BJD
British Journal of Dermatology

Development and initial validation of the HS-IGA: a novel hidradenitis suppurativa-specific investigator global assessment for use in interventional trials*

Amit Garg[1], Carla Zema[2], Katherine Kim[3], Weihua Gao[3], Naijun Chen[3], Gregor B.E. Jemec[4], Joslyn Kirby[5], Linnea Thorlacius[4], Bente Villumsen[6] and John R. Ingram[7]

[1]Donald and Barbara Zucker School of Medicine at Hofstra/Northwell, 1991 Marcus Avenue, Suite 300, New Hyde Park, NY, 11042, USA
[2]Zema Consulting, Madison, AL, 35757, USA
[3]AbbVie, Inc., 1 North Waukegan St, North Chicago, IL, 60064, USA
[4]Department of Dermatology, Zealand University Hospital, Sygehusvej 10, Roskilde, DK-4000, Denmark
[5]Penn State Health,500 University Drive, Hershey, PA, 17033, USA
[6]Patients' Association HS Denmark, Copenhagen, Denmark
[7]Division of Infection and Immunity, Cardiff University, Cardiff, UK

Abstract

Background Few validated instruments exist for use in hidradenitis suppurativa (HS) trials.
Objectives To develop a novel HS Investigator Global Assessment (HS-IGA) and to validate its psychometric properties.
Methods Development of HS-IGA involved discussion among stakeholders, including patients, within HISTORIC. Data from replicate phase III randomized controlled trials evaluating HS treatment were utilized. Multivariate models identified lesion type and body region as variables of importance. Classification and regression trees for ordinal responses were built. Validation included assessment of test–retest reliability, predictive validity, responsiveness and clinical meaningfulness.
Results There were 3024 unique measurements available in PIONEER I. Mean and median lesion counts by region were largely <10 and were highest in axillary and inguinal regions. The mean and median number of regions involved were ≤ 3 for individual lesions and combinations. Regardless of lesion type, axillary and inguinal regions most influenced the HS-IGA score. Accordingly, regions were combined into a six-point IGA based on the maximum lesion number in either upper or lower body regions with a score of 0 (0–1 lesions), 1 (2–5), 2 (6–10), 3 (11–15), 4 (16–20) and 5 (≥ 20 lesions). The intraclass correlation coefficient for test–retest reliability was 0·91 (95% confidence interval 0·87–0·94). Spearman's rank order correlations (SROCs) with HS-PGA and Hidradenitis Suppurativa Clinical Response (HiSCR) were 0·73 and 0·51, respectively ($P < 0·001$ for both comparisons). SROCs with Dermatology Life Quality Index (DLQI), pain numerical rating scale and HS-QoL were 0·42, 0·34 and −0·25, respectively ($P < 0·001$ for all comparisons). HS-IGA was responsive at weeks 12 and 36. Predictive convergent validity was very good with HS-PGA (area under the curve = 0·89) and with HiSCR (area under the curve = 0·82). Predictive divergent validity was low with DLQI and HS-QoL.
Conclusions HS-IGA has moderate-to-strong psychometric properties and is simple to calculate.

Correspondence
Amit Garg.
Email: amgarg@northwell.edu

Accepted for publication
3 March 2022

Funding sources
Financial support was provided by AbbVie, Inc. AbbVie participated in the interpretation of data, and review and approval of the manuscript. All authors contributed to the development of the manuscript and maintained control over the final content. No honoraria or payments were made for authorship.

Conflicts of interest
Conflicts of interest statements can be found in Appendix 1.

Data availability
The data that support the findings of this study are available from AbbVie. Restrictions apply to the availability of these data, which were used under licence for this study. Data are available from the authors with the permission of AbbVie.

Ethics statement
This study involved secondary analysis of deidentified clinical trial data and was exempt from IRB review.

*Plain language summary available online

DOI 10.1111/bjd.21236

Hidradenitis suppurativa (HS), also known as acne inversa, is an inflammatory disease arising from the follicular unit. It is a potentially debilitating disease that, in North America and Europe, disproportionately affects women and African American people.[1,2] HS is known to have substantial impact on general health-related and skin-specific quality of life (QoL).[3]

Nine in 10 patients describe recent pain associated with their disease, while six in 10 rate this pain as moderate to worst possible.[3] HS has also been linked to a significant comorbidity burden[3,4] and increased mortality.[5,6] The morbidity is likely exacerbated by the 10-year diagnostic delay that patients experience on average.[3]

