT for numbness and heaviness in arms and legs. In Study 404, ceiling effects were observed at baseline for numbness; at EOT for muscle cramps, muscle spasms, and physical fatigue; and at baseline and EOT for heaviness in arms and legs, and confusion. No floor effects (>40% of patients giving a score of 4) were observed for any of the HypoPT-SD Symptom subscale items.

Ability to Detect Change

Patients in Study 402 whose mean ACSC level changed from being above or below to within the reference range (7.5–10.6 mg/dL) (n = 5) had greater improvements in HypoPT-SD symptom subscale scores than patients with ACSC levels that remained stable (n = 22) or moved outside the reference range (n = 3) (F-statistic [2,27] 5.59, $P = 0.0093$).

The improvement in mean HypoPT-SD symptom subscale scores were greater among patients in Study 404 who reported an improvement in PGI-S score (n = 8) than in patients whose PGI-S score remained stable (n = 10) or deteriorated (n = 4) (F-statistic [2,19] 8.69, $P = 0.0021$).

Inter-Item Correlation at Baseline

Within the HypoPT-SD Symptom subscale, inter-item pairwise correlations at baseline were stronger than 0.5 for all except five pairs of items, using the Study 402 data set (heaviness of arms and legs, and tingling; heaviness of arms and legs, and numbness; and heaviness of arms and legs, and confusion; physical fatigue and tingling; physical fatigue and numbness) (Figure 2A). All pairwise inter-item correlations were stronger than 0.5 for the Study 404 data set (Figure 2B).

In Study 404, item pair correlations of 0.8 or higher were computed for muscle spasms and physical fatigue and tingling; numbness and heaviness in arms and legs; and physical fatigue and numbness, heaviness in arms and legs, and confusion. Correlations between the scores for numbness and tingling, and muscle spasms and muscle cramps, were above 0.8 in both studies.

Item–Total Correlations at Baseline

All of the individual items had strong correlations with the overall HypoPT-SD symptom subscale scores in both studies ($r \geq 0.5$). The strongest correlations with total scores were observed for muscle spasms in Study 402 ($r = 0.83$) and

