CORRELATION OF THE REFINED HURLEY CLASSIFICATION FOR HIDRADENITIS SUPPURATIVA WITH PATIENTS’ REPORTED QUALITY OF LIFE AND OBJECTIVE DISEASE SEVERITY ASSESSMENT


Angelique Rondags¹
Kelsey R. van Straalen²
Jelmer R. van Hasselt¹
Ineke C. Janse¹³
Christine B. Ardon²
Allard R.J.V. Vossen²
Errol P. Prens²
Hessel H. van der Zee²
Barbara Horváth¹

¹. Department of Dermatology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands.
². Department of Dermatology, Erasmus Medical Center, Rotterdam, the Netherlands.
³. Department of Dermatology, Meander Medical Center, Amersfoort, the Netherlands.
ABSTRACT

Background: Hidradenitis suppurativa (HS) is a chronic, debilitating, heterogeneous disease requiring different treatment approaches. Recently, we refined the classic Hurley classification into a seven-stage classification in order to guide these treatment choices. This new classification subdivides Hurley stage I and II into three substages, namely mild (A), moderate (B) and severe (C) HS disease. Hurley stage III is not subcategorized and is always severe.

Objectives: To investigate the correlation between the given severity grades of Hurley I and Hurley II in the refined Hurley classification, and the patient-reported quality of life and physician-assessed objective severity score.

Methods: In this cross-sectional study, patients with HS participating in the observational cohorts of two Dutch tertiary referral centres were included before June 2017. The patient-reported Dermatology Life Quality Index (DLQI) and physician-assessed International HS Severity Score System (IHS4) scores were compared between the refined Hurley stages.

Results: In total, 433 patients were analysed. DLQI and IHS4 scores increased within Hurley stage I and II from A through C. There was a significant positive correlation of DLQI and IHS4 with increasing refined Hurley substages [refined Hurley stage I (A, B and C) to DLQI: $r_s = 0.259$, $P < 0.001$ and refined Hurley stage II (A, B and C) to DLQI: $r_s = 0.185$, $P = 0.010$; refined Hurley stage I (A, B and C) to IHS4: $r_s = 0.603$, $P < 0.001$ and refined Hurley stage II (A, B and C) to IHS4: $r_s = 0.532$, $P < 0.001$].

Conclusions: The refined Hurley classification accurately correlates with HS severity assessed by both patients and clinicians. Therefore, the refined Hurley classification is a useful tool for the quick assessment of severity in HS.
INTRODUCTION

Hidradenitis suppurativa (HS) is a chronic, inflammatory, recurrent, debilitating skin disease (of the terminal hair follicle) that usually presents after puberty with painful, deep-seated, inflamed lesions, most commonly in the axillary, inguinal and anogenital regions (modified, Dessau definition, first International Conference on Hidradenitis Suppurativa, 30 March to 1 April 2006, Dessau, Germany).1-4

HS can have a tremendous negative influence on quality of life (QoL), owing to its chronic, recurrent nature and painful inflammatory nodules and abscesses in intimate body regions, which can lead to scarring and disfigurement.5 The exact pathogenesis of HS remains to be elucidated and a cure does not yet exist.2-4,6

Hidradenitis suppurativa is a heterogeneous disease, and there are different stages within the HS disease spectrum that require different therapeutic approaches. However, an accurate classification and severity assessment to define these different stages is still lacking. As the original Hurley classification was conceived to describe HS severity in a single affected body area, for surgical purposes only, it does not take into account the extent of the disease and the degree of inflammation in the entire patient.7 However, it is still frequently improperly used as a global severity assessment tool. In order to better classify and match the therapeutic approach, a modification of the Hurley classification was proposed in 2016, named the refined Hurley classification.8

The refined Hurley classification aims to incorporate all fundamental aspects of the disease within a patient: the presence of sinus tracts, the number of affected body regions, and the degree of inflammation (Figure 1).8 Based on the extent and inflammatory component, three subcategories can be distinguished within refined Hurley stage I and II (A, B and C), that represent mild, moderate and severe HS disease. Hurley III was redefined but not subcategorized and is always severe.