Nearly half of patients are dissatisfied with current treatments, most commonly due to perceived poor efficacy.[3] Accordingly, nearly half of patients express low optimism for having satisfactory control of symptoms in the near future.[3] At present, adalimumab is the only approved medication for the treatment of moderate-to-severe HS. While there is growing interest in drug development for HS, treatment represents the greatest unmet need for the disease. Relative lack of well-developed and validated measures of activity and response in HS may hinder further drug development.[7]

To address this fundamental gap, there is also a parallel international initiative to develop a core set of measures for trials in HS with the goal of improving truth, discrimination and feasibility in measurement of disease activity and treatment response, and to assess the comparative effectiveness of treatments. Hidradenitis Suppurativa Core Outcome Set Collaboration (HISTORIC) has established the core domain set (what to measure) in HS,[8] and it has highlighted significant challenges related to the question of 'how to measure' disease activity and responsiveness.[9] Global assessment represents one of the six core domains. However, there is no validated instrument for investigator global assessment (IGA) in HS. The purpose of this study was to develop and initially assess the psychometric properties of a novel IGA to measure disease severity and treatment response in HS.

Materials and methods

Development of the HS-IGA optimized psychometric measure properties based on several rounds of discussion from HS clinical experts, patient research partners, and outcomes measure development experts within HISTORIC. Development was carried out over more than 2 years in order to establish criteria for IGA development and to provide input into instrument development and validation.

Data from two phase III randomized controlled trials [PIONEER I (NCT01468207) and PIONEER II (NCT01468233)] evaluating HS treatment were used to develop and validate a draft conceptual framework of a simplified IGA. PIONEER I data were used for instrument development, while PIONEER II data were used for validation.

As part of the feasibility analysis in assessing all variables discussed among HISTORIC participants, lesion types, anatomical regions and combinations of these variables were included in the consideration of the initial HS-IGA framework. Lesion types included abscesses, fistulas, nodules, and all possible lesion combinations (i.e. A + F, A + N, F + N, A + F + N). Differentiation of inflammatory and noninflammatory nodules and of draining and nondraining fistulas was also examined. There were eight anatomical regions specified, namely axilla (left and right), inframammary (left and right), intermammary, buttock (left and right), inguinal (left and right), perianal, perineal and other. These eight regions were also combined to examine models with two, three or four regions (Figure 1).

Given the numerous combinations of types of lesions and regions to be included, the importance of the lesion type and region combinations was explored to determine whether any specific lesion type, region or combination influenced the potential instrument more than others (Figure 2) Multiple methods were used including standard error rate (ER), the ranked probability score (RPS), mean absolute error and mean square error. The R package 'party' was used for the computation of the importance measures ER and RPS.

Classification and regression trees for ordinal responses were used to develop the initial IGA scale. All possible combinations of lesion types, anatomical regions, and numbers of regions were explored in the models. The rpartScore R package was used to build classification trees for ordinal responses.[10] The weighed kappa was used to test the agreement between the IGA scale and scores for a physician-rated HS disease severity assessment known as the Hidradenitis Suppurativa Physician Global Assessment (HS-PGA).[11] Development was an iterative process that also included several rounds of HISTORIC workgroup input.

Validation of the IGA included assessment of test–retest reliability, construct validity, predictive validity, responsiveness and clinical meaningfulness to patients when assessing treatment effectiveness in controlling inflammatory signs and symptoms of HS, and, more specifically, HS lesions. The HS-

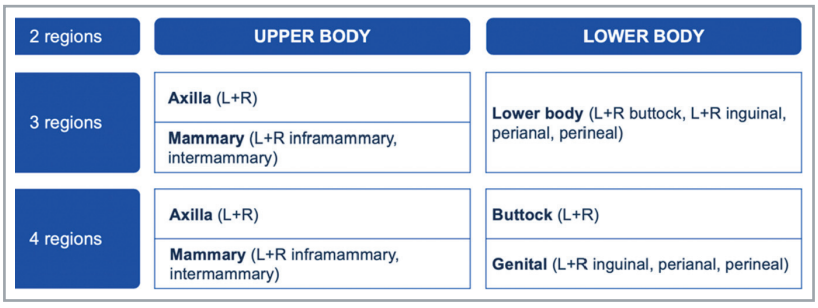

Figure 1 Anatomical regions and region combinations explored in development of the Hidradenitis Suppurativa Investigator Global Assessment. L, left; R, right.

