

# Responsiveness and interpretability of the Hand Eczema Severity Index\*

J.A.F. Oosterhaven  and M.L.A. Schuttelaar

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

**Linked Comment:** Howells. *Br J Dermatol* 2020; **182**:828–829.

## Summary

### Correspondence

Marie Louise Schuttelaar.

E-mail: m.l.a.schuttelaar@umcg.nl

### Accepted for publication

28 June 2019

### Funding sources

None.

### Conflicts of interest

None to declare.

\*Plain language summary available online

DOI 10.1111/bjd.18295

**Background** The Hand Eczema Severity Index (HECSI) is used to measure the severity of hand eczema. It is unclear what HECSI scores mean and what is the minimally important change (MIC). Furthermore, its responsiveness has not been studied.

**Objectives** To study the responsiveness and interpretability of the HECSI.

**Methods** This was a prospective study covering two time points: baseline and after 4–12 weeks. Responsiveness was assessed using a criterion approach, with at least one-step improvement on the ‘Photographic guide for severity of hand eczema’ as the anchor for important improvement. Interpretability of single scores was determined by defining severity bands based on agreement with the anchor. For change scores, the smallest detectable change (SDC) was calculated in patients indicating no change and the MIC was obtained in patients indicating that they had changed using three methods: mean cutoff, receiver operating characteristic (ROC) curve and 95% limit.

**Results** In total, 294 patients participated (160 male, mean age 44.9 years). HECSI scores improved or deteriorated in parallel with the anchor. The area under the ROC curve was 0.86 (95% confidence interval 0.81–0.91). The final severity band for single scores had a  $\kappa$ -coefficient of agreement of 0.694: clear, 0; almost clear, 1–16; moderate, 17–37; severe, 38–116; very severe,  $\geq 117$ . The SDC in 93 unchanged patients was 40.3 points. The obtained MIC values were all smaller than the SDC.

**Conclusions** The HECSI has good responsiveness. This study gives meaning to HECSI scores, which can be applied to clinical decision making and the design of clinical trials. We recommend that an improvement of 41 points on the HECSI is regarded as the minimally measurable true change.

### What’s already known about this topic?

- The Hand Eczema Severity Index (HECSI) is one of the most widely used measurement instruments to measure the severity of hand eczema.
- The HECSI has good reliability, but its responsiveness and interpretability have not been studied.

### What does this study add?

- This study shows good responsiveness of the HECSI.
- A severity grading for single scores is proposed.
- The smallest detectable change and minimally important change for improvement are determined.

### What are the clinical implications of this work?

- The obtained values help to interpret HECSI scores in daily practice and clinical studies, and may facilitate researchers in the calculation of sample sizes for clinical trials.

Multiple scoring systems exist for determining and monitoring the severity of hand eczema, but proper validation of most measurement properties is still lacking for many instruments.<sup>1</sup> The Hand Eczema Severity Index (HECSI) is one of the most widely used measurement instruments for hand eczema severity and its reliability has been studied by its developers.<sup>2</sup> However, the responsiveness (or sensitivity to change) and interpretability of the HECSI are not known.<sup>3</sup> A study from 2009 presented cutoff values for mild (0–11 points), moderate (12–27 points) and severe ( $\geq 28$  points) disease, although no reference was given.<sup>4</sup> These values most likely came from practical experience, not from a formally conducted interpretability study. Knowledge of whether an instrument can identify changes over time and an understanding of what single scores and change scores mean are vital to the proper use of such an instrument, both in clinical studies and in daily practice.<sup>5</sup> Therefore, in this study, we aimed to determine the responsiveness and interpretability of the HECSI.

## Patients and methods

### Study population and design

This prospective study was performed at the dermatology department of the University Medical Center Groningen, a tertiary referral centre for hand eczema. Adult patients (age  $\geq 18$  years) were included if they had hand eczema of at least 1-week duration, as diagnosed by a dermatologist according to current guidelines.<sup>6,7</sup> Patients were scored with the HECSI and an anchor for hand eczema severity (a 'photographic guide for assessing severity of chronic hand dermatitis')<sup>8</sup> at the outpatient clinic at two time points: at baseline ( $T_0$ ) and after 4–12 weeks ( $T_1$ ). Between  $T_0$  and  $T_1$ , patients were allowed to use any form of treatment. Recruitment was done consecutively between March 2017 and December 2018. The design of this study was in accordance with guidelines by the COSMIN group.<sup>5,9</sup> The institutional review board of the University Medical Center Groningen confirmed that this study did not fall under the scope of the Medical Research Involving Human Subjects Act (reference: METc 2014/391).

