Assessment tools and phenotype classification for hidradenitis suppurativa

Koerts, Nicole D.K.; Bouwman, Klasiena; Prens, Lisette M.; Horváth, Barbara

Published in:
Clinics in dermatology

DOI:
10.1016/j.clindermatol.2023.08.016

Publication date:
2023

Clinics in Dermatology (2023) **41**, 601–610

Clinics in Dermatology

Assessment tools and phenotype classification for hidradenitis suppurativa

Nicole D.K. Koerts, MD*, Klasiena Bouwman, MD, Lisette M. Prens, MD, PhD, Barbara Horváth, MD, PhD

Department of Dermatology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

Abstract Hidradenitis suppurativa (HS) is a heterogeneous chronic relapsing skin disease. Several assessment tools are used to assess disease severity and to classify disease phenotype; however, no consensus exists. This review evaluates the various assessment tools and phenotypes, assessing their validity and reliability. Numerous assessment tools and phenotype classifications have been proposed for identifying various subtypes within the hidradenitis suppurativa disease spectrum. Each has a different purpose, such as use in daily practice or in clinical trial settings. Several assessment tools and phenotype classifications have been validated but not always with satisfactory results and often with studies showing divergent intra-rater reliability results. A consensus is needed for a validated, easy-to-use, and timesaving assessment tool for routine daily practice. For clinical trials, a validated and extensive assessment tool that also measures response to treatment is also needed.

Introduction

Hidradenitis suppurativa (HS) is a chronic, inflammatory, and heterogeneous skin disease of the terminal hair follicle, and patients often experience a diminished quality of life.[1] HS has its onset in early adulthood, being characterized by painful lesions such as inflammatory nodules and abscesses in the intertriginous areas of the body. In the later stage, the formation of draining fistulas and scarring can occur.[2] HS has an estimated prevalence of around 1% and is associated with several comorbidities.[3,4] The true prevalence of HS is 2.1% in the Northern Netherlands according to a population-based Dutch cohort study.[5] Within the disease spectrum of HS, different stages and phenotypes have been identified due to the heterogenic nature of the disease, namely inflammation and scarring, resulting in several proposed assessment tools for HS.[6]

In a 2016 Cochrane review, 30 HS assessment tools from 12 randomized controlled trials were evaluated; 90% of these instruments lacked validation data.[7] Subsequently, various validated tools have been developed; however, not all assessment tools and phenotype classifications have validation data to support their use. Both reliability and validity are concepts that can be used to evaluate the quality of research. Reliability refers to the consistency of an assessment tool. The intra-rater reliability indicates the degree of similarity in staging guided by an assessment tool by the same researcher. In contrast, the inter-rater reliability indicates the degree of similarity in staging guided by an assessment tool used by various researchers. The intra-rater and inter-rater reliability can be expressed in Cohen's kappa or the intraclass correlation coefficient (ICC). Validity refers to the accuracy of an assessment tool and could be expressed by, for instance, face or construct validity.

Various phenotypes of HS have been proposed that seem to require different therapeutic approaches.[8,9] Because the treatment of HS often requires a multidisciplinary approach

* Corresponding author. Tel.: +31 (0)6 2564 6184.
E-mail address: n.d.k.koerts@umcg.nl (N.D.K. Koerts).

https://doi.org/10.1016/j.clindermatol.2023.08.016

with a combination of medication and surgery, a reliable HS assessment tool to determine treatment outcome is most important; moreover, research interest in both HS and HS clinical trials has grown in the past decade, resulting in several proposed assessment tools.[10] Each proposed assessment tool has both advantages and limitations, resulting in the lack of a universally accepted clinically relevant assessment tool.

Core outcome measures for HS trials are being developed, and the international HISTORIC (Hidradenitis Suppurativa Core Outcomes Set International Collaboration) initiative is evaluating outcome measurement instruments to determine the essential outcomes in HS.[11] Because no consensus exists, we have evaluated the different assessment tools and phenotype classifications for HS. We have also explored and summarized their validity and reliability where applicable.

Hurley classifications

Original Hurley classification

The original Hurley classification system is one of the most widely used HS assessment tools. It was first described in 1989 by Harry J. Hurley, Jr, MD (1926-2009) and stratified patients with HS into three stages of severity (Figure 1)[12]:

- Stage I: Abscesses and inflammatory nodules without sinus tracts and scarring
- Stage II: Recurrent abscesses with tract formation and scarring, single or multiple and widely separated lesions
- Stage III: Diffuse or multiple interconnected sinus tracts and abscesses across the entire area

Initially, the Hurley staging system was designed as a guide for surgical treatment because it was intended only to describe signs in one anatomic region. It contains characteristics such as scarring and formation of fistulas, treatable with surgery. Because inflammatory aspects of the disease are not included, this assessment tool is less useful for monitoring medical treatment. The original Hurley classification does not assess any extension of the disease. As a result, the Hurley classification is unsuitable for assessing disease severity in all sites that the patient may have. Unfortunately, this system is widely used.