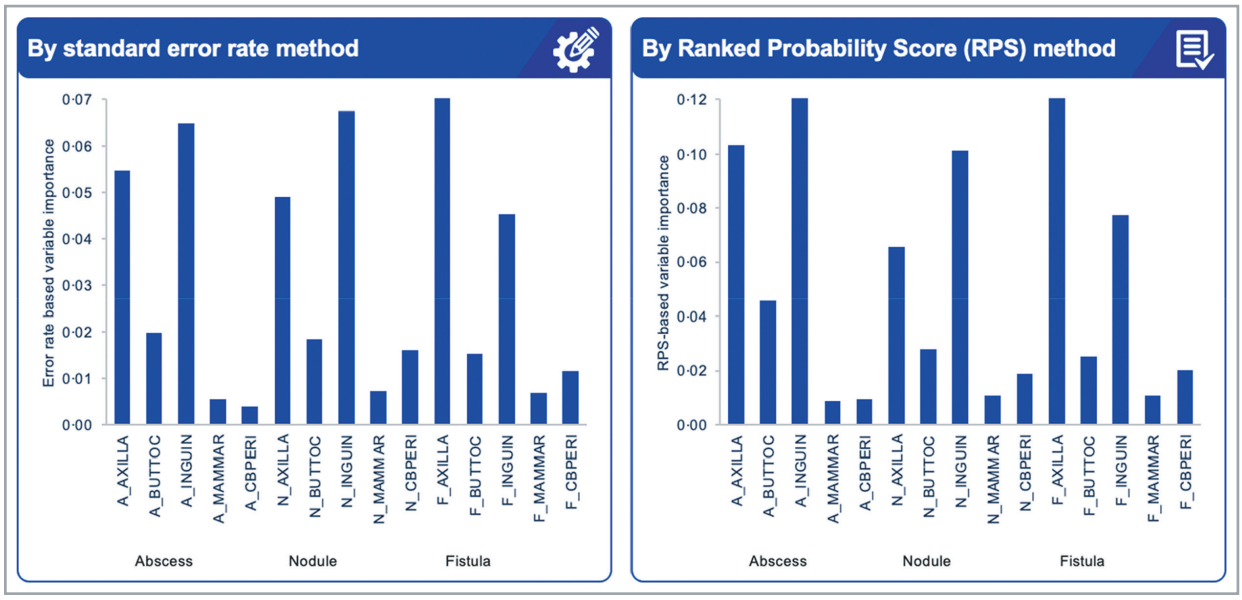

Figure 2 Type of lesion and region drivers of Hidradenitis Suppurativa Physician Global Assessment (HS-PGA). Multivariate regression models using the PIONEER I development dataset showing which lesion types and region combinations most influence HS-PGA, with consistent results using both (a) error rate and (b) ranked probability score.

PGA was used as a criteria measure for validation.[11] Responsiveness was assessed using the Hidradenitis Suppurativa Clinical Response (HiSCR).[12]

Test–retest reliability was assessed by examining the change in IGA categories, and correlating scores across appropriate study visits for patients who did not experience, or were not expected to experience, a change in their lesion counts during the intervening period (i.e. 'stable patients'). The IGA categories at baseline and the 2-week follow-up visit were used to assess test–retest reliability, with stable patients defined as those who reported 'no change' on the item between assessments. The intraclass correlation coefficient with a 95% confidence interval was calculated based on McGraw and Wong's one-way random effects ANOVA model.[13] A value of 0·70 or higher was considered to indicate strong reliability, while a value of 0·40–0·70 indicated moderate reliability.[14,15]

Construct validity was examined by evaluating convergence between the IGA and HiSCR, and between the IGA and HS-PGA. Spearman's rank-order correlations were conducted, with a value of 0·70 or higher considered to indicate strong correlation, and a value of 0·50–0·70 indicating moderate correlation.[16] Divergent validity was assessed using health-related QoL measures including the Dermatology Life Quality Index (DLQI), the Hidradenitis Suppurativa Quality of Life (HiSQoL) score, and a patient-reported pain measure on a numerical rating scale (pain NRS).

Responsiveness of the IGA was evaluated by examining the differences in change from baseline between HiSCR responders and nonresponders at weeks 12 and 36 (PIONEER II). The difference between the two groups' data distribution was quantified by Wilcoxon rank sum tests. Mean, SD, median and interquartile range were calculated, and we examined the cumulative distribution function (CDF) of responses (via empirical CDF curves) between treatment groups to characterize treatment effect and to examine the possibility that mean improvement reflected different responses in different patient subsets.

