Dermatol Ther (Heidelb)
https://doi.org/10.1007/s13555-025-01346-w

ORIGINAL RESEARCH

Hidradenitis Suppurativa Symptom Daily Diary (HSSDD) and Questionnaire (HSSQ): Psychometric Validation and Interpretation Threshold Derivation Using Phase 3 Study Data

John R. Ingram · Jérémy Lambert · Valerie Ciaravino · Robert Rolleri · Ingrid Pansar · Luke Peterson · Christopher G. Pelligra · Linnea Thorlacius

Received: December 3, 2024 / Accepted: January 21, 2025

ABSTRACT

***Introduction*:** Hidradenitis suppurativa (HS) is a chronic inflammatory disease characterised by painful skin lesions which negatively impact patients' physical and mental wellbeing. The HS Symptom Daily Diary (HSSDD) and HS Symptom Questionnaire (HSSQ) are patient-reported outcome (PRO) tools capturing patient-perceived severity of HS symptoms. Here, we report the psychometric properties of HSSDD and HSSQ along with score interpretation thresholds.

***Methods*:** Pooled data from patients with moderate to severe HS in two phase 3 studies (BE HEARD I II) were analysed. Test-retest reliability was evaluated using intraclass correlation coefficients (ICCs). Convergent validity was assessed between the HSSDD ($N=934$) and HSSQ ($N=1007$) compared with relevant PROs and clinician-reported outcomes (ClinROs) at baseline and Week (Wk)16. Known-groups validity was assessed, comparing HSSDD and HSSQ scores between participant subgroups pre-defined using PRO/ClinRO measures (Patient Global Impression [PGI] of HS severity, Hurley stage, International HS Severity Score System). Responsiveness was evaluated by correlating changes from baseline to Wk16 in HSSDD and HSSQ scores with changes in PGI scales. Clinically meaningful within-patient improvement thresholds were estimated using anchor- and distribution-based analyses. Symptom/impact severity thresholds were estimated using receiver operating characteristic curve analyses.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s13555-025-01346-w.

J. R. Ingram (✉)
Division of Infection and Immunity, Department of Dermatology & Academic Wound Healing, Cardiff University, Cardiff, UK
e-mail: ingramjr@cardiff.ac.uk

J. R. Ingram
European Hidradenitis Suppurativa Foundation (EHSF), Dessau, Germany

J. Lambert · V. Ciaravino
UCB, Colombes, France

R. Rolleri · L. Peterson
UCB, Morrisville, NC, USA

I. Pansar
UCB, Brussels, Belgium

C. G. Pelligra
Evidera, Atlanta, GA, USA

L. Thorlacius
Department of Dermatology, Zealand University Hospital, Roskilde, Denmark

L. Thorlacius
Health Sciences Faculty, University of Copenhagen, Copenhagen, Denmark

Published online: 28 March 2025

***Results*:** At Wk16, HSSDD and HSSQ completion rates were 70.1% and 90.2%, respectively. Test-retest reliability analyses demonstrated good score reproducibility (ICC: HSSDD: 0.80–0.86; HSSQ: 0.73–0.82). Correlations between HSSDD and HSSQ scores and other PROs/ClinROs were generally consistent with predefined hypotheses, indicating good convergent validity. HSSDD and HSSQ scores discriminated between pre-defined subgroups, confirming known-groups validity. Sixteen-wk changes from baseline in HSSDD and HSSQ scores and anchors were moderately to strongly correlated (>0.30), establishing responsiveness. Interpretation thresholds for both HSSDD and HSSQ were estimated.

***Conclusion*:** HSSDD and HSSQ item scores demonstrated good psychometric performance in participants with moderate to severe HS. The clinically meaningful severity thresholds defined here could be used to assess treatment efficacy.

***Clinical Trial registration*:** NCT04242446; NCT04242498.

Graphical Abstract:

Hidradenitis Suppurativa Symptom Daily Diary (HSSDD) and Questionnaire (HSSQ): Psychometric Validation and Interpretation Threshold Derivation Using Phase 3 Study Data

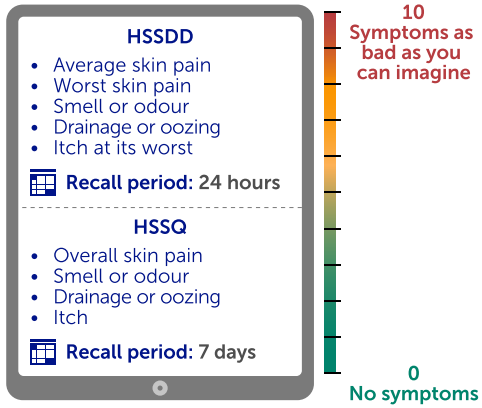
John R. Ingram, Jérémy Lambert, Valerie Ciaravino, Robert Rolleri, Ingrid Pansar, Luke Peterson, Christopher G. Pelligra, Linnea Thorlacius

Adis PEER-REVIEWED FEATURE

Introduction

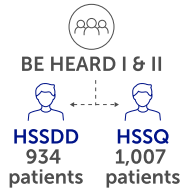
Hidradenitis suppurativa (HS) is a chronic inflammatory disease characterised by painful skin lesions which negatively impact patients' physical and mental wellbeing.

The **HS Symptom Daily Diary (HSSDD)** and **HS Symptom Questionnaire (HSSQ)** are patient-reported outcome tools developed to capture patient-perceived severity of symptoms.



Methods

Psychometric performance of the HSSDD and HSSQ total and subscale scores were assessed. Thresholds for **clinically meaningful within-patient improvements** and for different symptom/impact severity levels were determined using pooled blinded data from **two phase 3 (BE HEARD I & II)** trials evaluating bimekizumab in patients with **moderate to severe HS**.



Results

- HSSDD and HSSQ demonstrated:
  - ☑ Good score reproducibility
  - ☑ Convergent and known-groups validity
  - ☑ Good sensitivity to change over time
- Thresholds of none, mild, moderate, severe and very severe symptom/impact **severity levels** were identified.
- **Clinically meaningful within-patient improvement** thresholds were identified.

Worst and average skin pain	Smell or odour	Drainage or oozing	Itch
HSSDD: 3- to 4- HSSQ: 4- to 5- point decrease	HSSDD: 2- to 3- HSSQ: 3- to 4- point decrease	HSSDD: 2- to 3- HSSQ: 3- to 4- point decrease	HSSDD: 2- to 3- HSSQ: 3- point decrease

Conclusions

Both HSSDD and HSSQ are **valid, reliable and responsive** (i.e. fit-for-purpose) **patient-reported outcome** measures for HS symptoms. The **clinically meaningful within-patient thresholds** can assist with interpreting the scores. This study supports the use of HSSDD and HSSQ to assess treatment impacts and inform physician decisions.

The graphical abstract represents the opinions of the authors. For a full list of declarations, including funding and author disclosure statements, and copyright information, please see the full text online.

PLAIN LANGUAGE SUMMARY

Hidradenitis suppurativa (HS) is a chronic skin condition that causes lesions and painful lumps under the skin. HS can affect patients' lives by causing pain, emotional distress and difficulty completing daily activities. Currently, there are few medications to treat HS. To understand the impact and effectiveness of new treatments, it is important to look beyond clinical outcomes and capture patient experience. To measure the patient's perspective and more specifically symptom experience, self-completed questionnaires such as the HS Symptom Daily Diary (HSSDD) and HS Symptom Questionnaire (HSSQ) were developed. The HSSDD and HSSQ determine patients' perspective on the severity of their HS symptoms (pain, itch, smell or odour and drainage or oozing). Two phase 3 trials used HSSDD and HSSQ to investigate patients' perspective on the severity of their symptoms. We conducted a series of statistical analyses to assess the validity, reliability and robustness of both questionnaires. We found that HSSDD and HSSQ could assess patients' experience of symptoms. We showed that both questionnaires were sensitive enough to reveal changes over time. Furthermore, both questionnaires were able to distinguish between patient groups with different levels of HS symptom severity. We also established thresholds that will help clinicians determine whether an improvement in a patient's HSSDD/HSSQ scores are meaningful to the patient. The results from this study show HSSDD and HSSQ are reliable patient-completed questionnaires that could be useful in informing treatment choices.