Accordingly, the refined Hurley makes it possible for patients with Hurley stage I to have severe disease, based on the wide extent and high number of, especially migratory, inflammatory lesions. The refined Hurley classification enables the physician to quickly assess the severity of HS across different stages and helps to guide treatment, in particular whether surgery and/or anti-inflammatory treatment is indicated.8

In this study, we aim to investigate whether the refined Hurley classification accurately distinguishes three different severities of HS by correlating them with the patient-reported Dermatology Life Quality Index (DLQI) and physician-assessed International HS Severity Score System (IHS4).
Figure 1. The refined Hurley classification

HS, hidradenitis suppurativa; BSA, body surface area.

Previously published by Horváth et al. in Acta Derm Venereol 2017. Adapted with permission.

METHODS

Study population

In this multicentre study adult patients with a baseline visit before June 2017 within the Hidradenitis Suppurativa Registry (HiSURE) cohort of the University Medical Center Groningen (UMCG) and HiScreen cohort of the Erasmus Medical Center (EMC), both in the Netherlands, were included. Both cohorts are parallel longitudinal observational databases. Since 2015, adult patients who have adequate knowledge of the Dutch language, who visited the outpatient clinic of the dermatology department in the UMCG or EMC and have been diagnosed with HS by a dermatologist and are willing to participate, have been included and followed using a standardized protocol. The following patient characteristics were collected: sex, age, age of HS symptom onset, smoking status, body mass index (BMI), total DLQI score (range 0–30), and the refined Hurley stage. The IHS4 score was calculated for all patients using the raw data of presence of lesion types and counts.
Exclusion criteria were missing refined Hurley stage, missing components to derive the IHS4, and missing DLQI scores. For this type of study, a sample size calculation is not applicable. The HiSURE and HiScreen cohorts were approved by the local ethics committees of the UMCG and EMC, respectively. Medical ethical committee approval is not required for this type of analysis under Dutch law.

Data collection and comparison
The average DLQI scores and IHS4 scores of each refined Hurley stage were calculated. The refined Hurley classification is a seven-stage, discriminative classification system for patients with HS. In refined Hurley stage I and II, the letters A, B and C represent the severity grades mild, moderate and severe HS disease, respectively, and are based on the extent and the degree of inflammation of HS in the entire patient (Figure 1). The DLQI, a validated patient-reported dermatology-specific QoL questionnaire, consists of 10 questions covering six domains: symptoms and feelings, daily activities, leisure, work and school, personal relationships and treatment. A total score of 0–1 indicates no effect on QoL, 2–5 a small effect, 6–10 a moderate effect, 11–20 a very large effect and 21–30 an extremely large effect. The IHS4 is a physician-assessed dynamic HS severity tool, developed and validated by the European Hidradenitis Suppurativa Foundation Investigator group in 2017. The IHS4 score is the sum of the number of inflammatory nodules multiplied by 1; number of abscesses multiplied by 2; and number of draining tunnels (fistulae/sinuses) multiplied by 4. A score of ≤ 3 is considered mild, 4–10 moderate and ≥ 11 severe HS.

The following comparisons were made between the refined Hurley stages regarding DLQI and IHS4 scores: (i) differences in scores between Hurley stages, e.g. IA vs. IC and IIA vs. IIC; (ii) differences in scores between Hurley stages of the same severity category, e.g. IA vs. IIA and IC vs. IIC; (iii) correlation with the refined Hurley stage I (A, B and C) and stage II (A, B and C).

Statistical analysis
Statistical analysis was performed using IBM SPSS Statistics version 23.0 for Windows (SPSS Inc., Chicago, IL, U.S.A.). Results are presented as number of patients, n (%), mean ± SD for normally distributed data, and median (interquartile range) for non-normally distributed data. A Mann–Whitney U-test was performed to analyse differences between refined Hurley stages regarding DLQI and IHS4 scores. A Spearman correlation coefficient test was performed to analyse the correlation of the refined Hurley stage I (A, B and C) and stage II (A, B and C) to the DLQI and IHS4. A two-sided P-value < 0.05 was considered significant.
RESULTS

Patient characteristics
A total of 492 patients with HS were included from the combined cohorts. Fifty-nine patients were excluded because of missing data regarding refined Hurley stage, DLQI scores or data to calculate the IHS4 score, yielding 433 patients eligible for analysis: HiSURE 244 (56·4%) and HiScreen 189 (43·6%) patients. There was a female predominance (72·3%). Overall, 79·7% patients were current or former smokers, and the mean BMI was 28.5 ± 6·0 kg m\(^{-2}\) (Table 1). The median DLQI score was 10 (5·0–16·0) and IHS4 score 4·0 (1·0–11·0). The distribution of the refined Hurley classification showed that the majority of patients had stage IA (29·3%), followed by IIB (20·1%), IIC (18·0%), IC (11·3%), III (8·8%), IB (6·2%) and IIA (6·2%) disease (Table 2).