### Studied instrument

The HECSI is a physician-rated instrument covering six signs: erythema, infiltration/papulation, vesicles, fissures, scaling and oedema; and five anatomical locations: fingertips, fingers

(except tips), palm of hands, back of hands and wrists. Each sign is scored on each location with a score of 0–3, with 3 representing the most severe expression of the sign. For each location, the extent of total eczema involvement is scored 0–4 (0, 0%; 1, 1–25%; 2, 26–50%; 3, 51–75% and 4, 76–100%). The total sign score is multiplied by the extent score to obtain a total score, ranging from 0 to 360 points. The reliability of the HECSI was found to be good to excellent.<sup>2</sup>

### Anchor (external criterion)

The 'photographic guide for assessing severity of chronic hand dermatitis' (or 'Photoguide') was used as the anchor to determine the severity of hand eczema and to define which patients were changed or unchanged at  $T_1$ . It covers a five-point scale (clear, almost clear, moderate, severe, very severe), with photographic images depicting increasing severity and extent of hand eczema. It is a global assessment that incorporates a combination of clinical symptoms and extent. Its reliability was deemed adequate in a previous study.<sup>8</sup> In the current study, the Photoguide was scored by trained healthcare personnel.

### Responsiveness of the Hand Eczema Severity Index

To determine the responsiveness of the HECSI we chose a criterion approach, using the Photoguide as the anchor for change in hand eczema severity. We a priori hypothesized that change scores of the HECSI would correlate strongly (Spearman's rho  $> 0.7$ ) with changes on the anchor (Photoguide) and that mean change scores of the HECSI would increase or decrease in parallel with each step of improvement or deterioration on the Photoguide. Furthermore, we constructed a receiver operating characteristic (ROC) curve and assessed the area under the curve (AUC) to determine the discriminate ability of HECSI change scores. A dichotomized variable was computed including improved vs. unchanged patients on the Photoguide. Improvement was defined as at least one-step improvement. Patients whose condition deteriorated were not used in this analysis. We hypothesized that the AUC of the ROC curve would be  $> 0.7$ . If our a priori hypotheses were confirmed, we would consider the responsiveness of the HECSI to be good.<sup>5</sup>

### Interpretability of the Hand Eczema Severity Index

We studied the interpretability of the HECSI for use in single scores (e.g. 'what does a HECSI score of ... mean?') and in

change scores (e.g. 'what is an important difference in HECSI scores?').

### Single scores

We used an anchor-based method to define severity strata (or 'bands') for the HECSI with scores obtained at  $T_0$ , using the Photoguide as the severity anchor. Several numerical cutoffs of the HECSI score were tested against the severity categories of the Photoguide using a weighted kappa coefficient of agreement to determine the highest level of agreement between these. Numerical cutoff points were considered based on HECSI scores that corresponded to a one-step increase in mean, median and/or mode on the Photoguide. Sensitivity analyses consisted of tests for differences in sex and age distribution between patients for whom severity could be predicted based on the band with the highest  $\kappa$ -value and patients for whom scores disagreed with the predicted severity according to that band.<sup>10–12</sup>

### Change scores

The smallest detectable change (SDC) is defined as 'the smallest change in score that can be detected by the instrument, beyond measurement error'.<sup>5</sup> The SDC for the HECSI was determined in patients indicating no change at  $T_1$ , as identified using the Photoguide. For this, the standard error of measurement ( $SEM_{\text{agreement}}$ ) was obtained using the square root of the within-person total variance of an ANOVA analysis. The SDC was then calculated using the formula:  $SDC = 1.96 \times \sqrt{2} \times SEM_{\text{agreement}}$ .<sup>13</sup> This was done for the whole HECSI and also for HECSI scores within the severity bands that were determined using the Photoguide in the single-score analysis.