The predictive validity of the IGA was assessed using logistic regression to examine how well the IGA predicted outcomes: (i) HiSCR achievement (yes/no); (ii) being 'clear' or 'minimal' with HS lesions based on HS-PGA (vs. mild/moderate/severe/very severe); (iii) having 'no effect' or 'small effect' based on DLQI scores (vs. moderate/very large/extremely large); and (iv) 'best possible' quality of life based on HS-QoL scores (vs. worst possible). Baseline characteristics including age, sex, race, smoking status, body mass index and employment status were used in the regression models for predictive validity. The area under the receiver operating characteristic curve (AUC) was used to evaluate the performance of prediction, with a value ≥0·8 considered to be excellent or good, and a value of 0·7–0·8 considered to be fair.[16]

Finally, the patient-centredness of the measure was assessed by examining the differences in DLQI scores, HS-QoL and pain NRS between HS-IGA responders, as defined by ≥2-point improvement, and nonresponders. Statistical significance was determined by Wilcoxon rank sum test.

Results

Development

The baseline characteristics of patients participating in PIONEER I (n = 307) and PIONEER II (n = 326) used in the development and validation, respectively, of the HS-IGA are

Table 1 Baseline characteristics of clinical trial participants with hidradenitis suppurativa (HS)

Characteristic	PIONEER I (n = 307)	PIONEER II (n = 326)
Age (years)		
Mean (SD)	37·0 (11·1)	35·5 (11·1)
Median (range)	35·0 (18–67)	34·5 (18–69)
Body mass index (kg m^{-2})		
Mean (SD)	33·8 (7·8)	32·1 (7·7)
Median (range)	32·5 (16·4–69·8)	31·5 (16·7–60·1)
Female, n (%)	196 (63·8)	221 (67·8)
Race, n (%)		
White	234 (76·2)	273 (83·7)
Black	62 (20·2)	29 (8·9)
Other	11 (3·6)	24 (7·4)
Smoking, n (%)		
No	134 (43·7)	111 (34·0)
Yes	173 (56·4)	214 (65·6)
Unknown	0	1 (0·3)
Hurley stage, n (%)		
2	161 (52·4)	175 (53·7)
3	146 (47·6)	151 (46·3)
Previous systemic treatment, n (%)	134 (43·7)	158 (48·5)
Disease duration (years)		
Mean (SD)	11·5 (8·9)	11·6 (9·0)
Median (range)	9·2 (1·0–42·6)	9·3 (1·0–68·5)
Previous HS surgery, n (%)	34 (11·1)	45 (13·8)
Lesion counts, mean (SD); median (range)		
Total abscesses and inflammatory nodules	14·3 (13·4); 11 (3–141)	11·3 (9·7); 8 (3–66)
Abscesses	2·8 (3·6); 2 (0–24)	2·3 (3·0); 1 (0–16)
Inflammatory nodules	11·6 (12·5); 8 (0–138)	9·0 (8·4); 6 (0–62)
Draining fistulas	4·2 (4·8); 2 (0–20)	3·4 (4·7); 1 (0–20)
Patient global assessment of skin pain (pain numerical rating scale)		
Mean (SD)	5·0 (2·6)	4·5 (2·7)
Median (range)	5·1 (0–10)	4·4 (0–10)
Dermatology Life Quality Index		
Mean (SD)	16·1 (6·9)	14·5 (7·5)
Median (range)	16 (0–30)	14 (0–30)

described in Table 1. There were 3024 unique measurements across assessments at every timepoint in PIONEER I. Summary statistics for lesion counts by region, and region counts by lesion type are described in Tables 2 and 3, respectively. While maximum lesion counts were high in some patients, mean and median lesion counts by region were largely <10. Mean lesion counts were highest in axillary and inguinal regions (Table 2). Similarly, the mean and median number of regions involved were three or less for individual lesion types as well as for lesion combinations, including the combination of all lesion types. The highest mean region count was for nodules (individual lesion type), or fistula/nodule and abscess/fistula/nodule (lesion combinations) (Table 3).

Figure 2 shows the results of multivariate regression models showing the influence of lesion type and region combinations on the HS-PGA using both the ER and RPS methods. Regardless of lesion type, the axilla and inguinal regions most heavily influenced the HS-PGA score. The mean absolute error and mean square error methods found similar outcomes, reinforcing the robustness of the results. Accordingly, the results supported combining anatomical sites into two regions (upper and lower body) in the HS-IGA in order to simplify the calculation (Appendix S1; see Supporting Information).