Correlation of refined Hurley categories to Dermatology Life Quality Index
Overall, the refined Hurley classification correlated well to the patient-reported DLQI per stage (Table 2, Figure 2).

Comparison of Dermatology Life Quality Index between refined Hurley stages (I and II)
Within refined Hurley stage I, the median DLQI scores increased from IA through IB to IC. There were significant differences in DLQI scores between stage IA and IC (\(P < 0.001\)) and between IB and IC (\(P = 0.047\)). Similar to this, within refined Hurley stage II median DLQI scores increased from IIA through IIB to IIC and a significant difference was found between stage IIA and IIC (\(P = 0.022\)) (Table 3, Figure 2).

Comparison of Dermatology Life Quality Index to the refined Hurley stages of the same severity category (A, B and C)
Comparison of the refined Hurley stages of the same severity grades mild (IA vs. IIA), moderate (IB vs. IIB) and severe (IC vs. IIC) showed no differences in median DLQI scores (\(P = 0.784; P = 0.582; \) and \(P = 0.956, \) respectively), indicating similar DLQI scores between the refined Hurley stages of the same severity category. Refined Hurley stage III showed the highest median DLQI score. However, this was not significantly different from stage IC and IIC (\(P = 0.095\) and \(P = 0.104, \) respectively), indicating similarities between stage IC, IIC and III regarding patient-reported QoL (Figure 2).
### Table 1. Patient characteristics of included patients with hidradenitis suppurativa (n = 433)

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female sex</td>
<td>313 (72.3)</td>
</tr>
<tr>
<td>Age, years</td>
<td>39.0 ± 12.4</td>
</tr>
<tr>
<td>Age of symptom onset, years</td>
<td>22.3 ± 10.2</td>
</tr>
<tr>
<td>Smoking status*</td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>231 (54.6)</td>
</tr>
<tr>
<td>Former</td>
<td>106 (25.1)</td>
</tr>
<tr>
<td>Never</td>
<td>86 (20.3)</td>
</tr>
<tr>
<td>BMI*, kg m⁻²</td>
<td>28.5 ± 6.0</td>
</tr>
<tr>
<td>DLQI, score (range 0–30)</td>
<td>10.0 (5.0–16.0)</td>
</tr>
<tr>
<td>IHS4, score</td>
<td>4.0 (1.0–11.0)</td>
</tr>
<tr>
<td>IHS4, severity</td>
<td></td>
</tr>
<tr>
<td>Mild</td>
<td>205 (47.3)</td>
</tr>
<tr>
<td>Moderate</td>
<td>114 (26.3)</td>
</tr>
<tr>
<td>Severe</td>
<td>114 (26.3)</td>
</tr>
</tbody>
</table>

Values are presented as n (%) patients, mean ± standard deviation, or median (interquartile range).

BMI, body mass index; DLQI, Dermatology Life Quality Index; IHS4, International Hidradenitis Suppurativa Severity Score System.

* Missing data for BMI in n = 51 (11.8%), smoking status n = 10 (2.3%).

### Table 2. Patient distribution (n = 433); DLQI and IHS4 per refined Hurley classification stage

<table>
<thead>
<tr>
<th>Refined Hurley classification, stage</th>
<th>Patients</th>
<th>Female sex</th>
<th>Age, years</th>
<th>Age of symptom onset, years</th>
<th>DLQI score</th>
<th>IHS4 score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage IA</td>
<td>127 (29.3)</td>
<td>105 (82.7)</td>
<td>37.3 ± 12.5</td>
<td>22.3 ± 10.5</td>
<td>7.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Stage IB</td>
<td>27 (6.2)</td>
<td>19 (70.4)</td>
<td>39.7 ± 11.9</td>
<td>27.2 ± 12.5</td>
<td>9.0</td>
<td>2.0</td>
</tr>
<tr>
<td>Stage IC</td>
<td>49 (11.3)</td>
<td>38 (77.6)</td>
<td>38.8 ± 11.9</td>
<td>19.0 ± 7.3</td>
<td>13.0</td>
<td>5.0</td>
</tr>
<tr>
<td>Stage IIA</td>
<td>27 (6.2)</td>
<td>21 (77.8)</td>
<td>37.7 ± 11.9</td>
<td>23.2 ± 10.1</td>
<td>9.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Stage IIB</td>
<td>87 (20.1)</td>
<td>56 (64.4)</td>
<td>38.8 ± 11.6</td>
<td>22.2 ± 9.3</td>
<td>10.0</td>
<td>7.0</td>
</tr>
<tr>
<td>Stage IIC</td>
<td>78 (18.0)</td>
<td>54 (69.2)</td>
<td>39.8 ± 12.2</td>
<td>22.3 ± 10.0</td>
<td>13.0</td>
<td>12.5</td>
</tr>
<tr>
<td>Stage III</td>
<td>38 (8.8)</td>
<td>20 (52.6)</td>
<td>44.2 ± 14.9</td>
<td>22.9 ± 12.2</td>
<td>16.5</td>
<td>20.0</td>
</tr>
</tbody>
</table>