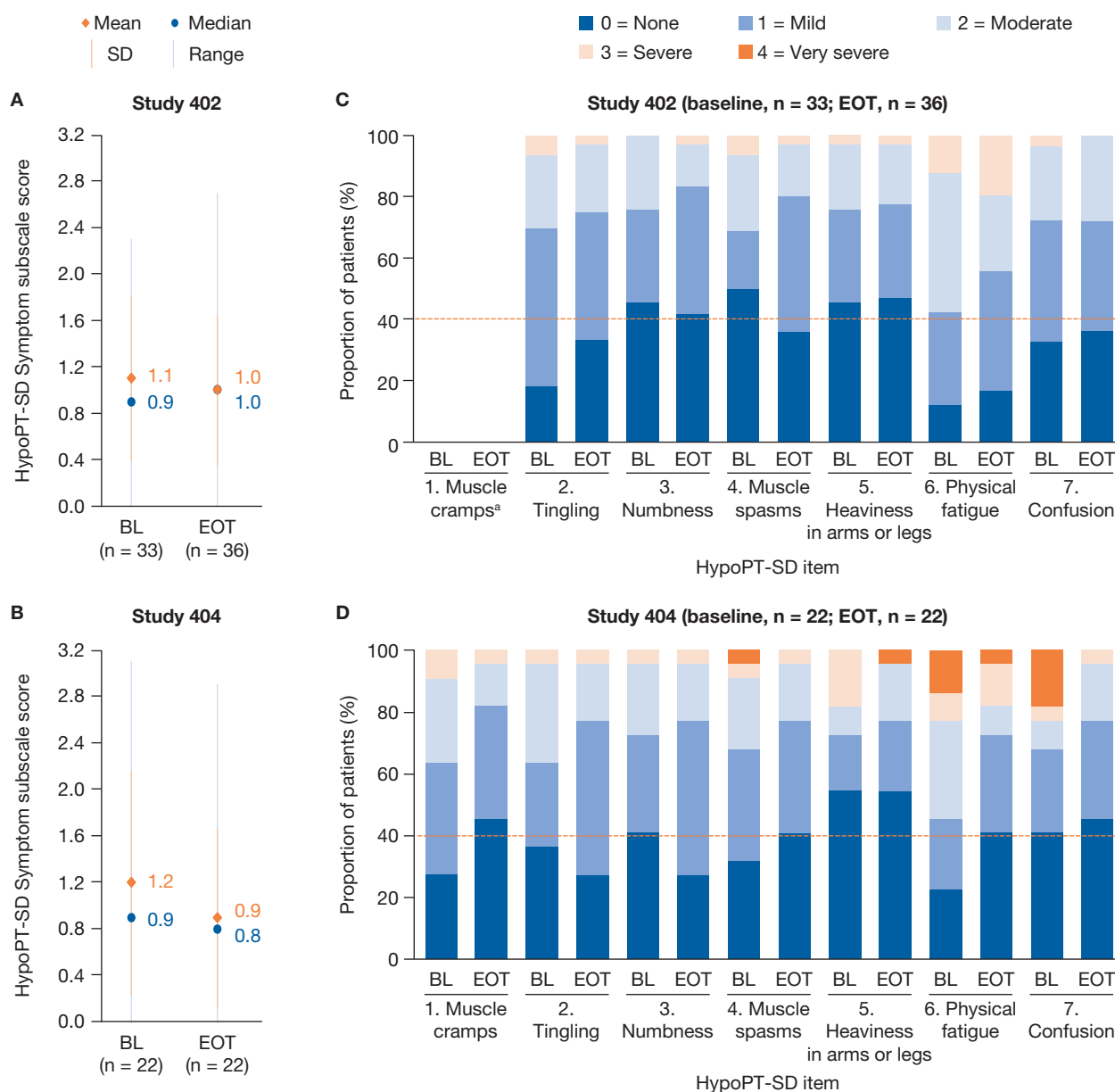


Figure 1 HypoPT-SD Symptom subscale-level score distributions (**A** and **B**) and item-level severity ratings (**C** and **D**) at baseline and EOT in the two studies.

Notes: EOT for Study 402 and Study 404 was at 36 months and 52 weeks, respectively. Orange dotted lines in panels (**C** and **D**) indicate the floor response threshold of 40% for a score of 0. ^aData are not shown for item 1 in Study 402 because the original item consisted of 11 points (0–10) and it was not included in the analysis plan.

Abbreviations: BL, baseline; EOT, end of treatment; HypoPT-SD, hypoparathyroidism symptom diary; SD, standard deviation.

physical fatigue and muscle spasms in Study 404 ($r = 0.85$ and 0.84 , respectively). The weakest correlations were observed for heaviness in arms and legs in both studies (0.55 in Study 402 and 0.70 in Study 404) (Figure 2).

Internal Consistency Reliability

Cronbach's α values for the HypoPT-SD Symptom subscale were 0.88 using data from Study 402 and 0.92 using data from Study 404.

Test–Retest Reliability

ICCs exceeded 0.70 using test–retest data from all patients in Study 402 and from the subgroup of patients with stable PGI-S scores from Study 404 (Figure 3). The ICC for all patients in Study 404 was 0.67 and the 95% confidence interval (CI) included 0.70 .

A Study 402

	HypoPT-SD item							
HypoPT-SD item	1	2	3	4	5	6	7	Total score
1. Muscle cramps	–							0.73
2. Tingling	0.55	–						0.68
3. Numbness	0.52	0.87	–					0.62
4. Muscle spasms	0.81	0.78	0.73	–				0.83
5. Heaviness in arms or legs	0.54	0.32	0.29	0.61	–			0.55
6. Physical fatigue	0.63	0.44	0.39	0.56	0.78	–		0.62
7. Confusion	0.63	0.75	0.64	0.74	0.35	0.60	–	0.68

B Study 404

	HypoPT-SD item							
HypoPT-SD item	1	2	3	4	5	6	7	Total score
1. Muscle cramps	–							0.74
2. Tingling	0.54	–						0.75
3. Numbness	0.52	0.96	–					0.80
4. Muscle spasms	0.89	0.80	0.77	–				0.84
5. Heaviness in arms or legs	0.71	0.66	0.85	0.65	–			0.70
6. Physical fatigue	0.78	0.75	0.82	0.80	0.82	–		0.85
7. Confusion	0.75	0.69	0.64	0.79	0.55	0.82	–	0.77