The minimally important change (MIC) is defined as 'the smallest change in the construct to be measured which is perceived as important' (by patients and/or clinicians).<sup>5</sup> The MIC for the HECSI score was determined in patients indicating that they had changed at  $T_1$ , as identified using the Photoguide. Raw change scores were calculated for the HECSI and Photoguide by subtracting the score at  $T_0$  from the score at  $T_1$ . Also, the percentage change compared with baseline was calculated using the formula:  $[(\text{HECSI } T_0 - \text{HECSI } T_1) / \text{HECSI } T_0] \times 100$ . Thus, negative scores and percentages correspond to a deterioration, and positive scores and percentages to an improvement of the hand eczema. We determined a change of one step in severity on the Photoguide as the cutoff for an important improvement. Three MIC values were determined for both raw change scores and percentage change compared with baseline, as follows:<sup>14,15</sup> (i) The MIC based on the mean change in HECSI value or percentage that corresponds with a one-step severity change on the Photoguide; (ii) the MIC of the ROC cutoff point, indicating the point closest to the upper left corner, where the sum of percentages of correctly classified patients is highest; (iii) the MIC based on the 95% upper-limit cutoff point of the patients indicating no change or no important change, which corresponds to  $\text{mean}_{\text{change}} + 1.645 \times SD_{\text{change}}$  of this group (or strictly to the  $\text{mean}_{\text{difference}}$  and  $SD_{\text{difference}}$  as this concerns patients indicating no change).

Values for the MIC of the raw HECSI change score were graphically presented using the visualized anchor-based MIC distribution method.<sup>15</sup> We determined the MIC only for patients indicating that they had improved. The MIC for deterioration was not feasible to assess, because too few patients ( $n < 50$ ) indicated that they had deteriorated compared with baseline to draw sound conclusions about this.

### Floor and ceiling effects

It was determined what proportion of patients achieved the highest and lowest possible scores on the HECSI. Floor and ceiling effects were considered to be present if the lowest or highest HECSI score was achieved by  $> 15\%$  of patients.<sup>16</sup>

### Data analysis

No sample-size calculation was performed. A general recommendation for interpretability studies is to use a minimum of 50 participants, but preferably at least 100, with a minimum of 50 in the smallest subgroup to calculate the MIC using the ROC method.<sup>5</sup> This study meets these recommendations. Spearman's rho was used to calculate correlations. The  $\chi^2$ -test and Student's t-test were used to calculate differences between groups. There were no missing values for the analysed patients. Analyses were performed with IBM SPSS Statistics for Windows, version 23.0 (IBM, Armonk, NY, U.S.A.) and GraphPad Prism version 7.02 for Windows (GraphPad Software, La Jolla, CA, U.S.A.).

### Subgroup analyses

The MIC has been found to vary depending on the baseline score.<sup>17</sup> Therefore, we performed a sensitivity analysis on the MIC, calculating it also for subgroups (low or high) based on the median of the baseline HECSI scores.

### Results

At baseline, 294 patients were included; their basic characteristics are presented in Table 1. At  $T_1$ , 215 patients were available for assessment. In the group of 79 patients who dropped out there were significantly more women (35% vs. 20%;  $\chi^2 = 8.43$ ,  $P < 0.01$ ) and they were significantly younger (mean 41.8 vs. 46.0 years;  $P < 0.05$ ).

### Responsiveness

There was a strong, significant correlation between the change on the Photoguide and the change on the HECSI: Spearman's rho = 0.75 ( $P < 0.001$ ), which was not significantly affected by age or sex. Table 2 shows the mean HECSI scores improving or deteriorating in parallel with the anchor. The AUC under the ROC curve was 0.86 (95% confidence interval 0.81–0.91) (Fig. 1). These results meet our hypotheses defined a priori; the HECSI showed good responsiveness.

**Table 1** Basic characteristics of the study population. There were no significant differences between male and female participants

	Male (n = 160)	Female (n = 134)	Total (n = 294)
Age (years)			
Mean $\pm$ SD	45.0 $\pm$ 14.5	44.8 $\pm$ 17.5	44.9 $\pm$ 15.9
Range	18–74	18–83	18–83
HECSI			
Mean $\pm$ SD	48.6 $\pm$ 41.1	40.9 $\pm$ 35.5	45.1 $\pm$ 38.8
Median (IQR)	36.0 (16.0–66.5)	29.5 (14.0–55.3)	33.5 (15.0–62.0)
Range	3–192	2–144	2–192
Photoguide severity (physician), n (%)			
Clear	0	0	0
Almost clear	37 (23.1)	40 (29.9)	77 (26.2)
Moderate	63 (39.4)	51 (38.1)	114 (38.8)
Severe	44 (27.5)	37 (27.6)	81 (27.6)
Very severe	16 (10.0)	6 (4.5)	22 (7.5)

HECSI, Hand Eczema Severity Index; IQR, interquartile range.