The classification and regression tree results initially produced a four-point scale based on the total number of lesions in the 'upper' or 'lower' regions. Clinical input from the HISTORIC expert group and consideration for the Food and Drug Administration definition of responsiveness requiring ≥2-point change directed the construct to a five-point scale. The responsiveness analyses supported a six-point HS-IGA scale with scores calculated by the sum of abscesses, nodules (inflammatory or noninflammatory) and fistulas (draining or nondraining) in either the upper body regions or the lower body regions, whichever is greater, with scores being assigned as 0 (0–1 lesions), 1 (2–5), 2 (6–10), 3 (11–15), 4 (16–20) and 5 (> 20 lesions) (Table 4; and Appendix S2; see Supporting Information). Distribution of patients across HS-IGA scores, and disease responsiveness to HS-IGA compared with HiSCR, at baseline and week 12 in PIONEER I, are shown in Table 4. Using the HS-IGA, 17% and 41% of patients were responsive at week 12 and at week 36, respectively, compared with 36% and 64% at week 12 and at week 36, respectively, using HiSCR.

Validation

In the validation dataset, the intraclass correlation coefficient score for test–retest reliability (baseline vs. week 2) was 0·91 (95% confidence interval 0·87–0·94). Spearman's rank order correlations between HS-IGA and HS-PGA and between HS-IGA and HiSCR were 0·73 and 0·51, respectively ($P < 0·001$ for both comparisons). Spearman's rank order correlations between HS-IGA and DLQI, pain NRS and HS-QoL were 0·42, 0·34 and −0·25, respectively ($P < 0·001$ for all comparisons).

Using HS-IGA, 24% and 40% of patients were responsive at week 12 and at week 36, respectively, compared with 46% and 68% at week 12 and at week 36, respectively, using HiSCR.

Predictive validity was very good between HS-IGA and HS-PGA ('clear' or 'minimal') (AUC = 0·89) and between HS-IGA and HiSCR (AUC = 0·82). Predictive validity was low between HS-IGA and DLQI ('little' or 'no' effect) (AUC = 0·68) and between HS-IGA and HS-QoL ('best possible') (AUC = 0·64).

HS-IGA had clinical meaningfulness to patients, as demonstrated by statistically significant differences in all patient-

Table 2 Lesion counts by region in the development dataset (PIONEER I)

Body region (3024 measurements)	Mean lesion count	Lower quartile	Median	Upper quartile	Maximum
Axilla	7·4	0	4	11	79
Inframammary	1·7	0	0	1	37
Mammary	2·3	0	0	2	40
Upper body	9·6	1	6	14	89
Buttock	3·3	0	0	3	66
Inguinal	8·6	2	6	12	81
Perianal/perineal	1·4	0	0	1	51
Lower body	13·3	3	9	17	147

Lesion types include abscesses, fistulas (draining, nondraining) and nodules (inflammatory, noninflammatory).

Table 3 Region counts by lesion type in the development dataset (PIONEER I)

Lesion type (3024 measurements)	Mean region count	Lower quartile	Median	Upper quartile	Maximum
Abscess	0·7	0	0	1	6
Fistula	1·8	1	2	3	6
Nodule	2·4	2	2	3	6
Abscess + fistula	2·0	1	2	3	6
Abscess + nodule	2·5	2	2	3	6
Fistula + nodule	2·7	0	0	1	6
Abscess + fistula + nodule	2·7	3	2	3	6

Regions include axilla, inframammary, intermammary, buttock, inguinal and perianal/perineal.

Table 4 Distribution of patients across Hidradenitis Suppurativa Investigator Global Assessment (HS-IGA) scores and by responsiveness compared with Hidradenitis Suppurativa Clinical Response (HiSCR) in the PIONEER I development dataset

		Total patients		Patients meeting endpoints for both HiSCR[b] and HS-IGA[c]		Patients meeting endpoint for HiSCR alone		Patients meeting endpoint for HS-IGA alone		Patients not meeting endpoints for either HiSCR or HS-IGA	
HS-IGA score	Lesion count[a]	Baseline	Week 12	Baseline	Week 12	Baseline	Week 12	Baseline	Week 12	Baseline	Week 12
0	0–1	0	9	0	5	0	4	0	0	0	0
1	2–5	22	53	0	16	8	17	0	2	14	18
2	6–10	49	62	3	11	18	23	0	4	28	24
3	11–15	68	40	13	5	20	2	2	2	33	31
4	16–20	38	33	8	0	0	8	4	0	26	25
5	> 20	111	91	13	0	21	13	2	0	75	78
Total		288		37 (13%)		67 (23%)		8 (3%)		176 (61%)	