In 2019, a group determined the inter- and intra-rater reliability of the system. They concluded that the Hurley classification is reliable for a rapid assessment of HS, with moderate inter-rater and substantial intra-rater reliability for all stages (Table 1).[13] In another study in which 12 HS experts assessed nine assessment tools for HS in 24 patients,[14] they found excellent inter-rater reliability for the gluteal and axillary region but moderate inter-rater reliability for the groin. Another group compared and assessed the reliability and reproducibility of six different assessment tools within a group of dermatology residents.[15] For the Hurley staging system, excellent intra-rater reliability and a slight inter-rater variability were found. In additional report, there is a study to evaluate inter-rater reliability in HS disease severity assessment using clinical and ultrasonography (US) techniques.[16] Twenty patients were assessed by two physicians using four different assessment tools before and after the US. Inter-rater agreement of each outcome measure before and after the US was obtained, implying an improvement of the overall inter-rater agreement using the US. The use of ultrasound has not been widely implemented among dermatologists.[17] An excellent inter-rater agreement was found for the original Hurley classification in the pre-US assessment.

Refined Hurley classification

In 2017, a Dutch HS expert group proposed modifying the original Hurley staging system, named the refined Hurley classification. The refined Hurley classification assesses the presence of sinus tracts, inflammation, and the number of affected body parts, resulting in seven stages with a subdivision in the first two stages into mild, moderate, and severe

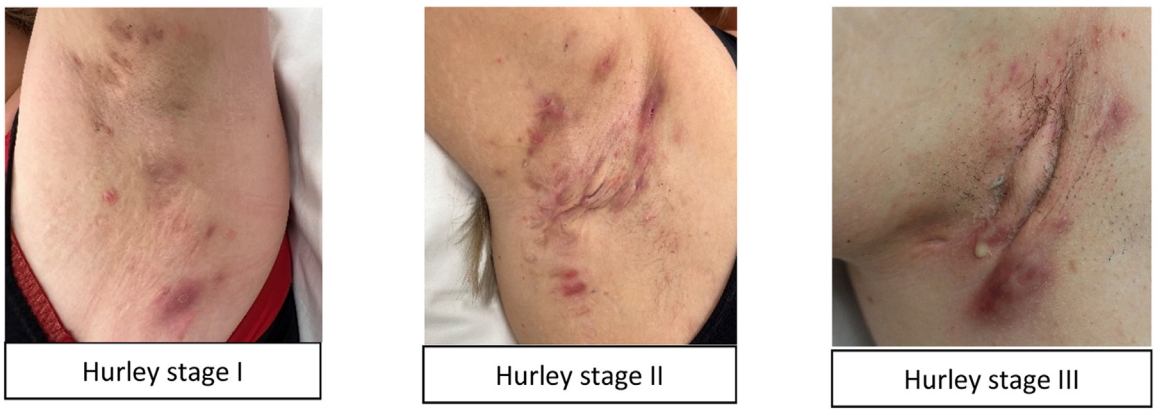

Fig. 1 Examples of the three stages of the original Hurley classification.

Table 1 Overview of HS assessment tools and phenotype classifications, including their validation.

Assessment	Characteristics	Goal	Validated by	Type of validation	Validation outcome
Classifications					
Hurley classification by Hurley (1989)[12]	Three severity stages and staging per region	HS severity classification and (surgical) treatment guidance	Ovadja et al (2018)[13]	Inter- and intra-rater reliability	Moderate to good inter-rater reliability (j = 0.59, 95% CI 0.48-0.70)
			Thorlacius et al (2019)[14]	Inter-rater reliability	Substantial inter-rater reliability (j = 0.65, 95% CI 0.58-0.72)
			Wlodarek et al (20120)[12]	Intra-rater reliability and inter-rater variability	Good inter-rater reliability (moderate for groin region)
			Lyons et al (2022)[16]	Inter-rater reliability	Hurley groin ICC: 0.55 (95% CI 0.44-0.67); gluteal ICC: 0.72 (95% CI 0.62-0.80); axillae ICC: 0.72 (95% CI 0.63-0.81)
					Excellent intra-rater reliability (ICC: 0.96)
					Slight inter-rater variability CVs 16.2 ± 9.4
					Excellent rater agreement (before ultrasonography) (ICC: 0.71, 95% CI 0.39-0.87)
Refined Hurley classification by Dutch HS expert group (2017)[18]	Seven severity stages and staging per patient	HS severity classification and treatment guidance	Rondags et al (2019)[19]	Inter-reliability	Moderate to high inter-rater and intra-rater agreement and reliability
			Thorlacius et al (2019)[14]	Inter-rater reliability	Real-life assessment inter-rater reliability ranged from $\alpha = 068$ (95% CI 0.32-0.95) to $\alpha = 0.92$ (95% CI 0.78-1.00); photograph assessment: 0.83 (95% CI 0.78-0.89)
			Wlodarek et al (2020)[15]	Intra-rater reliability and inter-rater variability	Fair inter-rater reliability (ICC: 0.51, 95% CI 0.35-0.68)
					Excellent intra-rater reliability (ICC: 0.95)
					Slight inter-rater variability CVs 19.3 ± 9.6
Canoui-Poitrine phenotypes by Canoui-Poitrine et al (2013)[8]	Three phenotypes	Identifying different phenotypes	Van Straalen et al (2018)[21]	Inter-rater reliability	Slight inter-rater reliability κ of 0.37 (95% CI 0.32-0.42)
			Frew et al (2019)[22]	Inter-rater reliability	κ of 0.44 (95% CI 0.39-0.48)
	Six phenotypes	Identifying different phenotypes	Frew et al (2019)[22]		κ of 0.82 (95% CI 0.80-0.83)
van der Zee and Jemec phenotypes by van der Zee and Jemec (2015)[9] and Dudink et al (2022)[23]	Four phenotypes	Identifying different phenotypes			No validation yet
HS severity assessment tools					
HS score by Sartorius et al (2003)[44]	Lesion and region count	Quantifying disease intensity	Wlodarek et al (2020)[15]	Intra-rater and inter-rater variability	Excellent intra-rater reliability (ICC: 0.97)
		Useful in clinical trials	Lyons et al (2022)[16]	Inter-rater reliability	Moderate inter-rater variability CVs 30.9 ± 5.4
Modified HS score by Sartorius et al (2009)[45]	Lesion and region count	Quantifying disease intensity			Good inter-rater agreement (ICC: 0.71, 95% CI 0.21-0.81)
		Useful in clinical trials			No validation yet