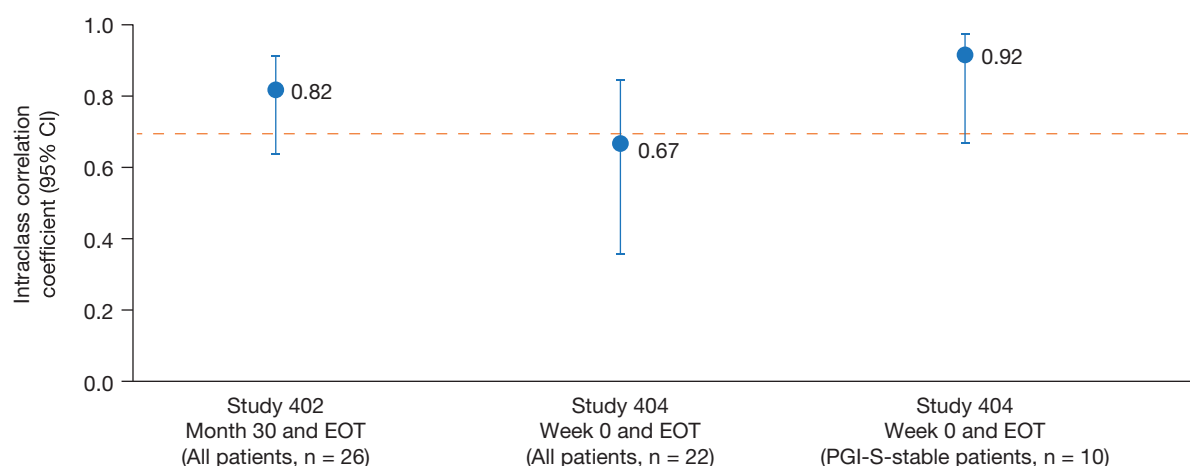
Figure 2 HypoPT-SD Symptom subscale inter-item and item-total correlations at baseline in (A) Study 402 and (B) Study 404.

Notes: According to guidelines for interpreting the strength of correlation coefficients, values ≥ 0.5 are strong (blue), 0.30–0.49 are moderate (grey), 0.10–0.29 are weak (grey), and < 0.10 are trivial.²⁷ Item 1 (muscle cramps) was rescaled from 0–10 to 0–4 points for the HypoPT-SD responses in Study 402.

Abbreviation: HypoPT-SD, hypoparathyroidism symptom diary.

Construct Validity

Using baseline and EOT data from Study 402, HypoPT-SD Symptom subscale scores had moderate (0.30–0.49) or strong (≥ 0.50) correlations with scores from other conceptually linked PRO tools, except for a weak correlation (0.10–0.29) with FACT-Cog perceived cognitive impairments scores at EOT (Figure 4A).²⁷ Hypotheses about the construct correlations (Figure 4A) were not met in three instances: the FACT-Cog perceived cognitive impairments correlation at EOT, and FACT-Cog impact on QoL correlations at baseline and EOT.

**Figure 3** HypoPT-SD Symptom subscale test–retest reliability for patients in Study 402 (month 30 and EOT) and Study 404 (week 0 and EOT).

Notes: EOT for Study 402 and Study 404 was at 36 months and 52 weeks, respectively. Orange horizontal line indicates the threshold above which multi-item scales are considered to be reliable.²⁶

Abbreviations: CI, confidence interval; EOT, end of treatment; HypoPT-SD, hypoparathyroidism symptom diary; PGI-S, Patient Global Impression of Severity.