Keywords: Bimekizumab; Hidradenitis suppurativa; Patient-reported outcomes; Psychometric validation; Symptoms

Key Summary Points

Why carry out this study?

The Hidradenitis Suppurativa Symptom Daily Diary (HSSDD) and HS Symptom Questionnaire (HSSQ) are patient-reported outcome (PRO) measures that have been specifically developed to capture patient-perceived severity of core HS symptoms (pain, itch, smell or odour and drainage or oozing) over the last 24 h (HSSDD) or 7 days (HSSQ).

It is essential to evaluate the psychometric properties of an outcome measure to ensure it is fit for purpose in the context of use to assess the efficacy of a treatment in patients with moderate to severe HS.

It is also important to determine clinically meaningful within-patient change and severity thresholds for that outcome measure to help interpret scores.

What was learned from the study?

The results from this study show that both HSSDD and HSSQ are valid, reliable and responsive (i.e. fit for purpose) PRO measures for HS symptoms. The study has also defined clinically meaningful within-patient thresholds that can assist with interpreting the scores derived from these measures.

DIGITAL FEATURES

This article is published with digital features, including a graphical abstract to facilitate understanding of the article. To view digital features for this article go to https://doi.org/10.6084/m9.figshare.28228928.

INTRODUCTION

Hidradenitis suppurativa (HS), which affects approximately 0.4–1.0% of the population globally, is a chronic, relapsing and debilitating inflammatory skin disease [1]. Characterised by painful inflammatory nodules, abscesses and draining tunnels, HS is associated with a significant detrimental impact on patients' quality of life (QoL) [2–4]. In particular, symptoms such as chronic pain, itch, smell or malodour and suppuration can negatively impact patients' physical, mental and social wellbeing [3–7].

Pain impacts a patient's daily life, bringing discomfort, immobility, difficulty sleeping, depressed mood, irritability, social isolation and decreased work productivity [8]. Itch and odour can also strongly impair patients' health-related QoL (HRQoL) [9].

Changes in patients' experiences of HS core symptoms, along with changes in physical, emotional and social functioning, are key factors in the holistic evaluation of efficacy of treatments for moderate to severe HS. An HS-specific core outcome set of domains (i.e. agreed minimum set of outcomes to measure in all clinical trials) has been established by the Hidradenitis SuppuraTiva cORe outcomes set International Collaboration (HiSTORIC) using a Delphi process involving both patients and health care providers [10]. This includes the concurrent measurement of five domains agreed upon by both patients and health care providers: pain, physical signs, HS-specific QOL, global assessment and progression of course. A sixth domain, symptoms, covering drainage and fatigue, was added because it received strong support from the patient stakeholder group. Whilst patient-reported outcome (PRO) measures that are skin disease specific (e.g. Dermatology Life Quality Index [DLQI]) or HS specific (e.g. Hidradenitis Suppurativa Quality of Life Questionnaire [HiSQOL]) exist, the items covered by these PROs are geared towards capturing the impact of the condition on patients' HRQoL [11] and may not capture the severity level of HS-specific symptoms.

The HS Symptom Daily Diary (HSSDD) and HS Symptom Questionnaire (HSSQ) are PRO tools that have been specifically developed to capture patient-perceived severity of core HS symptoms (pain, itch, smell or odour and drainage or oozing) over the last 24 h (HSSDD) or 7 days (HSSQ). These measures were developed based on an initial literature review that captured the core symptoms associated with HS, interviews with two clinicians with expertise in HS and a review of existing clinical outcome assessment (COA) measures from published and on-going trials. Items were generated to adequately capture the given symptom-related concepts and to ensure they were clearly defined, had clinical relevance and were appropriate for use in the context of pivotal clinical trials for the treatment and management of patients with moderate to severe HS. Cognitive debriefing interviews with 20 participants diagnosed with moderate to severe HS were conducted to confirm the relevance of the symptoms covered by the HSSDD and HSSQ and to determine the understandability and/or usability of the measures [12].

As a next step, it is essential to evaluate the psychometric properties of a measure to ensure it is fit for purpose in context of use to assess the efficacy of a treatment in patients with moderate to severe HS [13–15]. It is also important to determine clinically meaningful within-patient change and severity thresholds for the measure to help interpret scores and changes in scores [13–15].

In this article, we present the results of the assessment of the psychometric properties and derivation of interpretation thresholds of the HSSDD and HSSQ item scores using pooled blinded data from two phase 3 trials (BE HEARD I and II) evaluating bimekizumab efficacy and safety in moderate to severe HS.

METHODS

BE HEARD I and II Study Design and Patients

This psychometric analysis was conducted on blinded data, pooled from two identically designed phase 3 trials, BE HEARD I (NCT04242446) and II (NCT04242498) [16, 17]. The trials included an initial (Weeks 0–16) and

Table 1 Demographics and baseline disease characteristics of patients

Characteristic	HSSDD analysis set (*N* = 934)	HSSQ analysis set (*N* = 1007)
Age, years, mean (SD)	37.1 (12.2)	36.7 (12.2)
Female, *n* (%)	525 (56.2%)	569 (56.5%)
Race, n (%)		
White	722 (77.3%)	771 (76.6%)
Black or African American	88 (9.4%)	103 (10.2%)
Asian	37 (4.0%)	41 (4.1%)
Other or mixed[a]	44 (4.7%)	46 (4.6%)
Missing	43 (4.6%)	42 (4.2%)
Region, n (%)		
North America	350 (37.5%)	382 (37.9%)
Western Europe	271 (29.0%)	290 (28.8%)
Central and Eastern Europe	243 (26.0%)	260 (25.8%)
Asia and Australia	70 (7.5%)	75 (7.5%)
BMI		
< 25 kg/m^2	140 (15.0%)	155 (15.4%)
25 to < 30 kg/m^2	236 (25.3%)	253 (25.1%)
≥ 30 kg/m^2	555 (59.4%)	596 (59.2%)
Missing	3 (0.3%)	3 (0.3%)
Duration of disease, years, mean (SD)	8.0 (7.8)	7.9 (7.8)
Hurley stage,[b] *n* (%)		
II	524 (56.1%)	561 (55.7%)
III	409 (43.8%)	446 (44.3%)
Symptom item scores, mean (SD)		
Worst skin pain	5.5 (2.5)	N/A[c]
Average skin pain	4.8 (2.5)	N/A[c]
Skin pain	N/A[c]	5.8 (2.4)
Smell or odour	4.4 (3.0)	4.6 (3.0)
Itch	4.7 (2.7)	5.0 (2.8)
Drainage or oozing	4.5 (2.8)	5.0 (2.8)

Proportions may not add up to 100% due to rounding

BMI body mass index, *HSSDD* Hidradenitis Suppurativa Symptom Daily Diary, *HSSQ* Hidradenitis Suppurativa Symptom Questionnaire, *kg* kilograms, *N/A* not applicable, *SD* standard deviation

[a]Other or mixed category includes American Indian/Alaska Native and Native Hawaiian/other Pacific patients

[b]Only patients with Hurley stage II and III were included at baseline, as per the BE HEARD I and II inclusion and exclusion criteria

[c]HSSDD assesses worst skin pain and average skin pain, while HSSQ assesses skin pain

maintenance (Weeks 16–48) treatment period. Adult patients with moderate to severe HS were randomised to receive (initial/maintenance) bimekizumab 320 mg every 2 weeks (Q2W)/Q2W, bimekizumab Q2W/every 4 weeks (Q4W), bimekizumab Q4W/Q4W or placebo/bimekizumab Q2W. Full inclusion/exclusion criteria have been previously published [18].

Moderate to severe disease was defined as ≥ 5 inflammatory lesions (abscesses and/or inflammatory nodules) affecting ≥ 2 distinct anatomic areas, one of which was Hurley Stage II or III (at both screening and baseline visits). Patients had a diagnosis of HS based on clinical history and physical examination for ≥ 6 months prior to the baseline visit.