**Table 2** The mean change scores on the Hand Eczema Severity Index (HECSI), according to changes measured with the anchor (Photoguide)

Photoguide	HECSI				
Perceived change	n	Mean change	SD <sub>change</sub>	% change	SD <sub>% change</sub>
4-step improvement	0	NA	NA	NA	NA
3-step improvement	4	+71.5	42.9	+86%	12%
2-step improvement	18	+53.2	35.3	+69%	30%
1-step improvement (MIC)	65	+30.2	26.2	+59%	25%
No change	93	+1.2	20.6	–5% <sup>a</sup>	56% <sup>a</sup>
1-step deterioration	32	–23.3	19.6	–83%	93%
2-step deterioration	3	–68.3	55.2	–326%	194%
3-step deterioration	0	NA	NA	NA	NA

MIC, minimally important change; NA, not applicable. <sup>a</sup>One extreme outlying case was excluded (–300%).

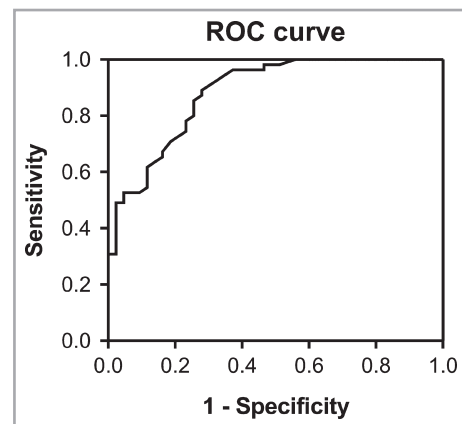
## Interpretability

### Single scores

There was a strong, significant correlation between the Photoguide and the HECSI: Spearman's  $\rho = 0.84$  ( $P < 0.001$ ), which was not significantly affected by age or sex. The distribution of HECSI scores stratified by Photoguide categories is shown in Figure 2. In total, 36 bands for severity were tested. Details of the testing of single scores are presented in Tables S1 and S2 (see Supporting Information). The band with the highest  $\kappa$ -coefficient of agreement ( $\kappa = 0.694$ ) was chosen as the final severity band: clear, 0; almost clear, 1–16; moderate, 17–37; severe, 38–116; very severe,  $\geq 117$ .

### Overview of Hand Eczema Severity Index scores falling outside the proposed banding

One patient (0.3%) had a Photoguide score  $> 1$  point outside that predicted by the final HECSI severity band. There were 57 patients (19.4%) patients with an actual Photoguide score 1


**Fig 1.** Receiver operating characteristic (ROC) curve for various cutoff points for change on the raw Hand Eczema Severity Index score.

point lower than the final HECSI severity band predicted. There were 33 patients (11.2%) with an actual Photoguide score 1 point higher than the final HECSI severity band predicted. There were no sex or age distribution differences

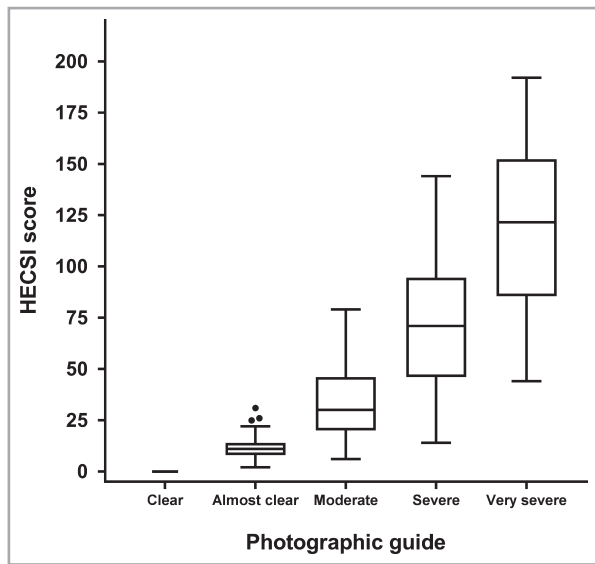


Fig 2. Box plots of the distribution of the Hand Eczema Severity Index (HECSI) by Photographic guide severity. Error bars are according to Tukey. Not the whole range of the HECSI is shown (0–360).

between the patients falling within the proposed banding and those falling outside it.<sup>12</sup>

**Change scores**

**Smallest detectable change.** The SDC of the HECSI was calculated based on the analysis of 93 patients indicating no change on the Photoguide between T<sub>0</sub> and T<sub>1</sub>. Its value was 40.3 points. Subsequently, the SDCs for the subgroups of severities were determined: almost clear (n = 26), 13.2 points; moderate (n = 37), 32.9 points; severe (n = 26), 58.1 points; very severe (n = 4), 67.2 points.