A Study 402

HypoPT-SD	FACIT-F total	FACT-Cog: perceived cognitive impairments	FACT-Cog: impact on quality of life	RAND-36 physical health composite	RAND-36 mental health composite
Baseline	−0.6	−0.4	−0.5	−0.6	−0.4
EOT	−0.5	−0.3	−0.4	−0.5	−0.4

B Study 404

HypoPT-SD	EQ VAS	PGI-S
Baseline	−0.7	0.9
EOT	−0.3	0.8

White	Convergent validity hypothesized and met	Orange	Convergent validity hypothesized, but <i>not</i> met
White	Divergent validity hypothesized and met	Black	Divergent validity hypothesized, but <i>not</i> met

Figure 4 Construct validity correlations among the HypoPT-SD Symptom subscale and other conceptually linked patient-reported outcome tools at baseline and EOT in (A) Study 402 and (B) Study 404.

Notes: According to guidelines for interpreting the strength of correlation coefficients, values ≥ 0.5 are strong, 0.30–0.49 are moderate, 0.10–0.29 are small, and < 0.10 are trivial.²⁷ Hypotheses: negative correlations were expected with FACIT-F (moderate), FACT-Cog (moderate with the perceived cognitive impairment scale and stronger than with the impact on quality of life scale), RAND-36 (moderate with the physical health composite scale and stronger than with the mental health composite scale), and EQ VAS (moderate) scales. Moderate positive correlations were expected with PGI-S.

Abbreviations: EOT, end of treatment; EQ VAS, EuroQol visual analog scale; FACIT-F, Functional Assessment of Chronic Illness Therapy – Fatigue; FACT-Cog, Functional Assessment of Cancer Therapy – Cognitive Function; HypoPT-SD, hypoparathyroidism symptom diary; PGI-S, Patient Global Impression of Severity; RAND-36, RAND 36-Item Short Form Health Survey.

Using data from Study 404, HypoPT-SD Symptom subscale scores were strongly correlated with PGI-S scores at baseline and EOT. The correlation with EuroQol visual analog scale scores was strong at baseline, but weak at EOT, and the correlation hypothesis was not met at this time point (Figure 4B).

Known Groups

HypoPT-SD Symptom subscale scores differed significantly among severity groups defined by other PRO instruments. In Study 402, HypoPT-SD Symptom subscale scores for patients with mild and severe symptoms as identified by the RAND-36 general health perception questionnaire were 0.7 and 1.5, respectively, at baseline (least-squares [LS] mean difference [95% CI] −0.7 [−1.2 to −0.3], $P = 0.0021$) and 0.7 and 1.4 respectively at EOT (LS mean difference [95%] −0.7 [−1.2 to −0.3], $P = 0.0011$) (Figure 5A and B). The corresponding scores for groups with mild and severe symptoms identified by the PGI-S in Study 404 were 0.6 and 2.0 at baseline (LS mean difference [95%] −1.4 [−2.0 to −0.8], $P < 0.0001$), and 0.6 and 1.7 at EOT (LS mean difference [95%] −1.1 [−1.7 to −0.5], $P = 0.0005$) (Figure 5C and D).

Discussion

In this study, the psychometric properties of the HypoPT-SD were evaluated in two prospective clinical trials involving 60 patients with HypoPT. At baseline, the item-levels scores were positively skewed (low symptom burden) suggesting that HypoPT was well managed in patients enrolled in the clinical trials. This finding is supported by the baseline data from Studies 404 and 402: mean albumin-corrected serum calcium levels were close to or within the target range for patients with hypoparathyroidism (1.94 mmol/L and 2.1 mmol/L, respectively; target range 2.0–2.25 mmol/L).²⁸ Furthermore, all the patients enrolled in Study 402 or 404 for whom data were available were receiving calcium and vitamin D supplementation. In addition to receiving conventional therapy, patients enrolled in Study 402 had received rhPTH(1–84) for several years before entering the trial, and patients in Study 404 had received rhPTH(1–84) for two days before the trial began.