HSSDD and HSSQ

The HSSDD consists of five items that assess worst skin pain, average skin pain, smell or odour, itch at its worst and drainage or oozing from HS lesions, experienced in the past 24 h. Each symptom item is rated on an 11-point numeric rating scale (NRS; from 0 ['no symptom'] to 10 ['symptom as bad as you can imagine']). Each HSSDD item score is derived as the weekly average of the daily scores from a given week (if ≥ 4 non-missing daily values are available, otherwise the item score is reported as missing). Higher scores indicate a higher level of symptomology.

The HSSQ consists of four items that assess overall skin pain, itch, smell or odour and drainage or oozing from HS lesions experienced in the past 7 days. Each symptom item is rated on an 11-point NRS, similar to the scale described above for HSSDD. Higher scores indicate a higher level of symptomology.

In BE HEARD I and II, the HSSDD was completed daily from screening to Week 16, and the HSSQ was completed at baseline, Week 16 and every other week to Week 48. The HSSDD and HSSQ were completed using electronic devices.

Other Assessments

Psychometric analysis utilised Hurley Stage (Stage II or III only included in the clinical studies) and HS lesion-based assessments, including the International Hidradenitis Suppurativa Severity Score System (IHS4) and HS Physician's Global Assessment (HS PGA), which assessed disease severity and activity from the clinician perspective. Patient global impression of HS severity (PGI-S-HS) and patient global impression of change in HS (PGI-C-HS) were used to measure patient perception of HS severity and its change over time. In addition, two measures were used that focused on patient perception of skin pain (patient global impression of severity of skin pain [PGI-S-SP] and patient global impression of change in severity of skin pain [PGI-C-SP]). The HiSQOL [19] and DLQI [11] are PRO measures used to capture HS-specific and skin-disease specific HRQoL, respectively, throughout the trials.

Psychometric Analyses

Psychometric analyses were conducted on the blinded HSSDD analysis set and the blinded HSSQ analysis set separately. The analysis sets for both measures were defined as all randomised study patients from both BE HEARD trials who had ≥ 1 non-missing weekly symptom item score of the HSSDD/HSSQ at any scheduled assessment visit.

Simulation analyses were conducted to assess the appropriateness of the current weekly scoring rule used to derive the weekly symptom item scores of the HSSDD (i.e. ≥ 4 out of 7 daily scores non-missing for a given week). This involved examining whether the variability of weekly symptom item scores of the HSSDD based on all missing-day scenarios (i.e. 1, 2, 3, 4, 5 or 6 days missing) would significantly differ from that based on the no missing-day rule (i.e. 7 out of 7 daily scores non-missing).

To assess convergent validity, Pearson's and Spearman's rank correlation coefficients and corresponding *p* values were calculated to assess the strength of associations between HSSDD and HSSQ item scores assessing similar (patient-reported DLQI and HiSQOL) and different (clinician-rated IHS4) concepts at baseline and Week 16. The correlation coefficient was used to interpret the strength of the correlation between

Table 2 Convergent validity for HSSDD and HSSQ at baseline and Week 16

Spearman s correlation, r		IHS4		DLQI Total		HiSQOL Total	
		Baseline	Week 16	Baseline	Week 16	Baseline	Week 16
HSSDD	n	846	654	833	652	834	652
	Worst skin pain	0.25	0.38	0.52	0.56	0.55	0.62
	Average skin pain	0.26	0.35	0.51	0.54	0.54	0.60
	Smell or odour	0.30	0.41	0.45	0.44	0.52	0.53
	Itch at its worst	0.20	0.30	0.45	0.45	0.52	0.51
	Drainage or oozing	0.34	0.48	0.46	0.47	0.54	0.56
HSSQ	n	997	904	994	908	996	908
	Skin pain	0.25	0.37	0.59	0.64	0.64	0.70
	Smell or odour	0.30	0.42	0.48	0.54	0.57	0.62
	Itch	0.18	0.31	0.49	0.55	0.57	0.59
	Drainage or oozing	0.35	0.49	0.49	0.58	0.58	0.67

Orange text indicates a weak correlation ($r < 0.30$), blue text indicates a moderate correlation ($0.30 \le r < 0.70$), and green text indicates a strong correlation ($0.70 \le r < 0.90$); $p < 0.001$ for all

DLQI Dermatology Life Quality Instrument, *HiSQOL* Hidradenitis Suppurativa Quality of Life questionnaire, *HSSDD* Hidradenitis Suppurativa Symptom Daily Diary, *HSSQ* Hidradenitis Suppurativa Symptom Questionnaire, *IHS4* International HS Severity Score System

two variables as weak ($r<0.3$), moderate ($r\ge0.3$ to <0.7), strong ($r\ge0.7$ to <0.9) or very strong ($r\ge0.9$) [20].

The ability of HSSDD and HSSQ to distinguish between groups known to be clinically different (i.e. known-groups validity) was assessed by analysis of variance (ANOVA), comparing mean HSSDD and HSSQ symptom item scores among patient subgroups with different clinical status defined by Hurley stage, IHS4, HS PGA, PGI-S-SP and PGI-S-HS at baseline and Week 16.

Test-retest reliability was evaluated using intraclass correlation coefficients (ICC), calculated for each item scores for HSSDD and HSSQ using a two-way mixed effect ANOVA model with week as a fixed effect. Test-retest reliability analysis was conducted in the subgroup of stable patients defined for HSSDD as those with no change in PGI-S-SP score between baseline and Week 4 when assessing the worst and average skin pain items. When assessing the other symptom item scores, stable patients were defined for HSSDD as those with no change in PGI-S-HS score between baseline and Week 4. Stable patients were defined for HSSQ as those with no change in IHS4 level (mild, moderate, severe) between Week 32 and Week 36. The PGI-S-SP and PGI-S-HS were not used for HSSQ, as they were not assessed at Week 36.

Responsiveness was assessed by correlating changes from baseline to Week 16 in both HSSDD and HSSQ item scores with changes in PGI scales (for both measures this includes PGI-S-HS and PGI-S-SP, and PGI-C-HS and PGI-C-SP) within that same time interval, using Spearman's rank correlation coefficient. A threshold of 0.30 Spearman's rank correlation was considered to demonstrate acceptable sensitivity to change over time.

Interpretation Thresholds

Following FDA guidance, proposed thresholds for clinically meaningful within-patient change were determined and assessed by triangulating threshold estimates from anchor- and distribution-based analysis, with anchor-based results as primary and distribution-based results as supporting evidence [13–15]. In the anchor-based analyses, using PGI-S-HS or PGI-S-SP, as well as the PGI-C-HS and PGI-C-SP, patients were classified into response groups based on the level of change on the PGI scales. A two-level improvement for PGI-S-HS/PGI-S-SP defined a patient as 'much better'. Descriptive statistics