Values are presented as n (%) patients, mean ± standard deviation, or median (interquartile range).

DLQI, Dermatology Life Quality Index; IHS4, International Hidradenitis Suppurativa Severity Score System.
There are significant differences in DLQI scores between stage IA and IC ($P < 0.001$), IB and IC ($P = 0.047$) and stage IIA and IIC ($P = 0.022$). There are no significant differences in DLQI scores between IA vs. IIA ($P = 0.784$), IB vs. IIB ($P = 0.582$) and IC vs. IIC ($P = 0.956$). The dotted arrows illustrate the positive correlation between the DLQI and refined Hurley stage I (A, B and C combined; $P < 0.001$) and refined Hurley stage II (A, B and C combined; $P = 0.010$).

**Table 3. Comparison of DLQI and IHS4 scores between the refined Hurley classification stages and severity grades**

<table>
<thead>
<tr>
<th>Refined Hurley classification, stages</th>
<th>Stages</th>
<th>Severity</th>
</tr>
</thead>
<tbody>
<tr>
<td>DLQI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IHS4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$P$-value</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$&lt; 0.001$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$0.001$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$0.01$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$0.02$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$0.05$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$0.1$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

DLQI, Dermatology Life Quality Index; IHS4, International Hidradenitis Suppurativa Severity Score System.
Correlation of the Dermatology Life Quality Index to the refined Hurley classification
There was a significant positive correlation between the DLQI and refined Hurley stage I (A, B and C combined), and refined Hurley stage II (A, B and C combined) ($r_s = 0.259, P < 0.001$ and $r_s = 0.185, P = 0.010$, respectively) (Figure 2).

Correlation of refined Hurley classification to physician-assessed severity measurement
International Hidradenitis Suppurativa Severity Score System
Overall, the refined Hurley classification correlated well to the median IHS4 scores per refined Hurley stage as shown in Table 2.

Comparison of International Hidradenitis Suppurativa Severity Score System between refined Hurley stages
Similar to the DLQI scores, the median IHS4 scores increased from stage IA through IB to IC and from stage IIA through IIB to IIC. Refined Hurley stage III showed the highest IHS4 score. There were significant differences in IHS4 scores between all seven refined Hurley stages (Table 3, Figure 3).

Comparison of the International Hidradenitis Suppurativa Severity Score System to the refined Hurley stages of the same severity category (A, B and C)
Regarding the severity grades, there was no difference in median IHS4 scores between the mild refined Hurley stages (IA vs. IIA, $P = 0.375$). For moderate (IB vs. IIB) and severe (IC vs. IIC) refined Hurley stages there were significant differences in median IHS4 scores (both $P < 0.001$) (Table 3, Figure 3).

Correlation of the International Hidradenitis Suppurativa Severity Score System to refined Hurley classification
A significant positive correlation was found between the IHS4 and refined Hurley stage I (A, B and C) and refined Hurley stage II (A, B and C) ($r_s = 0.603, P < 0.001$ and $r_s = 0.532, P < 0.001$, respectively) (Figure 3).

DISCUSSION

In this study, we investigated whether the three distinguished severity grades within Hurley stage I and II correlate with patient-reported outcome DLQI and the objective IHS4 scores. Our results show that there are increasing DLQI and IHS4 scores within refined Hurley stage I and II: both scores increased from A through B to C, and most of the
Figure 3. Distribution of the IHS4 scores per refined Hurley stage

There is no significant difference in median IHS4 scores between IA vs. IIA (P = 0.375). Between all other stages there are significant differences (all P < 0.001). The dotted arrows illustrate the positive correlation between the IHS4 and refined Hurley stage I (A, B and C combined; P < 0.001) and refined Hurley stage II (A, B and C combined; P < 0.001).