Strong inter-item correlations were observed for 88% of the symptom pairs and all individual items demonstrated strong correlations with overall Hypo-SD total scores. The strongest correlations between individual symptoms and total

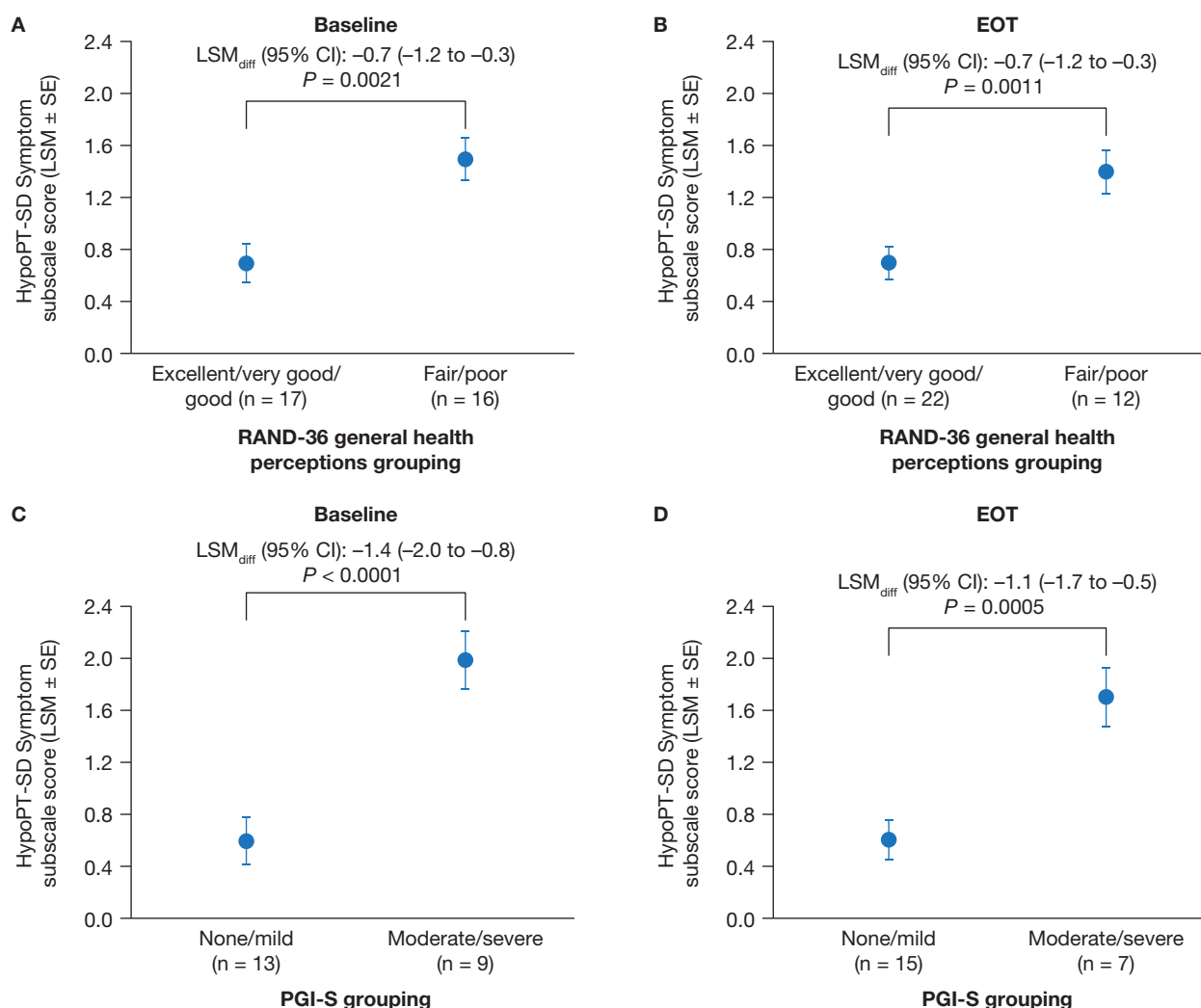


Figure 5 Known-groups validity for the HypoPT-SD Symptom subscale to discriminate between different severity groups on the RAND-36 general health perceptions scale in Study 402 (**A** and **B**) and on the PGI-S scale in Study 404 (**C** and **D**).

Notes: EOT for Study 402 and Study 404 was at 36 months and 52 weeks, respectively.

Abbreviations: CI, confidence interval; EOT, end of treatment; HypoPT-SD, hypoparathyroidism symptom diary; LSM, least-squares mean; LSM diff, least-squares mean difference; PGI-S, Patient Global Impression of Severity; RAND-36, RAND 36-Item Short Form Health Survey; SE, standard error.

scores were observed for muscle spasms and physical fatigue. The internal consistency of our tool was confirmed by Cronbach α values of 0.88 for data from Study 402 and 0.92 for data from Study 404, providing additional support.

