Acta Derm Venereol 2016; 96: 978–980

SHORT COMMUNICATION

Itch Assessment with Visual Analogue Scale and Numerical Rating Scale: Determination of Minimal Clinically Important Difference in Chronic Itch

Adam Reich[1#], Claudia Riepe[2#], Zografia Anastasiadou[3], Karolina Mędrek[1], Matthias Augustin[3], Jacek C. Szepietowski[1] and Sonja Ständer[2]

[1]Department of Dermatology, Venereology and Allergology, Wroclaw Medical University, Ul. Chalubinskiego 1, 50-368 Wroclaw, Poland, [2]Center for Chronic Pruritus, Department of Dermatology, University Hospital of Münster, Münster, and [3]Institute for Health Services Research in Dermatology and Nursing, University Medical Center Hamburg, Hamburg, Germany. E-mail: adam.reich@umed.wroc.pl

[#]These authors contributed equally to this study.

Accepted Apr 6, 2016; Epub ahead of print Apr 12, 2016

Unidimensional scales, such as the visual analogue scale (VAS) and the numerical rating scale (NRS), have for many years been applied in the evaluation of itch intensity, showing good construct validity, reproducibility and reliability, as well as a high correlation of achieved results (1, 2). In order to develop these tools further we founded a Special Interest Group: "Scoring itch in clinical trials" of the International Forum for the Study of Itch (www.itchforum.net). In a pilot study we determined the cut-off values of VAS in order to aid to the understanding of particular VAS scoring (2). It was provisionally suggested that these cut-offs should be utilized for itch assessment in clinical trials (3). Other researchers have tested that VAS is a valid method of itch evaluation and that the results are reproducible even in different populations and ethnic groups (4).

The question remains, however, as to whether and how these itch tools measure a clinical benefit or worsening of the symptom. A construct to determine this is the minimal clinically important difference (MCID), which is the smallest patient-reported outcome change that can be clinically detected by the patient and is considered as clinically meaningful (5).

The aim of this study was to investigate the MCID for the VAS and NRS for itch. The results should enable improved clinical care of patients, as well as improved statistical power calculation of study populations in clinical trials related to itch.

MATERIALS AND METHODS

Patients with pruritus were recruited consecutively from the cohorts of patients admitted to our departments for diagnostics and treatment of skin diseases and/or chronic itch. Inclusion criteria were: informed consent obtained from a patient to participate in the study; age over 18 years; presence of chronic itch (≥6 weeks) during consultation with the dermatologist; and neither cognitive nor motor problems that might preclude patients from understanding the scale or marking the line with a pen. A total of 398 Caucasian patients were included in the study. Basic characteristics of the patients are shown in Table I.

This was a prospective, longitudinal, observational study with 2 parts: 1 in Poland and 1 in Germany. Each study part was approved by local ethics committees. At baseline all subjects underwent a careful history and physical examination, including assessment of itch intensity. Next, patients received a standard of care treatment corresponding to the underlying disease and patient needs. Each patient was invited to attend a control visit approximately one week later, when all measurements performed at baseline were repeated for the second time.

Currently, the amplitude (mean or maximal itch) and recall periods of VAS, NRS and verbal rating scale (VRS) assessment do not follow an international standard and are utilized differently in different countries and clinical trials (3). In order to compare the MCID of different amplitude and recall periods, we investigated in parallel 2 large cohorts of patients using 2 different assessments: a longer one (3 days) with higher amplitude (maximal intensity) used in Poland, and a shorter one (24 h) with lower amplitude (mean intensity) used in Germany. The endpoints and number of assessments were the same in the 2 groups.

Poland. All patients ($n = 206$) were asked to rate their maximal itch intensity within the last 3 days prior to the visit according to the VAS, NRS and VRS in a random order, both at baseline (V1) and at the control visit (V2). The VAS was used as a 10-cm long horizontal line with a starting point as "no itch" (0 points) and ending with "worst imaginable itch" (6). Similarly, patients were asked to assess their itch verbally using the NRS, from 0 (no itch) to 10 points (worst imaginable itch). In addition, itch was assessed with a VRS, using the following grading system: no itch (0), mild (1), moderate (2) and severe itch (3).