(*continued on next page*)

Table 1 (*continued*)

Assessment	Characteristics	Goal	Validated by	Type of validation	Validation outcome
HS Physician's Global Assessment by Kimball et al (2012)[24]	Lesion count	Quantifying disease intensity Useful in clinical trials	Thorlacius et al (2019)[14] Wlodarek et al (20120)[15] Lyons et al (2022)[16]	Inter-rater reliability Intra-rater reliability and inter-rater variability Inter-rater reliability	ICC: 0.64 (95% CI 0.50-0.79) Excellent intra-rater reliability (ICC: 0.94) Slight inter-rater variability CVs 16.6 ± 8.2 Good rater agreement (ICC: 0.53, 95% CI 0.13-0.79)
Hidradenitis Suppurativa Clinical Response by Kimball et al (2014)[37]	Response after treatment	Quantifying treatment response Useful in clinical trials	Thorlacius et al (2019)[14] Lyons et al (2022)[16]	Inter-rater reliability Inter-rater reliability	Fair (ICC: 0.44, 95% CI 0.29-0.63) For AN "good" rater agreement: 0.69 (95% CI 0.37-0.86) For draining fistula count, "poor" rater agreement: 0.20 (95% CI 0.00-0.58)
International Hidradenitis Suppurativa Severity Score System by European HS Foundation (2017)[38]	Lesion count	Quantifying disease intensity Useful in clinical trials	Thorlacius et al (2019)[14] Wlodarek et al (2020)[15]	Inter-rater reliability Intra-rater reliability and inter-rater variability	Fair inter-rater reliability (ICC: 0.54, 95% CI 0.39-0.71) Excellent intra-rater reliability (ICC: 0.87) Slight inter-rater variability CVs 6.5 ± 10.1
Hidradenitis Suppurativa Severity Index by Grant et al (2010)[26]	Lesion and region count Subjective parameters	Quantifying disease intensity Useful in clinical trials	Thorlacius et al (2019)[14] Wlodarek et al (2020)[15]	Inter-rater reliability Intra-rater reliability and inter-rater variability	Good inter-rater reliability (ICC: 0.78, 95% CI 0.66-0.88) Mild/moderate/severe: Excellent intra-rater reliability (ICC: 0.92) Slight inter-rater variability CVs 11.6 ± 11.3 Points: Excellent intra-rater reliability (ICC: 0.93) Slight inter-rater variability CVs 13.4 ± 8.0
Acne Inversa Severity Index by Chiricozzi et al (2015)[40]	Lesion and region count Subjective parameters	Quantifying disease intensity Useful in clinical trials	Thorlacius et al (2019)[14]	Inter-rater reliability	Fair (ICC: 0.40, 95% CI 0.25-0.59)
Hidradenitis Suppurativa Area and Severity Index Revised by the HISTORIC initiative (2021)[43] HS Area and Severity Index (2020) Severity and Area Score for Hidradenitis (2020)	Lesion count, including body surface area	Quantifying disease intensity Useful in clinical trials	Goldfarb et al (2021)[43] Goldfarb et al (YYYY)[xx] Goldfarb et al (YYYY)[xx]	Inter-rater reliability and intra-rater reliability Inter-rater reliability Inter-rater reliability and intra-rater reliability	Moderate inter-rater reliability (ICC: 0.60) High intra-rater reliability (ICC: 0.91) Good inter-rater reliability (ICC: 0.86, 95% CI 0.49-0.99) Acceptable inter-rater reliability (ICC: 0.60, 95% CI 0.44-0.74) High intra-rater reliability (ICC: 0.98, 95% CI 0.94-1.00)

AN, abscesses and inflammatory nodules; CV; coefficient of variation, ; HISTORIC, Hidradenitis Suppurativa Core Outcomes Set International Collaboration; HS, hidradenitis suppurativa; ICC, intraclass correlation coefficient.

Fig. 2 QR codes for the hidradenitis suppurativa application.

based on the degree of inflammation and extent of the disease (Figure 2).[18]

Because inflammatory aspects were included in the refined staging system, this staging is valuable for monitoring medical treatment and is suitable for clinical trials. In addition, a trained eye assessment of a particular patient can be quick and easy.