Test-retest reliability of the HypoPT-SD was established by ICCs of 0.7 or higher when calculated for all patients in Study 402 and for sub group of patients in study 404 who were classified as stable based on PGI-S scores. Overall, the HypoPT-SD had moderate or strong correlations with all the other widely used PRO tools examined, and correlation hypotheses regarding the relationship between the HypoPT-SD and other PRO tools were met in over 70% of cases, lending support to the construct validity of our tool.

Significant differences in HypoPT were observed between groups of patients with different symptom severity (as established using other PRO instruments), demonstrating the ability of the tool to discriminate based on disease burden. Furthermore, initial analyses of ability to detect change suggest that the HypoPT-SD can capture alterations in disease severity.

Other HypoPT-specific PROs have been developed and tested. The HypoPT patient experience scale (HPES), which comprises a 17-item symptom scale and 26-item impact scale assessed using a 2-week recall period, has undergone psychometric evaluation using data from an observation study and a Phase 2 trial.^{16,17,29} Cronbach's α values exceeded

0.9 when calculated using data from either study. ICCs for patients whose treatment did not change during the observation period approached, but did not meet, the threshold of 0.7. Construct validity was assessed using data from the observational study only, correlations between HPES scores and those from contextually linked PRO instruments were moderate or strong.

The HypoPT questionnaire (HPQ), which comprises 28 questions across 5 domains, has been developed using an analytical-empirical approach and used to assess symptom load in patients with HypoPT.¹⁸ When employed in a cross-sectional study in Germany, HPQ scores related to pain and cramps, numbness and tingling, and heart palpitations were significantly higher in the postsurgical HypoPT group than a matched cohort who had undergone thyroid surgery but did not have HypoPT.³⁰ Furthermore, correlations between symptom burden and biochemical parameters were also observed—notably that calcium-phosphorus product was the strongest predictor of HypoPT symptom severity.^{18,30}

Both the HPES and the HPQ are considerably longer, and potentially more burdensome, than the HypoPT-SD.^{18,29} In addition, the 2-week recall period for the HPES may result in a less accurate picture of symptom severity than the shorter recall period used for our PRO instrument. Through focusing only on the symptoms that are most important to patients, the HypoPT-SD is brief, and therefore more suitable for daily use.¹⁹

Limitations

At baseline, few patients had severe HypoPT symptoms, which restricted the range of possible change in scores. Furthermore, the accuracy of the results may be limited due to differences in administration and recall periods (24 hours vs 7 days) between the two studies and the 11-point muscle cramp item requiring re-scaling to 5 points in Study 402.

Conclusion

There is an unmet need for practical, accurate and validated disease-specific tools that can measure HRQoL in patients with HypoPT. Such tools could track the burden of disease in patients and evaluate the impact of treatments on HRQoL in prospective trials with a greater degree of specificity and sensitivity than disease agnostic PRO instruments. We have developed and tested a short questionnaire to evaluate the physical and psychological burden of HypoPT and the impact it has on daily functioning. Future studies will evaluate the utility of this tool in larger patient samples; however, the results from this study indicate the reliability and validity of the disease-specific HypoPT-SD using data from prospective studies and support the use of this PRO instrument in future clinical trials.

Acknowledgments

This study was sponsored by Takeda Pharmaceuticals USA, Inc. The authors thank the study participants. The authors thank Miriam Durazo, PhD, of PharmaGenesis Cardiff, Cardiff, UK for providing medical writing support, which has been funded by Takeda Development Center Americas Inc., Lexington, MA, USA in accordance with Good Publication Practice 2022 (GPP 2022) guidelines (<https://www.ismpp.org/gpp-2022>). The Hypoparathyroidism Symptom Diary Response Scales in Table 1 are reprinted with permission of Shire Human Genetic Therapies, Inc. © 2016 Shire Human Genetic Therapies, Inc. All rights reserved.