Germany. All patients ($n = 192$) were asked to assess their mean itch intensity within the previous 24 h according to VAS, NRS and VRS in a random order, both at baseline (V1) and at the control visit (V2).

To determine MCID for VAS and NRS in itch assessment we used an anchor-based method using external anchoring measu-

Table I. *Patients' characteristics*

	Polish population $n = 206$	German population $n = 192$
Itch category	Dermatological, $n = 206$ (100%)	Dermatological, $n = 73$ (38.0%) Systemic cause, $n = 18$ (9.4%) Neurological, $n = 20$ (10.4%) Psychiatric, $n = 1$ (0.5%) Multifactorial, $n = 43$ (22.4%) Unclear, $n = 37$ (19.3%)
Sex, *n* (%)		
Female	90 (43.7)	96 (50.0)
Male	116 (56.3)	96 (50.0)
Age, years, mean ± SD (range)	47.0 ± 16.6 (18–83)	57.8 ± 14.8 (19–84)
Baseline itch scoring, scores		
NRS, mean ± SD	6.2 ± 2.3	5.8 ± 2.6
VAS, mean ± SD	5.9 ± 2.6	5.3 ± 2.8

NRS: numerical rating scale; VAS: visual analogue scale; SD: standard deviation.

rement (VRS) as reference (5). In order to calculate MCID we analysed 2 measurement time-points (V1 and V2) and included patients from the cohort, who changed the category of anchor scale by one category (e.g. from moderate to mild itch, etc.). The difference in VAS and NRS scoring in these patients was calculated.

All data were described by distributional characteristics, such as the mean and standard deviation, median or frequencies, depending on the type of data. The non-parametric sign test (7) was used to test the equivalence of changes between the VAS and NRS in groups, as the data were not normally distributed. Statistical analysis was carried out using SAS 9.3 for Windows.

RESULTS

Regarding the Polish population, 2 patients did not complete the VAS assessment on the second visit and 2 other patients did not report itch intensity with NRS on the same visit. Among the remaining patients, 87 subjects (42.2%) improved the VRS category on the control visit by one degree compared with baseline. Considering the maximal VAS within the previous 3 days, the change in VAS_{max} in the group of patients improving by one category of anchor measurement was 2.8 ± 2.0 points. Similarly, to achieve the change in VRS by one category it was necessary to change NRS_{max} scoring by 2.7 ± 1.7 points. The differences between VAS_{max} and NRS_{max} assessments were not statistically significant (Table II).

In the German population all patients completed both NRS assessments, while 6 patients missed the VRS and 4 missed the VAS at baseline visit, and 4 patients missed the VRS and 3 the VAS on the second visit. In total, 38 patients (19.8%) improved the VRS by one category in 1 week. The $VAS_{average}$ change within 24 h was 3.0 ± 2.5 points when changing the category of anchor measurement by one. When considering the change in $NRS_{average}$ within 24 h, 39 patients changed VRS by one category. The change in $NRS_{average}$ was 2.7 ± 1.8 points. There were no statistically significant differences between changes in $VAS_{average}$ and $NRS_{average}$ (Table II).

Table II. *Minimal clinically important difference calculation. Change in visual analogue scale (VAS) and numerical rating scale (NRS) scoring between visits 1 and 2 for patients achieving improvement in the verbal rating scale (VRS) by one category and for patients maintaining the same VRS category at each visit*

	Polish population		German population	
	ΔVAS_{max} last 3 days	ΔNRS_{max} last 3 days	$\Delta VAS_{average}$ last 24 h	$\Delta NRS_{average}$ last 24 h
ΔVRS=1 (improvement by one category)				
n	87	87	38	39
Mean change	2.8 ± 2.0	2.7 ± 1.7	3.0 ± 2.5	2.7 ± 1.8
Median change	2.6	2.0	2.5	2.0
p	0.22		0.18	
ΔVRS=0 (no category change)				
n	67	67	66	66
Mean change	0.6 ± 1.9	0.67 ± 1.8	0.3 ± 1.9	0.4 ± 1.8
Median change	0.4	1.0	0	0
p	0.6		0.76	

DISCUSSION

The proper interpretation of any numerical instrument used for the assessment of changing clinical status should consider whether the observed change is of clinical relevance. In large populations, especially, small changes may be statistically significant, but can be irrelevant clinically. The MCID is defined as the smallest change in any scale scoring that can be noticed by the patient (8, 9). Several methods exist for determining MCID; however, currently no universal rule is established. We applied an anchoring technique, as it seemed to us to be the most reliable one; however, a Delphi process and distribution-based method can also be applied (5, 10).