An accurate correlation between the refined Hurley classification and HS severity assessed by both patients and clinicians was found in 2018.[19] Another study assessed 25 real-life patients and 15 digital cases and concluded an overall moderate to high inter-rater and intra-rater agreement and reliability in real-life as well as in the digital cases, with high face validity results being reported.[20] Confusing the problem, one group determined fair inter-rater reliability for the refined Hurley classification,[14] whereas another study showed slight inter-rater variability and excellent intra-rater reliability.[15]

Phenotype-based classifications

HS phenotypes proposal

Florence Canoui-Poitrine, a physician at Hôpital Henri Mondor in Paris, France, was the first to propose different phenotypes of HS in 2013, suggesting three subtypes of HS.[8] Considering the heterogeneous presentation, it is likely that several underlying subtypes could be classified. This could be useful because the identified subtypes could benefit from different therapeutic approaches. A latent class analysis on prospective clinical data of 618 patients was performed, resulting in three outlined phenotypes, namely the "axillary-mammary" (LC1), "follicular" (LC2), and "gluteal" (LC3) phenotypes. Patients with the axillary-mammary type have a high probability of axillary and mammary involvement, plus hypertrophic scars. In the follicular type, other body sides are affected, such as ears, chest, back, and legs. This type is more often related to follicular lesions and severe acne, whereas the gluteal type is highlighted by gluteal involvement.

In a study to evaluate the inter-rater reliability in a clinical setting using a panel of eight HS experts, there was an agreement for 23.3% of cases on the same phenotype, indicating slight inter-rater reliability.[21] Another group conducted a genotype-phenotype correlation study in 2019, in which the inter-rater reliability of four HS clinical phenotype classifications was assessed.[22] A panel of three independent experts was asked to classify different cases phenotypically. For the Canoui-Poitrine phenotypes, there is poor inter-rater reliability, limiting its clinical usefulness.

HS phenotypes proposed by van der Zee and Jemec

In 2015, an additional phenotype classification was proposed based on expert opinion. It uses six phenotypes[9]:

- Regular type
- Frictional furuncle type
- Scarring folliculitis type
- Conglobata type
- Syndromic type
- Ectopic type

The frictional furuncle type is characterized by multiple deep nodules and abscesses on frictional areas, where patients are often overweight. Patients with the scarring folliculitis type suffer from pustules, cysts, scarring, and double-ended comedones. Most frequently, the buttocks, inguinal, and pubic region are affected, and patients are often overweight and smokers. The conglobata type is characterized by cyst formation, especially on the back and the face. Patients are usually men and not overweight. The syndromic type may be characterized by other manifestations, such as pyoderma gangrenosum and arthritis. The ectopic type is marked by the involvement of the face and back. Lastly, if patients do not fit one of the phenotypes, they are considered to have the regular type, which has been thought to be the most common phenotype. According to another study, the highest Cohen kappa values were found for the van der Zee and Jemec phenotypes, with an agreement in 84% of the cases.[22]

In 2021, one center suggested that the ectopic and syndromic types do not have specific clinical features and could be classified as one of the other phenotypes.[23] They assessed the prevalence and patient characteristics of the four phenotypes in 935 Dutch patients with HS. The regular type was the most common variant, with 75.9% of the cases, followed by the frictional furuncle type (10.3%), scarring folliculitis (7.2%), and the conglobata phenotype (6.6%). A more tailored therapy for patients with HS could be offered by linking

the phenotype and genotype. To date, no validation of these phenotypes has taken place.

HS severity assessment tools

Hidradenitis Suppurativa Score by Sartorius et al

The Hidradenitis Suppurativa Score (HSS) was proposed in 2003 as a dynamic scoring system for HS, including the number of involved anatomical regions, the number and scores of the lesions, the longest distance between two lesions, and the separation between lesions by normal skin[24]:

- The anatomic regions involved (3 points per region)
- Number and scores of lesions (points per lesion: fistulas 4; nodules 2; scars 1; others 1)
- Longest distance between two lesions (<5 cm: 2; <10 cm: 4; >10 cm: 8)
- Separation of all lesions by normal skin (yes: 0/no: 6)

In 2009, this group modified the original version of the HSS with a change in lesions and the number of points given for each parameter:

- The anatomic regions involved (3 points per region)
- Number and scores of lesions (points per lesion: fistulas 6; nodules 1)
- Longest distance between two lesions (<5 cm: 1; <10 cm: 3; >10 cm: 9)
- Separation of all lesions by normal skin (yes: 0/no: 9)

The Sartorius HSS is a dynamic and detailed scoring system, and it assesses the severity of HS in mild cases relatively straightforwardly; however, in patients with more severe HS, it can be challenging to distinguish between lesions when interconnected confluent sinuses are present. Using the HSS is quite time-consuming; therefore, it is unsuitable for daily practice. Another group determined an excellent intra-rater and a moderate inter-rater variability.[15] An additional group found an excellent inter-rater agreement for the original Sartorius HSS.[16] For the Modified HSS, no validation has been performed yet.

Hidradenitis Suppurativa Physician's Global Assessment Scale

The Hidradenitis Suppurativa Physician's Global Assessment Scale (HS-PGA) is based on lesion count, developed in 2012.[24] Rating varies between clear and severe.

- Clear: 0 abscesses, 0 draining fistulas, 0 inflammatory nodules, 0 noninflammatory nodules
- Minimal: 0 abscesses, 0 draining fistulas, 0 inflammatory nodules, and the presence of noninflammatory nodules
- Mild: 0 abscesses, 0 draining fistulas, and 1 to 4 inflammatory nodules or 1 abscess or draining fistula and 0 inflammatory nodules
- Moderate: 0 abscesses, 0 draining fistulas, and >5 inflammatory nodules or 1 abscess or draining fistula and >1 inflammatory nodule or 2 to 5 abscesses or draining fistulas and <10 inflammatory nodules
- Severe: 2 to 5 abscesses or draining fistulas and >10 inflammatory nodules
- Very severe: >5 abscesses

The HS-PGA is relatively easy to use; however, by using the HS-PGA score, patients may experience clinical improvement without reducing their HS-PGA score. As a result, this assessment tool is less helpful in indicating disease severity related to treatment for HS. A good inter-rater reliability was shown by one center[14] and supported by another, where there was a good inter-rater reliability for the HS-PGA.[16] An additional center examined excellent intra-rater reliability but slight inter-rater variability.[15]

Outcome measures in clinical trial settings

With the increase of clinical trials for HS, some assessment tools have been developed predominantly for use in the clinical trial setting. This development has led to a shift in the purpose of the assessment tool from identifying disease severity in the individual patient to indicating the response to treatment in a group of patients. Examples are the HS-PGA, created for an adalimumab trial, and the Hidradenitis Suppurativa Clinical Response (HiSCR). An overview of assessment tools used in randomized controlled trials is shown in Table 2.