**Minimally important change.** The correlation between the change in HECSI score and the change in the anchor (rho = 0.75) is higher than the minimally recommended correlation of 0.3–0.5 for studying the MIC. Therefore, the Photoguide was considered to be an appropriate anchor.<sup>18,19</sup> Calculation of the MIC of improvement according to our three used methods resulted in the values shown in Table 3. The distribution of raw HECSI

Table 3 Minimally important change (MIC) for improvement on the Hand Eczema Severity Index (HECSI) using three methods

MIC method	Total HECSI		Low baseline HECSI scores		High baseline HECSI scores	
	Points	%	Points	%	Points	%
Mean cutoff	30.2	59	10.7	62	46.9	57
ROC curve	10.5	31	5.5	38	19.0	31
95% limit	35.1	87	16.3	95	50.1	71

MIC values are presented for the total HECSI, and for low (≤ 33) and high (> 33) scores. MIC values for percentage decrease from baseline (%) are presented. ROC, receiver operating characteristic.

change scores was visualized as the anchor-based distribution for patients indicating that they had improved or indicating no change, along with the three MIC values (Fig. 3).

**Sensitivity analysis for low and high baseline scores**

The median value of the HECSI at baseline was 33.5 points. Therefore, MIC values were separately calculated for patients with a low baseline HECSI score (≤ 33 points) and patients with a high baseline HECSI score (> 33 points). Details of the MIC values are provided in Table 3 (and Tables S3 and S4 and Figs S1 and S2; see Supporting Information).

**Floor and ceiling effects**

No floor or ceiling effects were seen: no patients achieved the highest HECSI score of 360 and only three patients achieved the lowest score of 0 at T<sub>1</sub>.