Disclosure

Lauren Nelson and Susan Martin are employed by RTI Health Solutions. Jia Ma is a former employee of RTI Health Solutions and is now affiliated with Pfizer, USA. Steven W Ing is an advisory board member for, and received research grant support from, Takeda Pharmaceuticals USA, Inc. Mishaela R Rubin is a study investigator for Takeda Pharmaceuticals USA, Inc. and Ascendis Pharma. Rohini Sen is a former employee of Takeda Pharmaceuticals USA, Inc. and is now affiliated with AbbVie, USA. Olulade Ayodele is an employee of Takeda Development Center Americas Inc. and holds stock/stock options in Takeda. The authors report no other conflicts of interest in this work.

References

1. Bilezikian JP, Khan A, Potts JT Jr, et al. Hypoparathyroidism in the adult: epidemiology, diagnosis, pathophysiology, target-organ involvement, treatment, and challenges for future research. *J Bone Miner Res*. 2011;26(10):2317–2337. doi:10.1002/jbmr.483
2. Shoback D. Clinical practice. Hypoparathyroidism. *N Engl J Med*. 2008;359(4):391–403. doi:10.1056/NEJMc0803050
3. Bollerslev J, Rejnmark L, Zahn A, et al. European expert consensus on practical management of specific aspects of parathyroid disorders in adults and in pregnancy: recommendations of the ESE educational program of parathyroid disorders. *Eur J Endocrinol*. 2021;186(2):R33–R63. doi:10.1530/EJE-21-1044
4. Shoback DM, Bilezikian JP, Costa AG, et al. Presentation of hypoparathyroidism: etiologies and clinical features. *J Clin Endocrinol Metab*. 2016;101(6):2300–2312. doi:10.1210/jc.2015-3909
5. Gosmanova EO, Chen K, Ketteler M, et al. Risk of cardiovascular conditions in patients with chronic hypoparathyroidism: a retrospective cohort study. *Adv Ther*. 2021;38(8):4246–4257. doi:10.1007/s12325-021-01787-7
6. Gosmanova EO, Houillier P, Rejnmark L, Marelli C, Bilezikian JP. Renal complications in patients with chronic hypoparathyroidism on conventional therapy: a systematic literature review: renal disease in chronic hypoparathyroidism. *Rev Endocr Metab Disord*. 2021;22(2):297–316. doi:10.1007/s11154-020-09613-1
7. Fonseca OA, Calverley JR. Neurological manifestations of hypoparathyroidism. *Arch Intern Med*. 1967;120(2):202–206. doi:10.1001/archinte.1967.00300020074009
8. Arlt W, Fremerey C, Callies F, et al. Well-being, mood and calcium homeostasis in patients with hypoparathyroidism receiving standard treatment with calcium and vitamin D. *Eur J Endocrinol*. 2002;146(2):215–222. doi:10.1530/eje.0.1460215
9. Siggekkow H, Clarke BL, Germak J, et al. Burden of illness in not adequately controlled chronic hypoparathyroidism: findings from a 13-country patient and caregiver survey. *Clin Endocrinol (Oxf)*. 2020;92(2):159–168. doi:10.1111/cen.14128
10. Buttner M, Krogh D, Siggekkow H, Singer S. What are predictors of impaired quality of life in patients with hypoparathyroidism? *Clin Endocrinol (Oxf)*. 2022;97(3):268–275. doi:10.1111/cen.14701
11. Astor MC, Lovas K, Debowska A, et al. Epidemiology and health-related quality of life in hypoparathyroidism in Norway. *J Clin Endocrinol Metab*. 2016;101(8):3045–3053. doi:10.1210/jc.2016-1477
12. Vokes T. Quality of life in hypoparathyroidism. *Bone*. 2019;120:542–547. doi:10.1016/j.bone.2018.09.017
13. Buttner M, Musholt TJ, Singer S. Quality of life in patients with hypoparathyroidism receiving standard treatment: a systematic review. *Endocrine*. 2017;58(1):14–20. doi:10.1007/s12020-017-1377-3
14. Sikjaer T, Moser E, Rolighed L, et al. Concurrent Hypoparathyroidism is associated with impaired physical function and quality of life in hypothyroidism. *J Bone Miner Res*. 2016;31(7):1440–1448. doi:10.1002/jbmr.2812