[a]Sum of abscesses, nodules (inflammatory or noninflammatory) and fistulas (draining or nondraining) in either the upper body regions or the lower body regions, whichever is greater. [b]HiSCR response is defined as ≥50% reduction in the total abscess and nodule count with no increase in abscess count and no increase in draining fistula count relative to baseline. [c]HS-IGA response is defined as ≥2-point improvement in HS-IGA score from baseline.

reported outcomes (PROs) between HS-IGA responders and nonresponders. The mean (SD) overall DLQI score was 7·8 (6·2) for HS-IGA responders and 11·2 (7·5) for nonresponders ($P < 0·001$). The mean (SD) overall HS-QoL score was 5·5 (2·7) for HS-IGA responders and 4·9 (2·3) for nonresponders ($P < 0·001$). The mean (SD) overall pain NRS was 2·8 (2·6) for HS-IGA responders and 3·9 (2·7) for nonresponders ($P < 0·001$).

Discussion

There are few validated instruments measuring disease activity or treatment response in HS. HS is complex in its presentation, owing to the significant heterogeneity in lesion types, anatomical areas and surface area involved. This may result in substantial challenges to disease severity measurement, including reliability and feasibility. Moreover, the complexity of some

existing instruments also limits their adoption in clinical practice.[17] Accordingly, there is an urgent need to develop measures that reflect disease activity in HS, that are responsive, and that raters can use with accuracy and efficiency. The Physician's Global Assessment (PGA) was introduced for dermatological disease assessment in 1998 by a US Food and Drug Administration panel as the preferred tool to measure the severity of psoriasis in interventional studies.[18,19] While one other PGA-like measure exists for HS,[11] it has not undergone rigorous assessment for its operational properties. It is also complex to calculate, making it difficult to implement in trials or in clinical practice.

Herein, we have described the development and initial validation of a novel HS-specific IGA for use as a disease activity and response measure in interventional trials. The HS-IGA utilizes the familiar construct of a six-point scale, with response defined as ≥2-point improvement from baseline. The score is based on objective lesion counts, although it limits the requirement to count beyond 21 lesions. The HS-IGA does not exclude any lesion type. The score includes noninflammatory nodules, which HISTORIC patient participants have discussed can turn into inflammatory nodules, and back again, and are thus important to measure. Inclusion of noninflammatory nodules may also make distinction from inflammatory nodules by erythema easier among patients of colour. Scars, as readily distinguishable from lesions, are not included in the score. Specification of lesion types and distinction between other difficult to discern lesions (i.e. inflammatory nodule vs. abscess, or draining abscess vs. draining fistula) are not required by the investigator, which may support measurement accuracy. The measure also accounts for anatomical regions of involvement, but it simplifies this concept further by aggregating into upper or lower body regions. This is supported by a study suggesting that subclassification according to upper and lower regions is relevant.[20] The HS-IGA showed high test–retest reliability, moderate-to-strong construct validity with HS-PGA and HiSCR, very good predictive validity with HiSCR, and responsiveness to change. The HS-IGA also showed the expected divergent validity and low predictive validity with PROs.

The HS-IGA may overcome some limitations of existing disease activity and response instruments. With HiSCR, lesion counts are limitless and lesion type distinction is required – both of which may contribute to poor operational performances for HiSCR and other commonly used instruments.[21] In an inter-rater agreement and reliability exercise among dermatologists experienced in HS, observed intervals for limits of agreement were wide relative to the ranges of the scales of all of the measurement instruments tested, including HiSCR.[9] Additionally, HiSCR response does not permit an increase in abscess or draining fistula count relative to baseline, even when >50% reduction in total abscess and nodule count has been achieved. As an example, a patient with an 80% decrease in abscess and nodule count with an increase of one fistula would be considered a HiSCR nonresponder. However, an increase in any lesion type would not alone disqualify a response as measured by the HS-IGA, provided that the overall number of lesions is improved. Moderate-to-strong construct validity with HiSCR and HS-PGA suggests that the performance of the HS-IGA is similar to that of these existing instruments while being simpler to measure.

There are also some important considerations that may limit the performance of HS-IGA as a response measure. Patients with low lesion counts (fewer than six) at baseline cannot achieve response, as defined by a two-point change on the scale, with the HS-IGA. In recent phase II and phase III interventional trials for moderate-to-severe HS, patients with as few as three abscesses and/or inflammatory nodules across two or more regions have met the inclusion criteria. While there is no standardized definition of moderate-to-severe disease in clinical trials, HISTORIC, including many HS experts and patients, are in agreement that three to five lesions may be representative of patients with mild disease in HS. Ensuring that trial cohorts are reflective of real-world patients with respect to moderate-to-severe disease activity will help ensure that approved drugs have the intended effectiveness in practice.