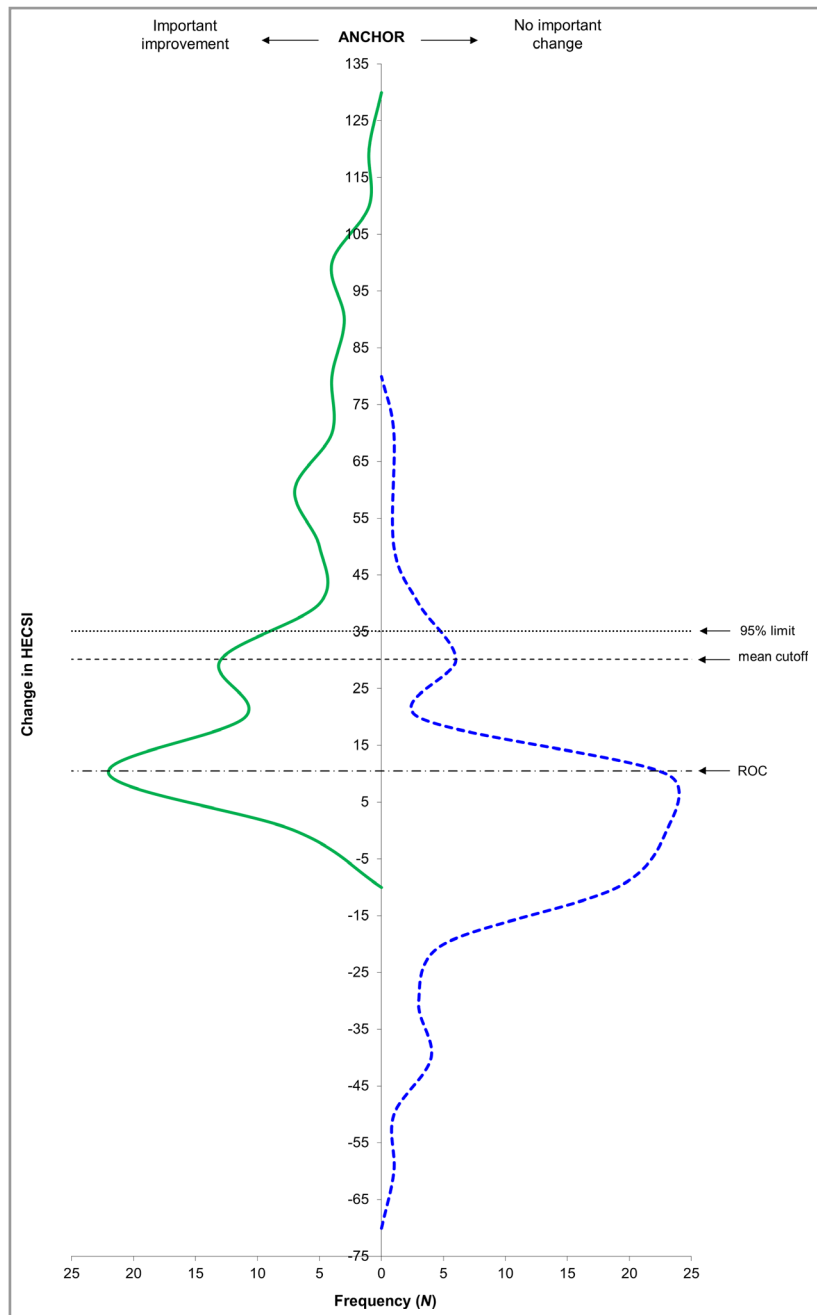
**Discussion**

In this study we found good responsiveness of the HECSI and report values to aid interpretability of HECSI scores.

For interpretability of single scores, the anchor-based approach gives an idea of what level of severity is represented by a HECSI score. It is clear that the distribution of scores is quite unequal: as severity rises, a gradually larger range of HECSI points covers the severity categories. This corresponds to what is found for the Eczema Area and Severity Index, a very similar instrument used to measure severity in atopic dermatitis.<sup>12</sup>

The anchor used in this study, the Photoguide, is a global instrument that crudely measures hand eczema severity. The high correlation between the Photoguide and the HECSI, along with detailed knowledge of both instruments, gives sufficient evidence that they measure many similar aspects of the assessed construct: hand eczema severity. Therefore, the Photoguide is a suitable anchor for the HECSI for single scores. However, when longitudinal measurements are performed, some aspects need consideration.

To determine the SDC, patients indicating no change were identified using the ordinal Photoguide. However, patients can change substantially on the HECSI while their baseline Photoguide category does not change, particularly within the ‘severe’ and ‘very severe’ categories. This inflates the SDC because of the large variance in HECSI scores that is found within the group of patients indicating no change after 4–12 weeks (T<sub>1</sub>). In many cases this results in an SDC larger than the MIC, which makes the MIC meaningless because it then falls within the range of values likely to be measurement error. Furthermore, the large SDC is problematic because it means that there may be a small but clinically meaningful change likely to be taking place, but the HECSI is unable (according to the currently used anchor) to detect this change beyond measurement error, and therefore it cannot detect these small but clinically important changes.



**Fig 3.** Visual anchor-based distribution of raw Hand Eczema Severity Index (HECSI) change scores for patients with improved (green line) and unchanged (blue dashed line) scores on the anchor (Photoguide), along with the minimally important change (MIC) values obtained using three different methods. ROC, receiver operating characteristic.

For the change scores, a change of one step in severity on the Photoguide as the cutoff for important change is an arbitrary choice, which may be valid in hand eczema of lower severities. However, when hand eczema is for example ‘very severe’, a decrease in HECSI score that does not correspond to a one-step decrease on the Photoguide may still count as a MIC for a patient. This would make the MIC even smaller than the values that we found. To determine this, we would have had to use a different anchor, for example asking patients whether they would rate their change as *important*. This would have been a more patient-guided MIC, while we chose a clinician’s perspective on the MIC. From that perspective, the one-step improvement on the Photoguide was the smallest possible change to

measure. In future studies, it would be interesting to see SDC and MIC values obtained with a patient-guided anchor, possibly also for higher and lower baseline severities separately.

Of the three MIC values obtained with the three methods, the MIC obtained with the ROC curve may be the most valid. The mean change score method does not actually represent a *minimally important difference*, but more of a mean *clear clinical change*, which inflates it compared with the cutoff point between patients indicating no change and those indicating that they had changed.<sup>20</sup> The 95% limit method is based on patients indicating no change, for whom the crude ordinal categories of the Photoguide result in large SDs of the HECSI score and an inflation of this method. Regardless, one should

realize that all three mean MIC values fall within the limits of agreement of the HECSI (at least when the Photoguide is used as anchor), as currently represented by the SDC of 40.3 points. Therefore, we recommend to regard an improvement of 41 points on the HECSI as the minimally measurable true change.