15. Cusano NE, Rubin MR, McMahon DJ, et al. The effect of PTH(1-84) on quality of life in hypoparathyroidism. *J Clin Endocrinol Metab*. 2013;98(6):2356–2361. doi:10.1210/jc.2013-1239
16. Brod M, McLeod L, Markova D, et al. Psychometric validation of the Hypoparathyroidism Patient Experience Scales (HPES). *J Patient Rep Outcomes*. 2021;5(1):70. doi:10.1186/s41687-021-00320-2
17. Brod M, Waldman LT, Smith A, Karpf D. Assessing the patient experience of hypoparathyroidism symptoms: development of the Hypoparathyroidism Patient Experience Scale-Symptom (HPES-Symptom). *Patient*. 2020;13(2):151–162. doi:10.1007/s40271-019-00388-5
18. Wilde D, Wilken L, Stamm B, et al. The HPQ—development and first administration of a questionnaire for hypoparathyroid patients. *JBMR Plus*. 2020;4(1):e10245. doi:10.1002/jbmr.10245
19. Martin S, Chen K, Harris N, Vera-Llonch M, Krasner A. Development of a patient-reported outcome measure for chronic hypoparathyroidism. *Adv Ther*. 2019;36(8):1999–2009. doi:10.1007/s12325-019-00999-2
20. US Food and Drug Administration. Guidance for industry. Patient-reported outcome measures: use in medical product development to support labeling claims; 2009. Available from: <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf>. Accessed December 5, 2022.
21. Coles T, Chen K, Nelson L, et al. Psychometric evaluation of the hypoparathyroidism symptom diary. *Patient Relat Outcome Meas*. 2019;10:25–36. doi:10.2147/PROM.S179310
22. ClinicalTrials.gov. A study of extended use of recombinant human parathyroid hormone (rhPTH(1-84)) in hypoparathyroidism NCT02910466; 2021. Available from: <https://clinicaltrials.gov/ct2/show/NCT02910466>. Accessed October 5, 2022.
23. ClinicalTrials.gov. Safety and efficacy study of rhPTH(1-84) in subjects with hypoparathyroidism NCT03364738; 2022. Available from: <https://clinicaltrials.gov/ct2/show/NCT03364738>. Accessed October 5, 2022.
24. Bland JM, Altman DG. Statistics notes: Cronbach's alpha. *BMJ*. 1997;314(7080):572. doi:10.1136/bmj.314.7080.572
25. Streiner DL. Starting at the beginning: an introduction to coefficient alpha and internal consistency. *J Pers Assess*. 2003;80(1):99–103. doi:10.1207/S15327752JPA8001_18
26. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155–163. doi:10.1016/j.jcm.2016.02.012
27. Cohen J. A power primer. *Psychol Bull*. 1992;112(1):155–159. doi:10.1037/0033-2909.112.1.155
28. National Health Service. Hypoparathyroidism; 2021. Available from: <https://www.nhs.uk/conditions/hypoparathyroidism/>. Accessed March 29, 2023.
29. Brod M, Waldman LT, Smith A, Karpf D. Living with hypoparathyroidism: development of the Hypoparathyroidism Patient Experience Scale-Impact (HPES-Impact). *Qual Life Res*. 2021;30(1):277–291. doi:10.1007/s11136-020-02607-1
30. Wilde D, Wilken L, Stamm B, et al. Quantification of symptom load by a disease-specific questionnaire HPQ 28 and analysis of associated biochemical parameters in patients with postsurgical hypoparathyroidism. *JBMR Plus*. 2020;4(7):e10368. doi:10.1002/jbmr.10368

Patient Related Outcome Measures

Dovepress

Publish your work in this journal

Patient Related Outcome Measures is an international, peer-reviewed, open access journal focusing on treatment outcomes specifically relevant to patients. All aspects of patient care are addressed within the journal and practitioners from all disciplines are invited to submit their work as well as healthcare researchers and patient support groups. The manuscript management system is completely online and includes a very quick and fair peer-review system. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <http://www.dovepress.com/patient-related-outcome-measures-journal>