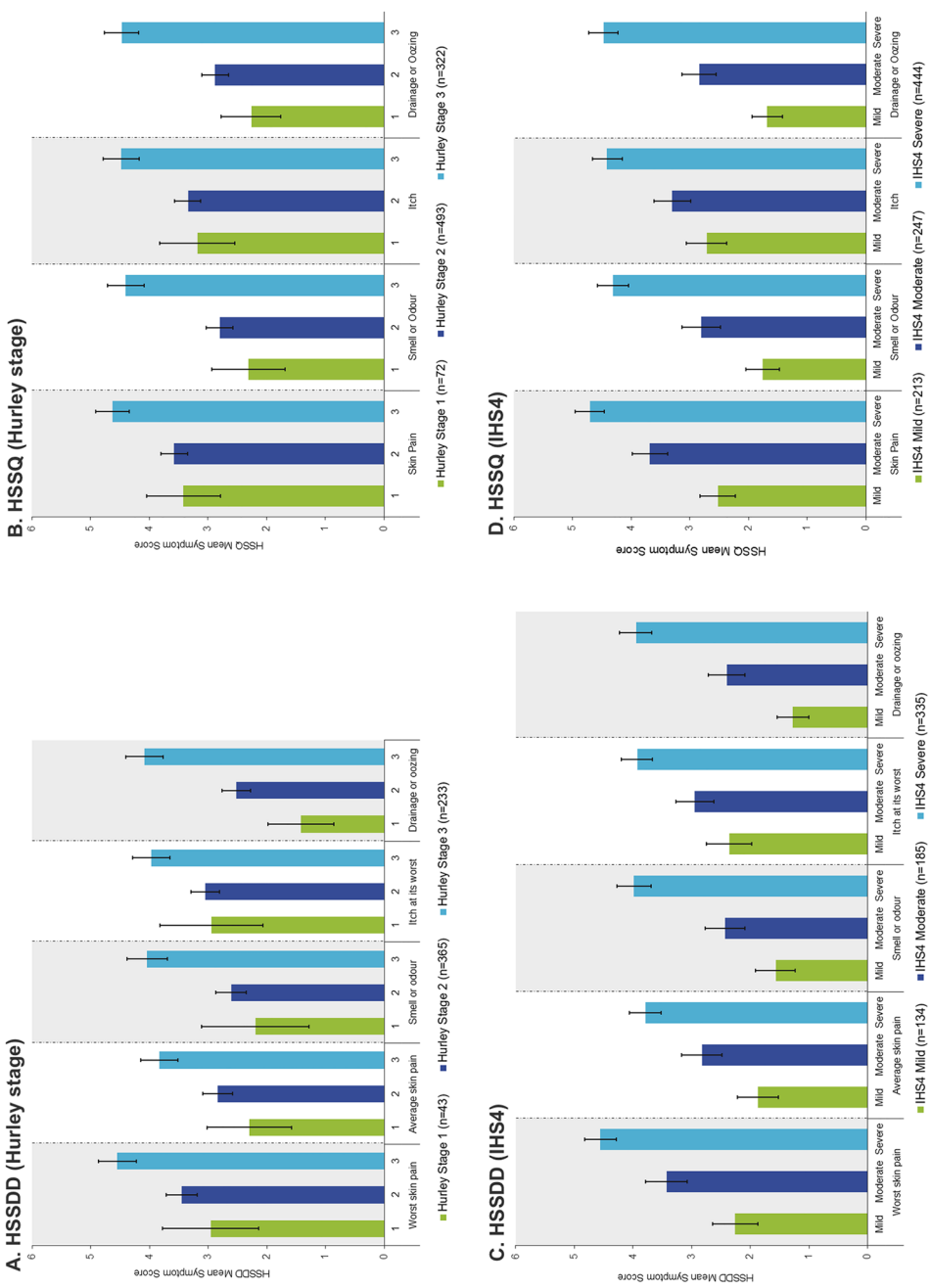


Fig. 1 HSSDD and HSSQ symptom item scores in HS subgroups as defined by Hurley stage (**A**, **B**) and IHS4 (**C**, **D**) at Week 16 (known-groups validity). Error bars show 95% confidence limits. *P* values estimated using the Kruskal-Wallis test were all $p < 0.001$ at Week 16. *HS* hidradenitis suppurativa, *HSSDD* Hidradenitis Suppurativa Symptom Daily Diary, *HSSQ* Hidradenitis Suppurativa Symptom Questionnaire, *IHS4* International Hidradenitis Suppurativa Severity Score System

Table 3 Test-retest reliability for HSSDD total score between baseline and Week 4 and HSSQ total score between Week 32 and Week 36

		ICC
HSSDD (Baseline vs Week 4)[a]	*n*	934
	Worst skin pain	0.83
	Average skin pain	0.84
	Smell or odour	0.85
	Itch at its worst	0.80
	Drainage or oozing	0.84
HSSQ (Week 32 vs Week 36)[b]	n	1007
	Skin pain	0.73
	Smell or odour	0.82
	Itch	0.80
	Drainage or oozing	0.76

[a]$n = 260$ for change from baseline to Week 4; [b]$n = 475$ for change from Week 32 to Week 36

HSSDD Hidradenitis Suppurativa Symptom Daily Diary, *HSSQ* Hidradenitis Suppurativa Symptom Questionnaire, *ICC* intraclass correlation coefficients

were calculated for each item score changes from baseline to Week 16 within these response groups. Effect sizes were calculated as the mean change from baseline to Week 16 divided by the overall baseline standard deviation (SD). Empirical cumulative distribution function (eCDF) and probability density function (PDF) curves of changes in item scores from baseline to Week 16 were plotted separately for each response group within each of the selected anchors to further guide the selection of the thresholds. Supportive distribution-based analyses (one standard error of measurement and half of the baseline SD) were also conducted.

Cut-off thresholds for different levels of severity were derived for each symptom item scores using the severity levels for the PGI-S-HS or PGI-S-SP (none, mild, moderate, severe or very severe) as anchors. Four separate receiver-operating characteristic (ROC) analyses per target item were employed to determine severity cut-off thresholds and meaningful score categories for the symptom item using the PGI-S-HS as an anchor (PGI-S-SP for pain items). The optimal cut-off threshold for a given severity level was estimated from the highest Youden Index of the ROC curve, using data pooled across all available visits between baseline and Week 48.

Ethical Approval

The study protocol, amendments and patient informed consent were reviewed by a national, regional or Independent Ethics Committee (IEC) or Institutional Review Board (IRB). This study was conducted in accordance with the current version of the applicable regulatory and International Conference on Harmonisation (ICH)-Good Clinical Practice (GCP) requirements, the ethical principles that have their origin in the principles of the Declaration of Helsinki, and the local laws of the countries involved.

Role of the Funding Source

UCB contributed to study design, participated in data collection, completed the data analysis and participated in data interpretation. UCB participated in writing, review and approval of the manuscript. All authors had full access to the data, reviewed and approved the final version, and were responsible for the decision to submit for publication. A medical writing agency, employed by UCB, assisted with manuscript preparation under the authors' direction.

Table 4 Correlation between changes in HSSDD and HSSQ item scores and PGI scales from baseline to Week 16 (responsiveness)

		External Anchors			
Spearman s correlation, r		**PGI-S-SP**	**PGI-C-SP**	**PGI-S-HS**	**PGI-C-HS**
HSSDD	**n**	603	613	602	613
	Worst skin pain	0.63	0.51	0.57	0.50
	Average skin pain	0.62	0.48	0.58	0.46
	Smell or odour	0.44	0.37	0.40	0.37
	Itch at its worst	0.36	0.34	0.36	0.36
	Drainage or oozing	0.50	0.40	0.47	0.41
HSSQ	**n**	896	898	894	898
	Skin pain	0.73	0.53	0.68	0.52
	Smell or odour	0.46	0.39	0.43	0.39
	Itch	0.43	0.40	0.43	0.40
	Drainage or oozing	0.52	0.42	0.51	0.43

Orange text indicates a weak correlation ($r < 0.30$), blue text indicates a moderate correlation ($0.30 \le r < 0.70$), and green text indicates a strong correlation ($0.70 \le r < 0.90$); $p < 0.001$ for all

HSSDD Hidradenitis Suppurativa Symptom Daily Diary, *HSSQ* Hidradenitis Suppurativa Symptom Questionnaire, *PGI-S-HS/SP* Patient Global Impression of HS Severity/Skin Pain, *PGI-C-HS/SP* Patient Global Impression of Change in HS/Skin Pain

RESULTS

Patient Disposition and Baseline Characteristics

Baseline characteristics were taken from 1010 patients enrolled in the two phase 3 trials. In total, 934 patients were included in the HSSDD analysis set and 1007 patients were included in the HSSQ analysis set. Mean (SD) age was 37.1 (12.2) years in the HSSDD analysis set and 36.7 (12.2) years in the HSSQ analysis set at baseline, with most patients being female (56.2% and 56.5%) and White (77.3% and 76.6%; Table 1). At baseline, 59.4% and 59.2% of patients had a BMI of ≥30 kg/m^2 (considered obese) [21]. In the HSSDD and HSSQ analysis sets, patients had a mean (SD) HS disease duration of 8.0 (7.8) years and 7.9 (7.8) years, respectively. Patients categorised as Hurley stage II were 56.1% and 55.7%; Hurley stage III were 43.8% and 44.3% (Table 1).