The MIC and SDC values obtained in this study can help interpret change scores on the HECSI and might facilitate researchers in the calculation of sample sizes for clinical trials. It is advised to use the MIC values expressing percentage change for this, as MIC values are more stable this way, especially when there is much variability in baseline values.<sup>21</sup> However, one should keep in mind that the MIC corresponding to a one-step improvement on the Photoguide is probably not always a sufficient clinical improvement when comparing treatments.<sup>22</sup> Various studies in hand eczema have defined 'response to treatment' as an improvement of at least two steps or achieving 'clear' or 'almost clear' on the used ordinal scale (Photoguide, or other Physician's Global Assessment).<sup>23,24</sup> To reach this Photoguide stage with the HECSI, an estimated reduction of 54 points or 70% is needed, based on the results of the current study (Table 2). Therefore, we suggest to use  $\geq 75\%$  improvement in HECSI score (HECSI 75) to define a truly, clinically relevantly improved patient. This corresponds to how the measurement of atopic dermatitis is currently reported with  $\geq 75\%$  improvement in the Eczema Area and Severity Index. Of note, although HECSI 75 can be reported to aid interpretation, it should not necessarily be used for analysis and trial design, as it turns the HECSI score into a binary outcome, which is less efficient than using the HECSI in its original continuous form.<sup>25</sup> Reporting HECSI 50 is not useful, as a 50% improvement is below all mean MIC values found in the current study.

A strength of this study is that it was conducted according to guidelines by COSMIN.<sup>5,9</sup> Furthermore, we believe that the results are generalizable to the white population of adult patients with hand eczema (the used instruments are mainly developed for use in white populations), as we included patients from all severity groups. However, a degree of inter- and intrarater variability exists,<sup>2</sup> so the MIC values that we report are in part dependent on how HECSI is scored in our centre. To improve this, a detailed, consensus-based scoring instruction for the HECSI should be made available in the future.<sup>26</sup> A limitation of the study is that the subgroup analyses for low and high HECSI baseline values did not include the minimum of 50 patients in each group of patients indicating that they had improved or indicating no change. This could have influenced these results.

The values presented in this study provide a sense of how HECSI scores can be interpreted. This should give researchers a better understanding of the meaning of HECSI scores in daily practice and stimulate its use in clinical trials. Furthermore, the study adds information on responsiveness as a measurement property of the HECSI, establishing a firmer base on which the HECSI may be adopted in a core outcome set for hand eczema studies in the future.<sup>27</sup>