Hidradenitis Suppurativa Clinical Response

The HiSCR was retrospectively developed with data from a phase 2 randomized controlled trial assessing the role of adalimumab in the treatment of HS in 2014.[37] The HiSCR mainly focuses on the inflammatory aspect of the disease and is used in trials to determine the response of treatment. The definition of responders, also called HiSCR achievers, to treatment is as follows:

- More than 50% reduction of the total of inflammatory nodules and abscesses
- No increase in the number of abscesses relative to the baseline
- No increase in the number of draining fistulas relative to the baseline

Because the total of sinuses is not included in this score, it could be considered the main limitation, because the presence of sinuses, not only those draining, could be significant. Another limitation of the HiSCR is that it is designed for the

Table 2 Overview of assessment tools used in RCTs for HS.

Study (first author, year)	Study design	Interventions	Used assessment tools
Adams, 2010[25]	2-arm, parallel-group RCT	Etanercept/placebo	HS-PGA Patient's Global Assessment DLQI
Grant, 2010[26]	2-arm, parallel-group RCT	Infliximab/placebo	HSSI DLQI VAS Static PGA
Highton, 2011[27]	Within-patient RCT	Intense pulsed light/untreated control side	Sartorius score
Miller, 2011[28]	2-arm, parallel-group RCT	Adalimumab/placebo	Sartorius score Hurley classification PGA scar scoring VAS DLQI
Kimball, 2012[24]	3-arm, parallel-group RCT	Adalimumab 40 mg EOW/adalimumab 40 mg weekly/placebo	HS-PGA MSS VAS DLQI
Yildiz, 2016[29]	Prospective RCT	Adjunctive hyperbaric oxygen therapy in patients treated with clindamycin/rifampicin	MSS HSSI DLQI VAS
Tzanetakou, 2016[30]	Placebo-controlled RCT	Anakinra/placebo	HiSCR Sartorius score VAS DLQI
Kimball, 2016 (PIONEER I and II studies)[31]	Phase 3, multicenter, double-blind RCTs	Adalimumab/placebo	HiSCR MSS
Vossen, 2019[32]	Placebo-controlled RCT	Apremilast/placebo	HiSCR DLQI
Andersen, 2020[33]	Single-blind RCT	Intense pulsed light hair removal/untreated control side	HiSCR MSS HS-PGA
Glatt, 2021[34]	Phase 2, double-blind, placebo-controlled RCT	Bimekizumab/adalimumab/placebo	HiSCR IHS4 Patient's Global Assessment DLQI
Bechara, 2021 (SHARPS trial)[35]	Phase 4, double-blind, placebo-controlled RCT	Adalimumab/placebo	HiSCR DLQI Patient's Global Assessment
Schultheis, 2022 (RELIEVE study)[36]	Multicenter RCT	Topical clindamycin combined with LAight therapy/topical clindamycin	HiSCR IHS4

DLQI, Dermatology Life Quality Index; EOW, every other week.; HiSCR, Hidradenitis Suppurativa Clinical Response; HS, hidradenitis suppurativa; HS-PGA, Hidradenitis Suppurativa Physician's Global Assessment Scale; HSSI, Hidradenitis Suppurativa Severity Index; IHS4, International Hidradenitis Suppurativa Severity Scoring System; MSS, modified Sartorius score; PGA, Physician's Global Assessment; RCT, randomized controlled trial; VAS, visual analogue scale.

trial setting only. It was created to measure change over time, but it is impossible to measure cross-sectional disease severity. One center has determined fair inter-rater reliability.[14] More recently, another group found good inter-rater reliability for inflammatory nodules and abscesses but poor inter-rater reliability for draining fistulas.[16]

International Hidradenitis Suppurativa Severity Scoring System

The European HS Foundation developed the International Hidradenitis Suppurativa Severity Scoring System (IHS4).[38] The IHS4 is a dynamic assessment tool that can be used both in real life and in clinical trial settings. The IHS4 score is obtained by the number of nodules (multiplied by 1), the number of abscesses (multiplied by 2), and the number of draining tunnels (multiplied by 4). After the summation of the score is obtained, a subdivision can be made between the following:

- Mild HS: ≤ 3
- Moderate HS: 4 to 10
- Severe HS: ≥ 11

Similar to the HiSCR in the IHS4 score, only draining tunnels are considered, whereas non-draining sinuses could also be bothersome for patients.