At baseline, HSSDD and HSSQ completion rates were 90.7% (n = 847/934) and 99.0% (n = 997/1,007), respectively. At Week 16, the completion rates were 70.1% (n = 655/934) and 90.2% (n = 908/1,007), respectively. For HSSDD, the baseline mean (SD) item scores for worst skin pain and average skin pain were 5.5 (2.5) and 4.8 (2.5). For HSSQ, the mean (SD) skin pain item score was 5.8 (2.4). The mean (SD) scores for the remaining HSSDD and HSSQ items are given in Table 1.

Floor and ceiling effects were minimal at baseline for the HSSDD item scores, with percentage of study patients with a score of 0 ranging from 1.5% (worst skin pain) to 8.4% (smell or odour) and percentage of study patients with a score of 10 ranging from 1.9% (average skin pain and itch at its worst) to 2.7% (drainage or oozing). Similarly, floor and ceiling effects were minimal at baseline for the HSSQ item scores, with percentage of study patients with a symptom item score of 0 ranging from 1.6% (skin pain) to

Table 5 Anchor-based clinically meaningful within-patient change threshold—observed changes from baseline at Week 16 by PGI-S one- and two-level improvement

Median change from baseline to Week 16 (*n*)		PGI-S-SP/PGI-S-HS[a] improvement from baseline to Week 16	
		One-level improvement	Two-level improvement
HSSDD	Worst skin pain	−2.17 (223)	−3.86 (91)
	Average skin pain	−2.00 (223)	−3.29 (91)
	Smell or odour	−1.11 (231)	−2.00 (97)
	Itch at its worst	−1.29 (231)	−2.00 (97)
	Drainage or oozing	−1.43 (231)	−2.71 (97)
HSSQ	Skin pain	−2.00 (329)	−5.00 (149)
	Smell or odour	−1.00 (336)	−3.00 (143)
	Itch	−1.00 (336)	−3.00 (143)
	Drainage or oozing	−2.00 (336)	−4.00 (143)

[a]PGI-S-SP used for pain items; PGI-S-HS used for all other items

HSSDD Hidradenitis Suppurativa Symptom Daily Diary, *HSSQ* Hidradenitis Suppurativa Symptom Questionnaire, *PGI-S-HS/SP* Patient Global Impression of HS Severity/Skin Pain, *PGI-C-HS/SP* Patient Global Impression of Change in HS/Skin Pain

9.7% (smell or odour) and percentage of study patients with a score of 10 ranging from 4.4% (smell or odour) to 5.7% (drainage and oozing).

Confirmation of the Scoring Rule of the HSSDD

The simulation analysis results showed that for each HSSDD symptom item, across all pre-specified assessment timepoints and overall, the standard deviation of the weekly symptom item score increased as the maximum allowed number of missing days within a given week increased. Still, there were no significant differences in standard deviations of weekly symptom item scores observed when comparing the current scoring rule with the zero-day missing scenario across different timepoints (baseline, Weeks 2, 4, 6, 8, 10, 12, 14 and 16). Thus, it was confirmed that the current weekly scoring rule (≥ 4 out of 7 daily scores non-missing for a given week) was appropriate. A less stringent rule (e.g. ≥ 3 out of 7 daily scores non-missing for a given week) may even be employed and would be unlikely to impact results (Supplementary Table 1).

Construct Validity

At baseline all convergent correlations between HSSDD/HSSQ and outcome measures were positive for both HSSDD and HSSQ, with moderate correlations observed between DLQI total score and all HSSDD and HSSQ symptom item scores (Table 2). For HSSDD, a weak correlation was observed between IHS4 scores and the worst skin pain ($r = 0.25$), average skin pain ($r = 0.26$) and itch at its worst ($r = 0.20$) items. For HSSQ, a weak correlation was observed between IHS4 scores and the skin pain ($r = 0.25$) and itch at its worst ($r = 0.18$) items. All correlations were moderate between HiSQOL total score and HSSDD and HSSQ symptom item scores. Correlations with other measures were generally stronger at Week 16 than at baseline for HSSDD and HSSQ item scores (Table 2).

At Week 16, both HSSDD and HSSQ symptom items were able to discriminate between HS subgroups as defined by Hurley stage, demonstrating good known-groups validity (Fig. 1A, B). As Hurley stage increased (indicating higher disease severity), a corresponding increase in mean symptom scores across all HSSDD and HSSQ items was observed at Week 16 (Fig. 1A, B). Similar findings were found using other anchors to define patient subgroups. Mean HSSDD and HSSQ symptom item scores in groups with higher severity according to IHS4 (Fig. 1C, D), PGI-S-HS (Supplementary Fig. 1A–B) and HS PGA (Supplementary Fig. 1C–D) were generally higher than in groups with lower severity.