IHS4, International Hidradenitis Suppurativa Severity Score System.

*P < 0.001

scores were significantly different between the stages. Moreover, the DLQI scores between stage IA and IIA (both considered mild HS), IB and IIB (moderate disease) and IC, IIC and III (severe HS) did not differ statistically, indicating a comparable QoL impairment for severity categories between the stages. A significant positive correlation with the DLQI and IHS4 was observed in refined Hurley stage I (A, B and C combined) and Hurley stage II (A, B and C combined). Results from this study confirmed the construct validity of the severity subcategories of the refined Hurley classification. These results clearly demonstrate that the refined Hurley classification is able to define specific severity subtypes within Hurley I such as moderate (IB) or severe (IC), which were previously considered as mild cases. Moreover, within Hurley stage II also, a mild (IIA), moderate (IIB) and severe (IIC) patient population can be classified.
The DLQI is a commonly used questionnaire to measure the impact of HS on QoL.\textsuperscript{11-14} One study showed that patients with HS have the highest DLQI scores of almost all dermatological diseases.\textsuperscript{13} This impact is due to the painful and draining lesions, which can soil clothing, can be malodorous, and can be very itchy.\textsuperscript{3, 15} Furthermore, lesions are often located in intimate body areas such as the inguino-genital area and buttocks and may be unsightly and disfiguring, increasing feelings of embarrassment.\textsuperscript{3, 4} Thus, it is not surprising that HS can have a profound negative impact on a patient’s professional and private life and depression is a not uncommon comorbidity.\textsuperscript{4, 16}

However, previous data have shown that the DLQI score increases with the ‘classic’ Hurley stages.\textsuperscript{13, 14, 17} We hypothesized that the impact of HS on QoL is influenced by the number of lesions, the number (and area) of affected body regions, and the presence of inflammation, rather than by sinus tracts alone. The results from our study confirm this hypothesis. Therefore, the classic Hurley classification lacks important items to validly assess HS severity.

Although the IHS4 scores increased within refined Hurley stage I and II from severity subgroups A through C, IHS4 scores for moderate (IIB and IIB) and severe (IC and IIC) were different. This difference is a result of the construction of the IHS4 score. A single draining tunnel contributes four times more than an inflammatory nodule, and two times more than an abscess, resulting in higher scores for patients at refined Hurley stage IIB, IIC and III.\textsuperscript{10} Hurley stage IA and IIA did show similar IHS4 scores, because in both stages the IHS4 score is determined by the number of inflammatory nodules and abscesses only; draining (inflammatory) sinus tracts are not present in both stage IA and IIA (only nondraining in IIA).

Summarizing, the refined Hurley classification allows the identification of specific severity subtypes of HS. For example, there is a distinct severe HS subtype (Hurley stage IC) within the classical Hurley stage I, previously defined as mild. These patients experience a high burden of disease, which is reflected by the high DLQI scores that are comparable with refined Hurley stage IIC and III as seen in this study. Regarding therapy, surgery is not an appropriate option for these migratory lesions and antibiotics might be inefficient. Therefore, by recognizing these patients as having severe HS, they are eligible for treatment with biologics (e.g. tumour necrosis α inhibitors). We hypothesize that acknowledgement of these subpopulations within HS contributes to a better understanding of the disease and more appropriate treatment decisions and outcomes in HS.

Our study does have some limitations. The number of patients in each refined Hurley stage differed: stage IA had the highest number of patients (n = 127) and stage IB and IIA the lowest (both n = 27). Furthermore, this study was performed with patients visiting tertiary referral centres for HS; thus, it is possible that the median DLQI and IHS4 are
higher than in the general HS population. However, the median DLQI score in this study was similar to other publications. In addition, the patient characteristics in this study are comparable with the general HS population, meaning our results could be extrapolated to the general HS population. Another point of discussion is that there is currently no validated HS disease-specific QoL measurement tool.

We have shown that in the refined Hurley classification stage I and II, three different subclasses of severity are distinguishable in both patient-reported outcomes and physician-reported objective levels. Therefore, we conclude that the refined Hurley classification accurately indicates the severity of HS and thus seems valid.

**Acknowledgments**

We thank all the patients with HS who are participating in the HiSURE and HiScreen cohorts.
References