## References

- Weistenhofer W, Baumeister T, Drexler H, Kutting B. An overview of skin scores used for quantifying hand eczema: a critical update according to the criteria of evidence-based medicine. *Br J Dermatol* 2010; **162**:239–50.
- Held E, Skoet R, Johansen JD, Agner T. The Hand Eczema Severity Index (HECSI): a scoring system for clinical assessment of hand eczema. A study of inter- and intraobserver reliability. *Br J Dermatol* 2005; **152**:302–7.
- Charman CR, English JS. Getting to grips with hand eczema: measuring skin disease severity objectively. *Br J Dermatol* 2005; **152**:199–201.
- Hald M, Agner T, Blands J *et al.* Clinical severity and prognosis of hand eczema. *Br J Dermatol* 2009; **160**:1229–36.
- de Vet HCW, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine*, 1st edn. New York: Cambridge University Press, 2011.
- Menné T, Johansen JD, Sommerlund M, Veien NK. Hand eczema guidelines based on the Danish guidelines for the diagnosis and treatment of hand eczema. *Contact Dermatitis* 2011; **65**:3–12.
- Diepgen TL, Andersen KE, Chosidow O *et al.* Guidelines for diagnosis, prevention and treatment of hand eczema. *J Dtsch Dermatol Ges* 2015; **13**:e1–22.
- Coenraads PJ, Van Der Walle H, Thestrup-Pedersen K *et al.* Construction and validation of a photographic guide for assessing severity of chronic hand dermatitis. *Br J Dermatol* 2005; **152**:296–301.
- Mokkink LB, Terwee CB, Patrick DL *et al.* The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010; **63**:737–45.
- Hongbo Y, Thomas CL, Harrison MA *et al.* Translating the science of quality of life into practice: what do Dermatology Life Quality Index scores mean? *J Invest Dermatol* 2005; **125**:659–64.
- Charman CR, Venn AJ, Ravenscroft JC, Williams HC. Translating Patient-Oriented Eczema Measure (POEM) scores into clinical practice by suggesting severity strata derived using anchor-based methods. *Br J Dermatol* 2013; **169**:1326–32.
- Leshem YA, Hajar T, Hanifin JM, Simpson EL. What the Eczema Area and Severity Index score tells us about the severity of atopic dermatitis: an interpretability study. *Br J Dermatol* 2015; **172**:1353–7.
- de Vet HC, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. *J Clin Epidemiol* 2006; **59**:1033–9.
- van der Roer N, Ostelo RW, Bekkering GE *et al.* Minimal clinically important change for pain intensity, functional status, and general health status in patients with nonspecific low back pain. *Spine (Phila PA 1976)* 2006; **31**:578–82.
- de Vet HC, Ostelo RW, Terwee CB *et al.* Minimally important change determined by a visual method integrating an anchor-based and a distribution-based approach. *Qual Life Res* 2007; **16**:131–42.
- McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? *Qual Life Res* 1995; **4**:293–307.
- de Vet HCW, Foumani M, Scholten MA *et al.* Minimally important change values of a measurement instrument depend more on baseline values than on the type of intervention. *J Clin Epidemiol* 2015; **68**:518–24.
- Cella D, Hahn EA, Dineen K. Meaningful change in cancer-specific quality of life scores: differences between improvement and worsening. *Qual Life Res* 2002; **11**:207–21.
- Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol* 2008; **61**:102–9.

- 20 Schram ME, Spuls PI, Leeflang MMG *et al.* EASI, (objective) SCORAD and POEM for atopic eczema: responsiveness and minimal clinically important difference. *Allergy* 2012; **67**: 99–106.
- 21 Farrar JT, Young JP, LaMoreaux L *et al.* Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain* 2001; **94**:149–58.
- 22 Roos EM. 3 steps to improve reporting and interpretation of patient-reported outcome scores in orthopedic studies. *Acta Orthop* 2018; **89**:1–2.
- 23 Ruzicka T, Lynde CW, Jemec GB *et al.* Efficacy and safety of oral alitretinoin (9-*cis* retinoic acid) in patients with severe chronic hand eczema refractory to topical corticosteroids: results of a randomized, double-blind, placebo-controlled, multicentre trial. *Br J Dermatol* 2008; **158**:808–17.
- 24 Oosterhaven JAF, Schuttelaar MLA. Study protocol: efficacy of oral alitretinoin versus oral cyclosporine A in patients with severe recurrent vesicular hand eczema (ALICsA): a randomised prospective open-label trial with blinded outcome assessment. *BMJ Open* 2018; **8**:e020192.
- 25 Altman DG, Royston P. The cost of dichotomising continuous variables. *BMJ* 2006; **332**:1080.
- 26 Castorr AH, Thompson KO, Ryan JW *et al.* The process of rater training for observational instruments: implications for interrater reliability. *Res Nurs Health* 1990; **13**:311–18.
- 27 Rönisch H, Apfelbacher C, Brans R *et al.* Which outcomes have been measured in hand eczema trials? A systematic review. *Contact Dermatitis* 2019; **80**:201–7.

## Supporting Information

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

**Fig S1.** Receiver operating characteristic curve for various cutoff points for change on the raw Hand Eczema Severity Index (HECSI) score for patients with a low baseline HECSI score ( $\leq 33$ ).

**Fig S2.** Receiver operating characteristic curve for various cutoff points for change on the raw Hand Eczema Severity Index (HECSI) score for patients with a high baseline HECSI score ( $> 33$ ).

**Table S1.** Number of patients with each Hand Eczema Severity Index score.

**Table S2.** Kappa coefficients of agreement for different proposed sets of Hand Eczema Severity Index severity bands.

**Table S3.** The mean change scores on the Hand Eczema Severity Score (HECSI), according to changes measured with the anchor (Photoguide) for patients with a low baseline HECSI score ( $\leq 33$ ).

**Table S4.** The mean change scores on the Hand Eczema Severity Score (HECSI), according to changes measured with the anchor (Photoguide) for patients with a high baseline HECSI score ( $> 33$ ).