Regarding the reliability of the IHS4, one group has determined fair inter-rater reliability.[14] Another center has observed slight inter-rater reliability but excellent intra-rater reliability.[15]

Hidradenitis Suppurativa Severity Index

The Hidradenitis Suppurativa Severity Index (HSSI) has been used in several clinical trials.[26] This score incorporates either objective or subjective patient-reported variables.[39] The HSSI includes the number of affected sites, the body surface area (BSA), the number of lesions (erythematous or painful), drainage by assessing dressing changes during working hours, and the pain experienced by the patient by using the visual analogue scale (VAS) score.

The scoring system (points) is a total of the following:

- The BSA (%) [palm(s)]: 0 (0 points); 1 (1 point); 2 to 3 (2 points); 4 to 5 (3 points); >5 (4 points)
- Number of lesions: 0 (0 points); 1 to 2 (1 point); 2 to 3 (2 points); 4 to 5 (3 points); >5 (4 points)
- Drainage: 0 (0 points); 1 (2 points); >1 (3 points)
- Pain (VAS): 0 to 1 (0 points); 1 to 2 (1 point); 2 to 4 (2 points); 5 to 7 (3 points); 8 to 10 (4 points)

A score between 0 and 7 is considered mild HS, 8 to 12 as moderate HS, and >13 as severe HS. A unique aspect of the HSSI is including the VAS score in the overall rating. It makes the HSSI suitable for obtaining a comprehensive picture of the disease severity and indicates disease burden. Adding a subjective component to the score makes this assessment tool, logically, also less objective.

One study found good inter-rater reliability for the HSSI,[14] and this was in opposition to another report, which found slight inter-rater variability and excellent intra-rater reliability for the scoring in points and the categories mild, moderate, and severe.[15]

Acne Inversa Severity Index

The Acne Inversa Severity Index (AISI) was designed to define disease severity.[40] First, the kind of apparent lesion is assessed, and per lesion, it is supposed to be multiplied by the overall number of sites where the lesion occurs:

1. point for a comedogenic lesion
2. points for an abscess or inflammatory nodule
3. points for a sinus tract
4. points for keloid and fibrotic adherence
5. points for a fibrosclerotic inflammatory plaque

The VAS score is added to the sum of the lesions for obtaining the total score. Mild HS is considered less than 10 points, moderate HS is between 10 and 18, and severe HS is greater than 18.

Similar to the HSSI, a subjective parameter is included in this assessment tool, resulting in an extensive but less objective scoring system for HS. One center assessed the inter-rater reliability of the AISI, and fair inter-rater reliability was found.[14]

Hidradenitis Suppurativa Area and Severity Index Revised

The Hidradenitis Suppurativa Area and Severity Index Revised (HASI-R) is a collaboration of two groups who simultaneously created assessment tools with the same goal.

In a 2020 publication, the HS Area and Severity Index (HASI) tool was based on the Psoriasis Area and Severity Index concept.[41] This tool incorporates the signs of HS inflammation of various body locations with the estimated involved BSA. The proposers also determined the inter-rater reliability of HASI using three patients with HS and a panel of seven dermatologists, and a good inter-rater reliability was found for the HASI.

An additional group created a novel outcome instrument for the assessment of HS named the Severity and Area Score for Hidradenitis (SASH).[42] The SASH, developed from qualitative interviews and focus groups of clinicians, includes inflammatory color change, induration, and the amount of open skin surface with an estimate of involved BSA. In the evaluation stage, the SASH was assessed by 10 clinicians in 23

patients. There was acceptable inter-rater reliability and high intra-rater reliability.

Because both assessment tools had limitations, such as the absence of tunnel assessment, the HISTORIC initiative created the HASI-R in 2021[43]:

First, a site-specific score is calculated:

1. The sum of the four intensity scores: inflammatory color change, inflammatory induration, open skin surface, and extent of tunnels. The four variables are scored on a Likert scale from 0 to 3 (0 = none, 1 = limited/mid, 2 = moderate, 3 = severe/extensive) based on the average intensity for each body site (maximum intensity score of 12).

Multiplied by:

2. The BSA ordinal score, which is calculated as the proportion of the site involved by HS. (0 points = 0 BSA, 1 point = 1% to 3%, 2 points = 4% to 9%, 3 points = 10% to 20%, 4 points = 21% to 29%, 5 points = 30% to 50%, 6 points = >51%).

The maximum score for each body site is 72. The HASI-R assesses HS activity at 10 body sites so that the total can range between 0 and 720. The HASI-R was assessed by 20 raters who evaluated 15 patients with HS. There was moderate inter-rater variability and high intra-rater reliability. In addition, there was good construct validity. The HASI-R is, together with the HASI and SASH, the first validated assessment tool including BSA for HS; however, scoring a patient by using the HASI-R tool is time-consuming and probably less efficient in daily practice.

Conclusions

Many different assessment tools and phenotype classifications for HS have been proposed. When comparing these assessment tools and phenotype classifications, it is essential to keep their particular purpose in mind, because their use can differ between routine daily practice and a clinical trial setting. Subsequent to the systematic review in 2016, several assessment tools and phenotype classifications have been validated, but not always with satisfactory results. In addition, sometimes the inter-rater reliability between studies for the same assessment tool varies (Table 1).

what are table footnotesSome assessment tools are mainly focused on the classification of HS, like the original or refined Hurley classification, and could be a guide for (surgical) treatment. For phenotype classification systems, the main goal is to identify the different phenotypes of HS. Because there are different subtypes within the heterogeneous HS disease spectrum, it is conceivable that such subtypes could benefit from potentially different therapeutic approaches in the future. The facts remain that these assessment tools are more focused on classifying disease severity than on measuring response to therapy.