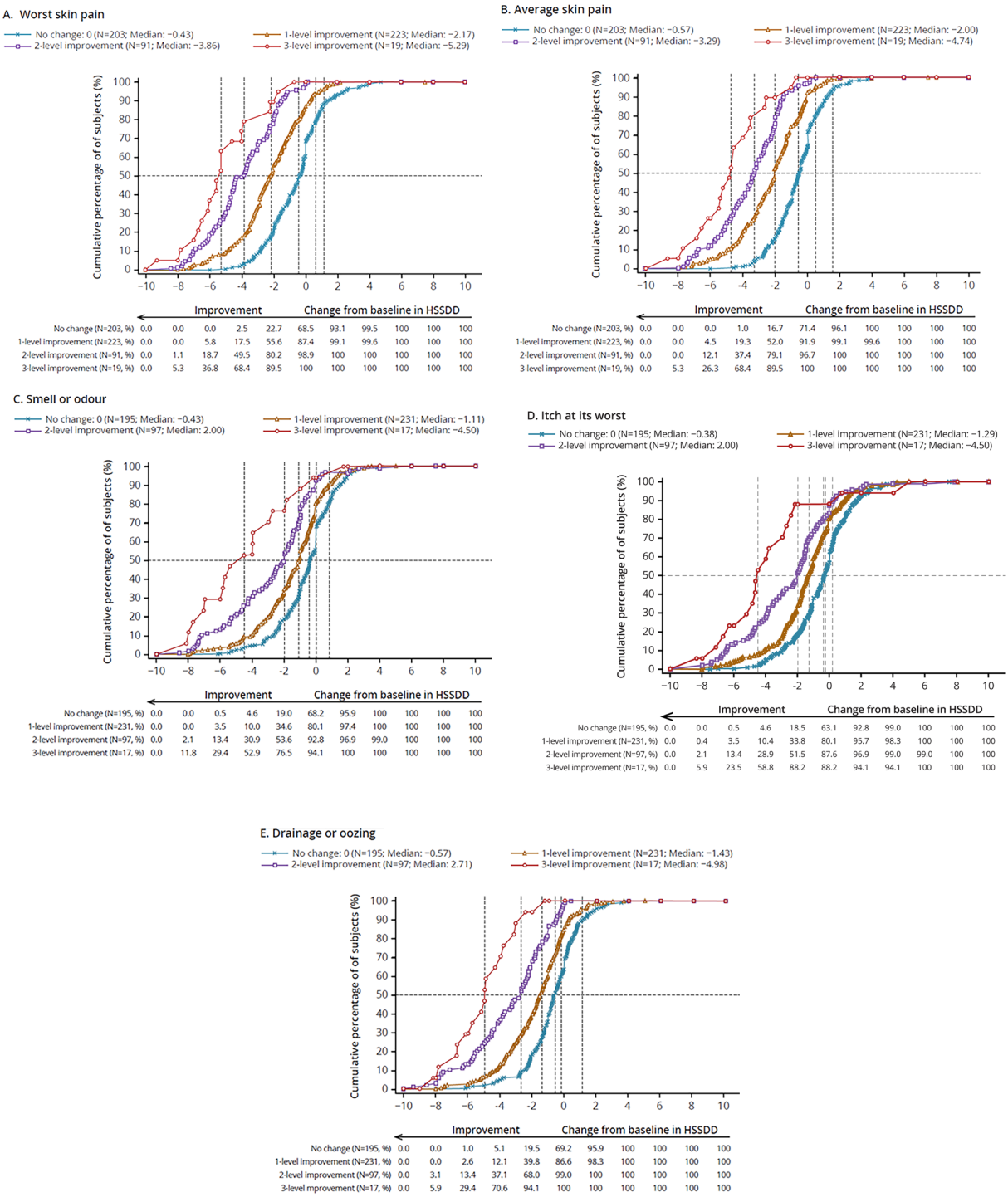


Fig. 2 Empirical cumulative distribution function curves of changes from baseline to Week 16 on the HSSDD: worst skin pain (**A**), average skin pain (**B**), smell or odour (**C**), itch at its worst (**D**) and drainage or oozing (**E**) items by change in levels of PGI-S-SP/HS response category (none, mild, moderate, severe, very severe; $N = 934$). PGI-S-SP was used as the anchor for the worst and average skin pain items; PGI-S-HS was used as the anchor for the other items. *eCDF* empirical cumulative distribution function, *HSSDD* Hidradenitis Suppurativa Symptom Daily Diary, *PGI-S-SP/HS* Patient Global Impression of Skin Pain/Hidradenitis Suppurativa Severity

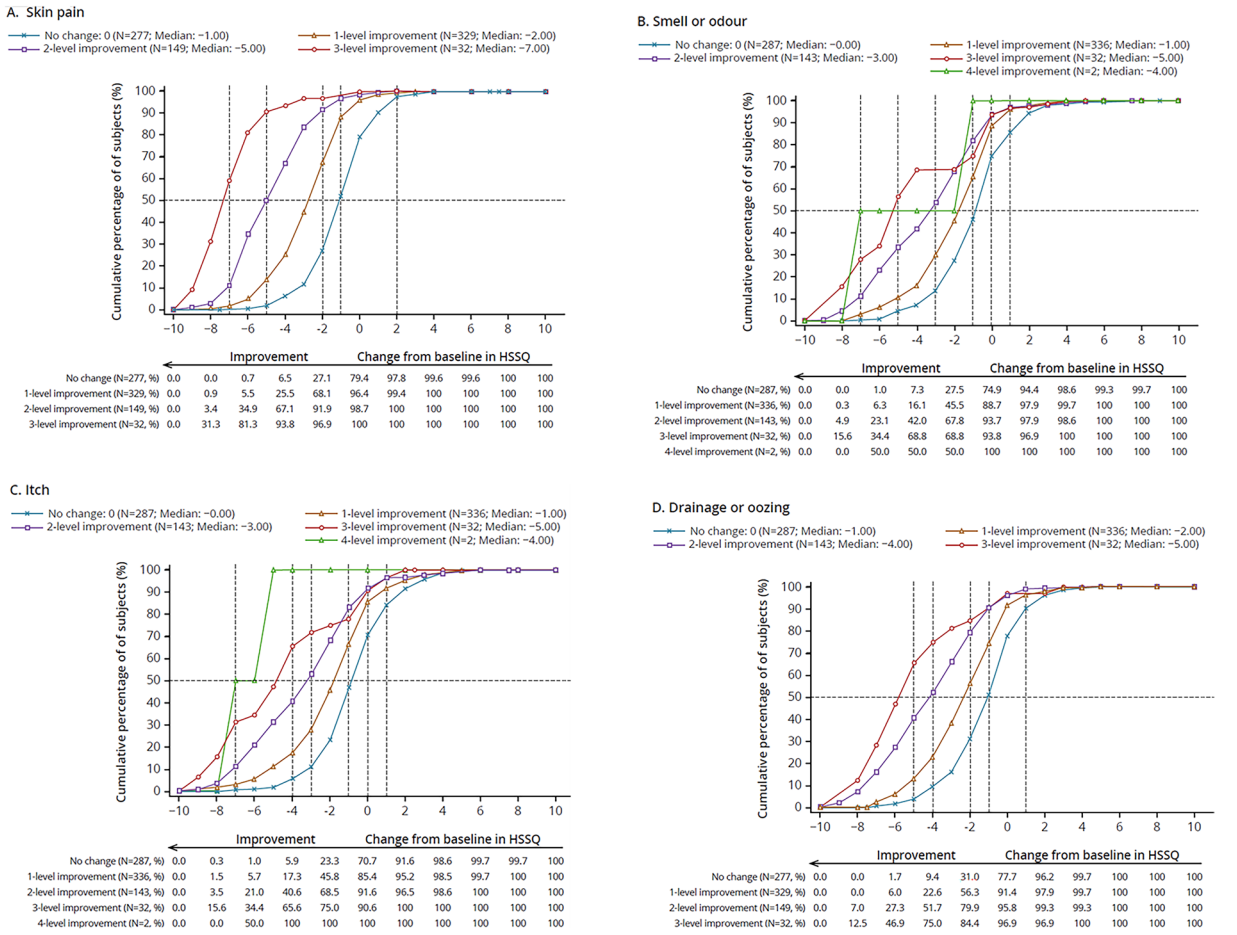


Fig. 3 Empirical cumulative distribution function curves of changes from baseline to Week 16 on the HSSQ: skin pain (**A**), smell or odour, (**B**) itch (**C**) and drainage or oozing (**D**) items by change in levels of PGI-S-SP/HS response category (none, mild, moderate, severe, very severe; N = 1007). PGI-S-SP was used as the anchor for skin pain item; PGI-S-HS was used as the anchor for the other items. *eCDF* empirical cumulative distribution function, *HSSQ* Hidradenitis Suppurativa Symptom Questionnaire, *PGI-S-SP/HS* Patient Global Impression of Skin Pain/Hidradenitis Suppurativa Severity

Baseline results were very consistent, although it should be noted that less variation in the anchors was observed at baseline as expected due to inclusion and exclusion criteria (data not shown).

Test-Retest Reliability

When test-retest reliability was assessed, observed ICC values were 0.80–0.85 between baseline and Week 4 across HSSDD items, and 0.73–0.82 between Week 32 and Week 36 across HSSQ items (Table 3). All scores showed acceptable test-retest reliability (ICC ≥ 0.70) [22].

Responsiveness

All correlation coefficients between changes in HSSDD and HSSQ item scores and changes in PGI-S-SP, PGI-S-HS, PGI-C-SP and PGI-C-HS were positive, as expected, with all p values < 0.001, and exceeded the threshold of 0.30 to demonstrate acceptable sensitivity (Table 4) [23]. Additionally, many correlation coefficients, particularly with skin pain items (PGI-S-SP and HSSQ skin pain), exceeded 0.50, indicating moderate correlations.

Table 6 Severity thresholds for HSSDD and HSSQ

	None	Mild	Moderate	Severe	Very severe
HSSDD					
Worst skin pain	< 1.67	≥ 1.67 to < 4.00	≥ 4.00 to < 5.50	≥ 5.50 to < 7.17	≥ 7.17
Average skin pain	< 1.40	≥ 1.40 to < 3.29	≥ 3.29 to < 4.50	≥ 4.50 to < 5.67	≥ 5.67
Smell or odour	< 1.00	≥ 1.00 to < 2.83	≥ 2.83 to < 4.50	4.50 to < 5.25	≥ 5.25
Itch at its worst	< 2.25	≥ 2.25 to < 3.40	≥ 3.40 to < 4.40	≥ 4.40 to < 4.57	≥ 4.57
Drainage or oozing	< 1.00	≥ 1.00 to < 3.00	≥ 3.00 to < 4.71	≥ 4.71 to < 6.00	≥ 6.00
HSSQ					
Skin pain	< 1.00	≥ 1.00 to < 3.00	≥ 3.00 to < 6.00	≥ 6.00 to < 7.00	≥ 7.00
Smell or odour	< 1.00	≥ 1.00 to < 2.00	≥ 2.00 to < 5.00	≥ 5.00 to < 6.00	≥ 6.00
Itch	< 1.00	≥ 1.00 to < 3.00	≥ 3.00 to < 4.00	≥ 4.00 to < 5.00	≥ 5.00
Drainage or oozing	< 1.00	≥ 1.00 to < 3.00	≥ 3.00 to < 4.00	≥ 4.00 to < 7.00	≥ 7.00