Finally, a greater percentage of patients were responsive to HiSCR compared with the HS-IGA, based on PIONEER II trial data. It is important to note that 22 of 288 patients (7·6%) in PIONEER I had a baseline score of 1 on the HS-IGA and could not have achieved response given the requirement for a two-point improvement in score. While the threshold for HS-IGA response is higher than for HiSCR, the instrument may also reduce the placebo response rate, which has been high in HS trials. Similarly, the instrument may be less susceptible to ceiling effects as more effective treatments are developed in the future.

Global assessments have been criticized for oversimplifying the multifaceted nature of dermatological conditions. In HS, instruments that rely on the presence or absence of lesions ignore aspects of the disease that are also critically important to patients, such as pain, magnitude of drainage, and odour. The proposed HS-IGA is designed to complement a Patient Global Assessment[22] as well as other PROs, included in the HISTORIC core measures set to ensure assessment of all aspects of the multifaceted disease deemed important to patients, experts and other stakeholders.[8] This is supported by the strong predictive validity with HiSCR and HS-PGA and low predictive validity with the PROs. Despite the low predictive validity with PROs, the HS-IGA is clinically meaningful to patients, as shown by the substantial differences in PRO scores between HS-IGA responders and nonresponders as expected. Responders as measured by the HS-IGA experienced less impact on quality of life (i.e. lower score on DLQI and higher score on HS-QoL) and less pain (lower score on pain NRS).

There are limitations to this study that warrant consideration. One of the objectives for developing an IGA was to eliminate counting of high numbers of lesions (i.e. ≥ 30, as some patients had >50 lesions in a single region). However, assessing validity based on anchoring to existing measures that rely on lesion counts may bias the performance of the HS-IGA towards the number of lesions. While development and

validation of the instrument were based on data from several clinical trials of patients with moderate-to-severe HS, instrument performance will need to be assessed among patients with greater variation in baseline severity.

In summary, we have developed and initially validated an HS-specific IGA to measure disease activity and responsiveness to an intervention. The IGA represents a frequent primary efficacy instrument to evaluate performance of interventions for dermatological diseases that are approved by the Food and Drug Administration. As such, we believe further work on the HS-IGA will be critical to drug development efforts on behalf of patients with HS, and potentially to applications in clinical practice.

References

1 Garg A, Kirby JS, Lavian J *et al*. Sex- and age-adjusted population analysis of prevalence estimates for hidradenitis suppurativa in the United States. *JAMA Dermatol* 2017; **153**:760–4.
2 Garg A, Lavian J, Lin G *et al*. Incidence of hidradenitis suppurativa in the United States: a sex- and age-adjusted population analysis. *J Am Acad Dermatol* 2017; **77**:118–22.
3 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.
4 Reddy S, Strunk A, Garg A. Comparative overall comorbidity burden among patients with hidradenitis suppurativa. *JAMA Dermatol* 2019; **155**:797–802.
5 Reddy S, Strunk A, Garg A. All-cause mortality among patients with hidradenitis suppurativa: a population-based cohort study in the United States. *J Am Acad Dermatol* 2019; **81**:937–42.
6 Tiri H, Jokelainen J, Timonen M *et al*. Substantially reduced life expectancy in patients with hidradenitis suppurativa: a Finnish nationwide registry study. *Br J Dermatol* 2019; **180**:1543–4.
7 Sartorius K, Emtestam L, Jemec GB, Lapins J. Objective scoring of hidradenitis suppurativa reflecting the role of tobacco smoking and obesity. *Br J Dermatol* 2009; **161**:831–9.
8 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.
9 Thorlacius L, Garg A, Riis PT *et al*. Inter-rater agreement and reliability of outcome measurement instruments and staging systems used in hidradenitis suppurativa. *Br J Dermatol* 2019; **181**:483–91.
10 Galimberti G, Soffritti G, Di Maso M. Classification trees for ordinal responses in R: the rpartScore package. *J Stat Softw* 2012; **47** (10):1–25.
11 Kimball AB, Kerdel F, Adams D *et al*. Adalimumab for the treatment of moderate to severe hidradenitis suppurativa: a parallel randomized trial. *Ann Intern Med* 2012; **157**:846–55.
12 Kimball AB, Jemec GB, Yang M *et al*. Assessing the validity, responsiveness and meaningfulness of the Hidradenitis Suppurativa Clinical Response (HiSCR) as the clinical endpoint for hidradenitis suppurativa treatment. *Br J Dermatol* 2014; **171**:1434–42.
13 McGraw K, Wong S. Forming inferences about some intraclass correlation coefficients. *Psychological Methods* 1996; **1**:30–46.
14 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; **33**:159–74.
15 Mukaka MM. Statistics corner: a guide to appropriate use of correlation coefficient in medical research. *Malawi Med J* 2012; **24**: 69–71.
16 Metz CE. Basic principles of ROC analysis. *Semin Nucl Med* 1978; **8**:283–98.
17 Ingram JR, Hadjieconomou S, Piguet V. Development of core outcome sets in hidradenitis suppurativa: systematic review of outcome measure instruments to inform the process. *Br J Dermatol* 2016; **175**:263–72.
18 Feldman S, Krueger G. Psoriasis assessment tools in clinical trials. *Ann Rheum Dis* 2005; **64**:ii65–8.
19 Spuls P, Lecluse L, Poulsen M *et al*. How good are clinical severity and outcome measures for psoriasis? Quantitative evaluation in a systematic review. *J Invest Dermatol* 2010; **130**:933–43.
20 Thorlacius L, Riis PT, List EK *et al*. Sub-classification of hidradenitis suppurativa: a cross-sectional study. *Arch Dermatol Res* 2022; **314**:207–12.
21 Sartorius K, Lapins J, Emtestam L, Jemec GB. Suggestions for uniform outcome variables when reporting treatment effects in hidradenitis suppurativa. *Br J Dermatol* 2003; **149**:211–13.
22 Kirby JS, Hereford B, Thorlacius L *et al*. Validation of global item for assessing impact on quality of life of patients with hidradenitis suppurativa. *Br J Dermatol* 2021; **184**:681–7.