Despite the number of assessment tools that have been proposed, no gold standard for assessing HS has been identified. To reach a consensus on which assessment tool to use for HS, we suggest distinguishing between daily practice and a clinical trial setting. Regarding use in daily practice, there is a need for a validated, easy-to-use, and timesaving assessment tool that indicates disease severity. The refined Hurley classification, the HS-PGA, and the IHS4 might have the suitable capacity for this purpose. An extended and validated tool measuring therapeutic response is recommended for clinical trials, such as the (modified) Sartorius score, the HiSCR, the IHS4, or the HASI-R. By incorporating subjective components, the HSSI and AISI are also motivating candidates to obtain a comprehensive image of disease severity and burden.

Declaration of Competing Interest

B. Horváth reports fees from Janssen-Cilag (advisory boards, educational grants, consultations, investigator initiative studies), AbbVie (advisory boards, educational grants, consultations, investigator initiative studies), Novartis Pharma (advisory boards, consultations, investigator initiative studies), UCB Pharma (advisory boards, consultations), Leo Pharma (consultation), Solenne B.V. (investigator initiative studies), Celgene (consultations, investigator initiative studies), Akari Therapeutics (consultations, investigator initiative studies), Philips (consultation), Roche (consultation), Regeneron (consultation), and Sanofi (consultation), the fees of which were paid to the institution. N.D.K. Koerts, K. Bouwman, and L.M. Prens have no conflicts of interest to declare.

References

1. Sabat R, Jemec GBE, Matusiak Ł, Kimball AB, Prens E, Wolk K. Hidradenitis suppurativa. *Nat Rev Dis Primers*. 2020;6:18.
2. Zouboulis CC, Del Marmol V, Mrowietz U, Prens EP, Tzellos T, Jemec GB. Hidradenitis suppurativa/acne inversa: criteria for diagnosis, severity assessment, classification and disease evaluation. *Dermatology*. 2015;231:184–190.
3. Revuz JE, Canoui-Poitrine F, Wolkenstein P, et al. Prevalence and factors associated with hidradenitis suppurativa: results from two case-control studies. *J Am Acad Dermatol*. 2008;59:596–601.
4. Revuz J. Hidradenitis suppurativa. *J Eur Acad Dermatol Venereol*. 2009;23:985–998.
5. Prens LM, Bouwman K, Troelstra LD, Prens EP, Alizadeh BZ, Horváth B. New insights in hidradenitis suppurativa from a population-based Dutch cohort: prevalence, smoking behaviour, socioeconomic status and comorbidities. *Br J Dermatol*. 2022;186:814–822.
6. Kirby JS. Unraveling the heterogeneity of hidradenitis suppurativa with phenotype schema. *J Invest Dermatol*. 2021;141:1136–1138.
7. 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;175:263–272.