Based on our study, our data suggest that the MCID for clinical improvement in itch, as rated on the VAS and NRS, ranks between a decrease of 2–3 points (for details, see Table II). MCID for improvement in a clinical condition in any study with active medical intervention must be higher than the effect of placebo. As shown recently by van Laarhoven et al. (11), the placebo effect may decrease itch by 1.3 points compared with baseline if the VAS is used for itch assessment (95% confidence interval 1.0–1.6). These results show that the placebo change on VAS/NRS and clinically meaningful change on VAS/NRS are distinct.

Interestingly, there was almost no numerical or statistically significant difference in MCID results between NRS and VAS, supporting the results of our previous studies on the comparability of these tools (1, 2). There is only a slight MCID difference of 0.1–0.6 points (in mean and median values) between VAS and NRS, with higher MCID in VAS, but this was not statistically significant.

These results must be viewed with caution, however, and should be considered preliminary, due to several limitations. The 2 populations included in our study were slightly different; the Polish group included dermatological itch only, while in the German population all types of chronic itch types were included. It is also not clear whether the variations between the Polish and German populations reflect cross-cultural differences, differences in disease severity due to inclusion of several different healthcare centres, and different itch assessments (maximal vs. mean itch). Different MCID values might also depend on different recall periods used in both populations (24 h vs. 3 days). Although we found almost no difference, it is likely that recall periods that are too long may result in a lower reliability of itch assessment. The definite determination of recall periods should be addressed in another prospective study.

REFERENCES

1. Phan NQ, Blome C, Fritz F, Gerss J, Reich A, Ebata T, et al. Assessment of pruritus intensity: prospective study on validity and reliability of the visual analogue scale, numerical rating scale and verbal rating scale in 471 pa-

tients with chronic pruritus. Acta Derm Venereol 2012; 92: 502–507.
2. Reich A, Heisig M, Phan NQ, Taneda K, Takamori K, Takeuchi S, et al. Visual analogue scale: evaluation of the instrument for the assessment of pruritus. Acta Derm Venereol 2012; 92: 497–501.
3. Ständer S, Augustin M, Reich A, Blome C, Ebata T, Phan NQ, et al. Pruritus assessment in clinical trials: consensus recommendations from the International Forum for the Study of Itch (IFSI) Special Interest Group Scoring Itch in Clinical Trials. Acta Derm Venereol 2013; 93: 509–514.
4. Kido-Nakahara M, Katoh N, Saeki H, Mizutani H, Hagihara A, Takeuchi S, et al. Comparative cut-off value setting of pruritus intensity in visual analogue scale and verbal rating scale. Acta Derm Venereol 2015; 95: 345–346.
5. Copay AG, Subach BR, Glassman SD, Polly DW Jr, Schuler TC. Understanding the minimum clinically important difference: a review of concepts and methods. Spine J 2007; 7: 541–546.
6. Furue M, Ebata T, Ikoma A, Takeuchi S, Kataoka Y, Takamori K, et al. Verbalizing extremes of the visual analogue scale for pruritus: a consensus statement. Acta Derm Venereol 2013; 93: 214–215.
7. Conover WJ. Chapter 3.4: The Sign Test. In: Practical nonparametric statistics. (3rd edn). New York: Wiley; 1999: p. 157–176.
8. Cook CE. The Minimal clinically important change score (MCID): a necessary pretense. J Man Manip Ther 2008; 16: E82–E83.
9. Hajiro T, Nishimura K. Minimal clinically significant difference in health status: the thorny path of health status measures? Eur Respir J 2002; 19: 390–391.
10. Ringash J, O'Sullivan B, Bezjak A, Redelmeier DA. Interpreting clinically significant changes in patient-reported outcomes. Cancer 2007; 110: 196–202.
11. van Laarhoven AI, van der Sman-Mauriks IM, Donders AR, Pronk MC, van de Kerkhof PC, Evers AW. Placebo effects on itch: a meta-analysis of clinical trials of patients with dermatological conditions. J Invest Dermatol 2015; 135: 1234–1243.