HSSDD Hidradenitis Suppurativa Symptom Daily Diary, *HSSQ* Hidradenitis Suppurativa Symptom Questionnaire

Interpretation Thresholds

Change in HSSDD and HSSQ symptom item scores for patients with a one- and two-level improvement on the PGI-S-HS or PGI-S-SP from baseline to Week 16 are shown in Table 5. A two-level improvement from baseline to Week 16 on the PGI-S-HS/PGI-S-SP was considered to represent a clinically meaningful within-patient improvement. Some study participants with a one-level improvement on the PGI-S-HS/PGI-S-SP, particularly those who responded very severe at baseline, reported at least some level of worsening on the HSSDD/HSSQ items, implying that one level of improvement on the PGI-S-HS/PGI-S-SP is not clinically meaningful.

eCDF curves supported the use of estimates from the group with two levels of improvement on the PGI-S-HS or PGI-S-SP as an anchor to derive the thresholds due to the larger degree of separation between no change and two levels of improvement compared to no change and one level (Figs. 2 [HSSDD] and 3 [HSSQ]).

PGI-C-HS/PGI-C-SP, considered as a supportive anchor only as it requires a patient to recall their status from baseline, provided further evidence on the selection of the thresholds (data not shown).

Triangulation of the various estimates from the anchor-based approaches indicated that the clinically meaningful within-patient improvement thresholds should be a 3- to 4-point decrease for worst and average skin pain item scores and a 2- to 3-point decrease for smell or odour, itch at its worst, and draining or oozing item scores. In HSSQ, a 4- to 5-point decrease for the skin pain, 3- to 4-point decrease for smell or odour and draining or oozing item scores and 3-point decrease for itch item scores was identified as clinically meaningful within-patient improvement thresholds.

Severity thresholds were identified for both HSSDD and HSSQ item scores from ROC analyses using PGI-S-SP and PGI-S-HS as anchors. Identified cut-off values for disease severity of none, mild, moderate, severe and very severe are presented in Table 6.

DISCUSSION

These analyses aimed to assess the psychometric properties and derive interpretation thresholds of the HSSDD and HSSQ, two HS symptom measures, using pooled, blinded data from two phase 3 trials (BE HEARD I and II) evaluating the efficacy and safety of bimekizumab in patients with moderate to severe HS. While high completion rates were observed for both HSSDD and HSSQ, higher rates were observed for HSSQ compared with HSSDD, reflecting the difference between on-site (HSSQ) and at-home (HSSDD) administration of the measures.

Overall, both HSSDD and HSSQ symptom item scores demonstrated good reliability, validity and responsiveness in a sample of patients with moderate to severe HS. Both HSSDD and HSSQ were shown to have strong construct validity. All convergent validity correlations were in the pre-specified direction and strength at baseline and Week 16 for IHS4 (r values at Week 16: HSSDD, 0.30–0.48; HSSQ, 0.31–0.49) and DLQI total score (r values at Week 16: HSSDD, 0.44–0.56; HSSQ, 0.54–0.64). The slightly stronger correlations observed for DLQI (patient-assessed) compared with the IHS4 (investigator-assessed) underscore the discrepancies between clinician-reported outcomes and PROs that have been reported in the literature across disease areas [24–27].

Known-groups validity assessment found HSSDD and HSSQ item scores discriminated successfully between subgroups as defined by Hurley stage, PGI-S-HS, HS PGA and IHS4 measures at baseline and Week 16. Mean HSSDD and HSSQ item scores in groups with more severe Hurley stage, PGI-S-HS, IHS4 and HS PGA were generally higher.

Test-retest reliability analyses demonstrated good item score reproducibility for both HSSDD and HSSQ item scores with ICC values in all cases exceeding the pre-specified threshold of acceptability (0.70). Additionally, changes from baseline to Week 16 in HSSDD and HSSQ item scores and PGI scale anchors were moderately to strongly correlated (>0.30), establishing satisfactory responsiveness of both measures.

This analysis defined clinically meaningful within-patient improvement thresholds using anchor-based analyses. The findings from the eCDF curves supported the use of estimates from the group with two levels of improvement on the PGI-S. For HSSDD, a 3- to 4-point decrease was identified as clinically meaningful within-patient improvement thresholds for worst and average skin pain item scores, 2- to 3-point decrease for smell or odour, itch at its worst and draining or oozing item scores. For HSSQ, a 4- to 5-point decrease was identified as clinically meaningful within-patient improvement thresholds for skin pain item score; 3- to 4-point decrease for smell or odour and draining or oozing item scores; and 3-point decrease for itch. Differences in thresholds capturing similar concepts across the two measures may be the results of slight variations in the recall periods (24 h versus 7 days) and in the concepts (worst skin pain/average skin pain vs skin pain; itch at its worst vs itch) in HSSDD and HSSQ, respectively.

The HiSTORIC initiative recently released their recommendation on the suitable use of the HiSQOL to assess patients' HRQoL in routine clinical practice [28]. The HiSQOL also has a symptoms domain and captures the impact of each symptom (pain, itch, drainage, odour) on patients' HRQoL. On the other hand, the HSSDD and HSSQ measure the level of severity of those four symptoms. HiSQOL and both the HSSDD and HSSQ, which are reliable and valid fit-for-purpose PRO measures, should thus be seen as complementary measures addressing multiple domains of the core outcomes set: HS-specific QOL, pain and symptoms.

Limitations

Validity assessments in this study were limited to the global regions involved in the BE HEARD studies; therefore, cultural validity could not be assessed. As the BE HEARD I and II studies were phase 3 trials which primarily aimed to evaluate the efficacy and safety of bimekizumab, the studies were not designed to assess the psychometric validity of HSSDD or HSSQ or to derive thresholds to define clinically meaningful

improvements. Furthermore, the BE HEARD studies enrolled patients with moderate to severe HS; psychometric assessments at Week 16 included patients who had reached a milder disease severity as measured by Hurley stage, to some extent providing evidence that the measures performed well across the disease severity spectrum. Further research into milder disease is needed to confirm the performance of the two symptom measures across the full spectrum of HS severity.

CONCLUSION

Using a substantial, representative sample of patients with moderate to severe HS in the phase 3 BE HEARD I and II trials, this study demonstrated that both HSSDD and HSSQ are fit-for-purpose PRO measures. Furthermore, this study supports the use of these measures to assess the impact of treatment interventions on key HS symptoms and inform physician treatment decisions in the management of patients with moderate to severe HS.

ACKNOWLEDGEMENTS

The authors thank the patients, the investigators and their teams who took part in this study. The authors also acknowledge Susanne Wiegratz, MSc, UCB, Monheim, Germany, for publication coordination and Sana Yaar, PhD, from Costello Medical, UK, for medical writing and editorial assistance based on the authors' input and direction. This study was funded by UCB. All the results presented in this article are in aggregate form, and no personally identifiable information was used for this study.

Medical writing/editorial assistance. Sana Yaar, PhD, from Costello Medical, UK, performed medical writing and editorial assistance based on the authors' input and direction.

Authors' contributions. Substantial contributions to study conception and design: John R. Ingram, Jérémy Lambert, Valerie Ciaravino, Robert Rolleri, Ingrid Pansar, Luke Peterson, Christopher G. Pelligra, Linnea Thorlacius; substantial contributions to analysis and interpretation of the data: John R. Ingram, Jérémy Lambert, Valerie Ciaravino, Robert Rolleri, Ingrid Pansar, Luke Peterson, Christopher G. Pelligra, Linnea Thorlacius; drafting the article or revising it critically for important intellectual content: John R. Ingram, Jérémy Lambert, Valerie Ciaravino, Robert Rolleri, Ingrid Pansar, Luke Peterson, Christopher G. Pelligra, Linnea Thorlacius; final approval of the version of the article to be published: John R. Ingram, Jérémy Lambert, Valerie Ciaravino, Robert Rolleri, Ingrid Pansar, Luke Peterson, Christopher G. Pelligra, Linnea Thorlacius.