8. Canoui-Poitrine F, Le Thuaut A, Revuz JE, et al. Identification of three hidradenitis suppurativa phenotypes: latent class analysis of a cross-sectional study. *J Invest Dermatol*. 2013;133:1506–1511.
9. van der Zee HH, Jemec GB. New insights into the diagnosis of hidradenitis suppurativa: clinical presentations and phenotypes. *J Am Acad Dermatol*. 2015;73(5 suppl 1):S23–S26.
10. Jemec GB. Clinical practice. Hidradenitis suppurativa. *N Engl J Med*. 2012;366:158–164.
11. 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;179:642–650.
12. Hurley HJ. Axillary hyperhidrosis, apocrine bromhidrosis, hidradenitis suppurativa, and familial benign pemphigus: surgical approach. In: Roenigk RK, Roenigk HH, eds. *Dermatologic Surgery: Principles and Practice*. New York: Marcel Dekker; 1989:729–739.
13. Ovadja ZN, Schuit MM, van der Horst CMAM, Lapid O. Inter- and intrarater reliability of Hurley staging for hidradenitis suppurativa. *Br J Dermatol*. 2019;181:344–349.
14. 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;181:483–491.
15. Włodarek K, Stefaniak A, Matusiak Ł, Szepietowski JC. Could residents adequately assess the severity of hidradenitis suppurativa? Inter-rater and intrarater reliability assessment of major scoring systems. *Dermatology*. 2020;236:8–14.
16. Lyons AB, Narla S, Kohli I, et al. Assessment of inter-rater reliability of clinical hidradenitis suppurativa outcome measures using ultrasonography. *Clin Exp Dermatol*. 2022;47:319–324.
17. Martorell A, Alfageme Roldán F, Vilarrasa Rull E, et al. Ultrasound as a diagnostic and management tool in hidradenitis suppurativa patients: a multicentre study. *J Eur Acad Dermatol Venereol*. 2019;33:2137–2142.
18. Horváth B, Janse IC, Blok JL, et al. Hurley staging refined: a proposal by the Dutch Hidradenitis Suppurativa Expert Group. *Acta Derm Venereol*. 2017;97:412–413.
19. Rondags A, van Straalen KR, van Hasselt JR, et al. Correlation of the refined Hurley classification for hidradenitis suppurativa with patient-reported quality of life and objective disease severity assessment. *Br J Dermatol*. 2019;180:1214–1220.
20. Prens LM, Rondags A, Volkering RJ, et al. The refined Hurley classification: the inter-rater and intrarater reliability and face validity. *Br J Dermatol*. 2019;181:1335–1337.
21. van Straalen KR, Verhagen T, Horváth B, et al. Poor interrater reliability of hidradenitis suppurativa phenotypes. *J Am Acad Dermatol*. 2018;79:577–578.
22. Frew JW, Hawkes JE, Sullivan-Whalen M, Gilleaudeau P, Krueger JG. Inter-rater reliability of phenotypes and exploratory genotype-phenotype analysis in inherited hidradenitis suppurativa. *Br J Dermatol*. 2019;181:566–571.
23. Dudink K, Aarts P, Ardon CB, et al. Prevalence and clinical characteristics of hidradenitis suppurativa phenotypes in a large Dutch cohort. *Dermatology*. 2022;238:600–602.
24. 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;157:846–855.
25. Adams DR, Yankura JA, Fogelberg AC, Anderson BE. Treatment of hidradenitis suppurativa with etanercept injection. *Arch Dermatol*. 2010;146:501–504.
26. Grant A, Gonzalez T, Montgomery MO, Cardenas V, Kerdel FA. Infliximab therapy for patients with moderate to severe hidradenitis suppurativa: a randomized, double-blind, placebo-controlled crossover trial. *J Am Acad Dermatol*. 2010;62:205–217.
27. Highton L, Chan WY, Khwaja N, Laitung JKG. Treatment of hidradenitis suppurativa with intense pulsed light: a prospective study. *Plast Reconstr Surg*. 2011;128:459–466.
28. Miller I, Lynggaard CD, Lophaven S, Zachariae C, Dufour DN, Jemec GB. A double-blind placebo-controlled randomized trial of adalimumab in the treatment of hidradenitis suppurativa. *Br J Dermatol*. 2011;165:391–398.
29. Yildiz H, Senol L, Ercan E, Bilgili ME, Karabudak Abuaf O. A prospective randomized controlled trial assessing the efficacy of adjunctive hyperbaric oxygen therapy in the treatment of hidradenitis suppurativa. *Int J Dermatol*. 2016;55:232–237.
30. Tzanetakou V, Kanni T, Giatrakou S, et al. Safety and efficacy of anakinra in severe hidradenitis suppurativa: a randomized clinical trial. *JAMA Dermatol*. 2016;152:52–59 Published correction appears in *JAMA Dermatol*. 2017;153:950.
31. Kimball AB, Okun MM, Williams DA, et al. Two phase 3 trials of adalimumab for hidradenitis suppurativa. *N Engl J Med*. 2016;375:422–434.
32. Vossen ARJV, van Doorn MBA, van der Zee HH, Prens EP. Apremilast for moderate hidradenitis suppurativa: results of a randomized controlled trial. *J Am Acad Dermatol*. 2019;80:80–88.
33. Andersen PL, Riis PT, Thorlacius L, et al. Intense pulsed light treatment for hidradenitis suppurativa: a within-person randomized controlled trial. *Eur J Dermatol*. 2020;30:723–729.
34. Glatt S, Jemec GBE, Forman S, et al. Efficacy and safety of bimekizumab in moderate to severe hidradenitis suppurativa: a phase 2, double-blind, placebo-controlled randomized clinical trial. *JAMA Dermatol*. 2021;157:1279–1288 Published correction appears in *JAMA Dermatol*. 2021;157:1384.
35. Bechara FG, Podda M, Prens EP, et al. Efficacy and safety of adalimumab in conjunction with surgery in moderate to severe hidradenitis suppurativa: the SHARPS randomized clinical trial. *JAMA Surg*. 2021;156:1001–1009.
36. Schultheis M, Staubach P, Nikolakis G, et al. LAight® therapy significantly enhances treatment efficacy of 16 weeks of topical clindamycin solution in Hurley I and II hidradenitis suppurativa: results from period A of RELIEVE, a multicenter randomized, controlled trial. *Dermatology*. 2022;238:476–486.
37. 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;171:1434–1442.
38. Zouboulis CC, Tzellos T, Kyrgidis A, et al. Development and validation of the International Hidradenitis Suppurativa Severity Score System (IHS4), a novel dynamic scoring system to assess HS severity. *Br J Dermatol*. 2017;177:1401–1409.
39. Amano M, Grant A, Kerdel FA. A prospective open-label clinical trial of adalimumab for the treatment of hidradenitis suppurativa. *Int J Dermatol*. 2010;49:950–955.
40. Chiricozzi A, Faleri S, Franceschini C, Caro RD, Chimenti S, Bianchi L. AISI: a new disease severity assessment tool for hidradenitis suppurativa. *Wounds*. 2015;27:258–264.
41. Goldfarb N, Ingram JR, Jemec GBE, et al. Hidradenitis Suppurativa Area and Severity Index (HASI): a pilot study to develop a novel instrument to measure the physical signs of hidradenitis suppurativa. *Br J Dermatol*. 2020;182:240–242.
42. Kirby JS, Butt M, King T. Severity and Area Score for Hidradenitis (SASH): a novel outcome measurement for hidradenitis suppurativa. *Br J Dermatol*. 2020;182:940–948.
43. Goldfarb N, Lowes MA, Butt M, King T, Alavi A, Kirby JS. Hidradenitis Suppurativa Area and Severity Index Revised (HASI-R): psychometric property assessment. *Br J Dermatol*. 2021;184:905–912.
44. Sartorius, Lapins J, Emtestam L, Jemec GB, et al.2003 Suggestions for uniform outcome variables when reporting treatment effects in hidradenitis suppurativa. *Br J Dermatol*. 2003;149:211–213.
45. 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;161:831–839.