Appendix 1: Conflicts of interest

A.G. is a consultant for AbbVie, Aclaris Therapeutics, Anaptys Bio, Aristea Therapeutics, Boehringer Ingelheim, Bristol Myers Squibb, Incyte, InflaRx, Insmed, Janssen, Novartis, Pfizer, UCB and Viela Biosciences and receives honoraria. He has grants from AbbVie and the National Psoriasis Foundation. He is a joint copyright holder of the HiSQoL. C.Z., K.K. and W.G. were previously employees of AbbVie for a portion of this study and may own AbbVie stock or stock options. N.C. is an employee of AbbVie and may own AbbVie stock or stock options. G.B.E.J. is Editor in Chief of *Dermatology* and reports grants and personal fees from AbbVie, personal fees from Coloplast, personal fees from Chemocentryx, personal fees from LEO Pharma, grants from LEO Foundation, grants from Afyx, personal fees from Incyte, grants and personal fees from InflaRx, grants from Janssen-Cilag, grants and personal fees from Novartis, grants and personal fees from UCB, grants from CSL Behring, grants from Regeneron, grants from Sanofi, personal fees from Kymera, and personal fees from Viela Bio. He is a joint copyright holder of the HiSQoL and Patient Global Assessment instruments for hidradenitis suppurativa. J.K. is a consultant for AbbVie, ChemoCentryx, Incyte, InflaRx, Janssen, Novartis, UCB and Viela Bio and receives honoraria. She is a speaker for AbbVie. She is a joint copyright holder of the HiSQoL and Patient Global Assessment instruments for hidradenitis suppurativa. L.T. has received speaker honoraria from UCB, had travel expenses covered by AbbVie and Janssen, and been an investigator for Regeneron. She is a joint copyright holder of the HiSQoL and Patient Global Assessment instruments for hidradenitis suppurativa. B.V. is a consultant for Boehringer Ingelheim and a joint copyright holder of the HiSQoL and Patient Global Assessment instruments for hidradenitis suppurativa. J.R.I. is Editor in Chief of the *British Journal of Dermatology* and receives an author honorarium from UpToDate. He is a consultant for UCB Pharma, Novartis, Boehringer Ingelheim and ChemoCentryx and has participated in advisory boards for Kymera Therapeutics and Viela Bio, all in the field of hidradenitis

suppurativa. He is a joint copyright holder of the HiSQoL and Patient Global Assessment instruments for hidradenitis suppurativa.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Appendix S1 Categorization of anatomical sites into upper and lower body regions in the Hidradenitis Suppurativa Investigator Global Assessment score.

Appendix S2 Hidradenitis Suppurativa Investigator Global Assessment score.