Funding. This study was sponsored by UCB. Support for third-party writing assistance for this article, provided by Sana Yaar, PhD, Costello Medical, UK, was funded by UCB in accordance with Good Publication Practice (GPP) 2022 guidelines (https://www.ismpp.org/gpp-2022). Funding for the journal's Rapid Service Fee and Open Access Fees was provided by UCB.

Data availability. Underlying data from this manuscript may be requested by qualified researchers 6 months after product approval in the US and/or Europe, or global development is discontinued, and 18 months after trial completion. Investigators may request access to anonymised individual patient-level data and redacted trial documents which may include: analysis-ready datasets, study protocol, annotated case report form, statistical analysis plan, dataset specifications and clinical study report. Prior to use of the data, proposals need to be approved by an independent review panel at www.Vivli.org and a signed data sharing agreement will need to be executed. All documents are available in English only, for a pre-specified time, typically 12 months, on a password protected portal.

Declarations

Conflicts of interest. John R Ingram: Receives a stipend as Editor-in-Chief of the

British Journal of Dermatology and an authorship honorarium from UpToDate; consultant for AbbVie, Boehringer Ingelheim, ChemoCentryx, Citryll, MoonLake, Novartis, UCB, and Union Therapeutics, and has served on advisory boards for Insmed, Kymera Therapeutics and Viela Bio; co-copyright holder of HiSQOL©, patient global assessment, and HS-IGA; his department receives income from copyright of the Dermatology Life Quality Instrument (DLQI) and related instruments. Jérémy Lambert, Valerie Ciaravino, Robert Rolleri, Ingrid Pansar, Luke Peterson: Employees and shareholders of UCB. Christopher G Pelligra: Employee of Evidera, a part of ThermoFisher Scientific that receives funding for research from UCB. Linnea Thorlacius: Received speaker honoraria from UCB and is co-copyright holder of HiSQOL© and HS-IGA.

Ethical approval. The study protocol, amendments, and patient informed consent were reviewed by a national, regional, or Independent Ethics Committee (IEC) or Institutional Review Board (IRB). This study was conducted in accordance with the current version of the applicable regulatory and International Conference on Harmonisation (ICH)-Good Clinical Practice (GCP) requirements, the ethical principles that have their origin in the principles of the Declaration of Helsinki, and the local laws of the countries involved.

REFERENCES

1. Zouboulis CC, Del Marmol V, Mrowietz U, et al. Hidradenitis suppurativa/acne inversa: criteria for diagnosis, severity assessment. Classif Dis Eval Dermatol. 2015;231:184–90.
2. Sabat R, Jemec GBE, Matusiak Ł, et al. Hidradenitis suppurativa. Nat Rev Dis Primers. 2020;6:18.
3. Matusiak Ł. Profound consequences of hidradenitis suppurativa: a review. Br J Dermatol. 2020;183:e171–7.
4. Garg A, Neuren E, Cha D, et al. Evaluating patients' unmet needs in hidradenitis suppurativa: results from the global survey of impact and healthcare needs (VOICE) project. J Am Acad Dermatol. 2020;82:366–76.
5. Patel ZS, Hoffman LK, Buse DC, et al. Pain, psychological comorbidities, disability, and impaired quality of life in hidradenitis suppurativa [corrected]. Curr Pain Headache Rep. 2017;21:49.
6. Wright S, Strunk A, Garg A. New-onset depression among children, adolescents, and adults with hidradenitis suppurativa. J Am Acad Dermatol. 2020;83:1360–6.
7. Wright S, Strunk A, Garg A. Prevalence of depression among children, adolescents, and adults with hidradenitis suppurativa. J Am Acad Dermatol. 2022;86:55–60.
8. Orenstein LAV, Salame N, Siira MR, et al. Pain experiences among those living with hidradenitis suppurativa: a qualitative study. Br J Dermatol. 2023;188:41–51.
9. Molina-Leyva A, Cuenca-Barrales C. Pruritus and malodour in patients with hidradenitis suppurativa: impact on quality of life and clinical features associated with symptom severity. Dermatology. 2020;236:59–65.
10. Thorlacius L, Ingram JR, Villumsen B, et al. A core domain set for hidradenitis suppurativa trial outcomes: an international Delphi process. Br J Dermatol. 2018;179:642–50.
11. Finlay AY, Khan G. Dermatology life quality index (DLQI)—a simple practical measure for routine clinical use. Clin Exp Dermatol. 1994;19:210–6.

12. Ingram JR, Ciaravino V, Rolleri R, et al. Electronic patient-reported outcomes in hidradenitis suppurativa: content validity and usability of the electronic hidradenitis suppurativa symptom daily diary, hidradenitis suppurativa symptom questionnaire, and hidradenitis suppurativa quality of life questionnaire. Dermatology. 2024;240:65–76.

13. US Food and Drug Administration. Patient-focused drug development: Methods to identify what is important to patients. Volume 2023.

14. US Food and Drug Administration. Patient-focused drug development: Guidance 4—incorporating clinical outcome assessment into endpoint for regulatory decision making. Volume 2023.

15. US Food and Drug Administration. Patient-reported outcome measures: Use in medical product development to support labeling claims. Volume 2023.

16. ClinicalTrials.gov. A study to evaluate the efficacy and safety of bimekizumab in study participants with moderate to severe hidradenitis suppurativa (BE HEARD I). https://clinicaltrials.gov/NCT04242446. Accessed 12 Feb 2025.

17. ClinicalTrials.gov. A study to evaluate the efficacy and safety of bimekizumab in study participants with moderate to severe hidradenitis suppurativa (BE HEARD II). https://clinicaltrials.gov/NCT04242498. Accessed 12 Feb 2025.

18. Kimball AB, Jemec GBE, Sayed CJ, et al. Efficacy and safety of bimekizumab in patients with moderate-to-severe hidradenitis suppurativa (BE HEARD I and BE HEARD II): two 48-week, randomised, double-blind, placebo-controlled, multicentre phase 3 trials. Lancet. 2024;403:2504–19.

19. Kirby J, Thorlacius L, Villumsen B, et al. The hidradenitis suppurativa quality of life (HiSQOL) score: development and validation of a measure for clinical trials. Br J Dermatol. 2020;183:340–8.

20. Hinkle DE, Wiersma W, Jurs SG. Applied statistics for the behavioral sciences. 5th ed: Houghton Mifflin. 2003.

21. Centers for Disease Control and Prevention (CDC). Defining adult overweight & obesity. Volume 2023.

22. Hays RD, Farivar SS, Liu H. Approaches and recommendations for estimating minimally important differences for health-related quality of life measures. COPD. 2005;2:63–7.

23. Cappelleri JC, Zou KH, Bushmakin AG, et al. Patient-reported outcomes measurement, implementation and interpretation: Chapman and Hall/CRC. 2014.

24. Barbieri JS, Gelfand JM. Patient-reported outcome measures as complementary information to clinician-reported outcome measures in patients with psoriasis. JAMA Dermatol. 2021;157:1236–7.

25. Bily W, Jauker J, Nics H, et al. Associations between patient-reported and clinician-reported outcome measures in patients after traumatic injuries of the lower limb. Int J Environ Res Public Health. 2022;19:8.

26. Long V, Chen Z, Du R, et al. Understanding discordant perceptions of disease severity between physicians and patients with eczema and psoriasis using structural equation modeling. JAMA Dermatol. 2023;159:811–9.

27. Xiao C, Polomano R, Bruner DW. Comparison between patient-reported and clinician-observed symptoms in oncology. Cancer Nurs. 2013;36:E1–16.

28. Mastacouris N, Tannenbaum R, Strunk A, et al. Outcome measures for the evaluation of treatment response in hidradenitis suppurativa for clinical practice: a historic consensus statement. JAMA Dermatol. 2023;159:1258–66.