The aim of this study was to evaluate the visual analogue scale (VAS) as a method of pruritus assessment. A total of 310 subjects with pruritic dermatoses (148 Caucasian subjects and 162 Asian subjects) were recruited. The patients assessed pruritus intensity using the horizontal and vertical VAS, numeric rating scale (NRS) and verbal rating scale (VRS). All scales showed very good reproducibility (intraclass coefficient (ICC) > 0.8). No significant differences were found between the horizontal and vertical VAS (5.3 ± 2.9 vs. 5.3 ± 3.0 points, p = 0.34). Using NRS, patients rated their pruritus significantly higher than with VAS (5.7 ± 2.6 points, p < 0.01). VRS showed the highest correlation with NRS (R = 0.82, p < 0.001), followed by horizontal (R = 0.75, p < 0.001) and vertical VAS (R = 0.74, p < 0.001). Based on detailed analysis following VAS categories were proposed: 0 = no pruritus, > 0–< 4 points = mild pruritus, ≥ 4–< 7 points = moderate pruritus, ≥ 7–< 9 points = severe pruritus, and ≥ 9 points = very severe pruritus. In conclusion, the VAS is a valuable method of pruritus measurement. Key words: itch; pruritus; measurement; validation.

(Accepted August 31, 2011.)


Adam Reich, Department of Dermatology, Venereology and Allergology, Wroclaw Medical University, Ul. Chalubinskiego 1, PL-50-368 Wroclaw, Poland. E-mail: adi_medicalis@poczta.onet.pl

Itching is an unpleasant sensation that leads to the desire to scratch (1). As pruritus is a subjective feeling, the objective measurement of its intensity remains a challenge. There is a range of methods of pruritus evaluation, but none can be considered as standard. In general, assessments of pruritus may be divided into two major groups: subjective evaluations of itch, and measurement of scratching. The first group includes simple assessments of itch severity (i.e. visual analogue scale (VAS), numeric rating scale (NRS), verbal rating scale (VRS)), itch questionnaires providing data on itch quality, computerized analysing systems, and measurement of pruritus perception threshold. Scratching may be assessed with the help of observation of excoriations and degree of lichenification, infrared video-recording, limb meters (wrist activity monitors, pressure sensors), fingernail vibration transducers (piezo film sensors, pruritometer) and acoustic evaluation system of scratching. In addition, functional imaging techniques (functional magnetic resonance, positron emission tomography) have been used to analyse brain activity during itching episodes (2, 3). In clinical studies on pruritus it is usually recommended to use at least two independent methods of itch assessment. However, this recommendation could be too time-consuming for daily clinical use, and a simple and reliable method of itch intensity measurement is highly desirable.

VAS seems to be one of the most commonly used methods of pruritus severity assessment, as it provides an easy and rapid estimation of itch (4). It was developed originally to assess the intensity of pain, but subsequently it was also adopted for pruritus evaluation. A number of studies dealing with pain have demonstrated that VAS, despite some limitations (e.g. being awkward to use for patients with motor problems, being difficult for some patients to understand, and the time-consuming need to transform a graphic result to a metric one), is a reliable method of pain severity measurement (5–11). By analogy, it was recognized that VAS should also be suitable for pruritus assessment. However, until now no studies have been performed to validate this scale as an instrument for the measurement of itch intensity. Therefore, on behalf of the International Forum for the Study of Itch (IFSI), we performed a study evaluating VAS as a method of pruritus assessment, also defining a set of VAS bands referring to mild, moderate, severe and very severe pruritus.

MATERIALS AND METHODS

Patients

Pruritic patients were consecutively recruited from the cohorts of patients admitted to our departments for diagnostics and treatment of skin diseases between November and December 2008 (Poland)
and November and December 2009 (Japan). The inclusion criteria were as follows: informed consent obtained from a patient to participate in the study, age over 18 years, presence of dermatological itch according to the newest pruritus classification (1), and neither motor nor cognitive problems that might preclude patients from understanding the scale or marking the line with a pen. A total of 310 patients with various dermatological diseases were recruited. At the time of the study, 175 (57%) individuals took anti-pruritic drugs, mostly antihistamines (83.2%). There were 148 (47.7%; male/female ratio: 72/76) Caucasian subjects (from Poland) and 162 (52.3%; male/female ratio: 89/72) Asian subjects (from Japan) with pruritus recruited. Of note, patients from Poland were significantly older (Poland: 52.7 ± 16.4 years vs. Japan: 46.5 ± 19.7 years, p = 0.003) and more commonly had psoriasis (Poland: 33.8% vs. Japan: 9.3%, p < 0.001) and lichen planus (Poland: 5.4% vs. Japan: 0.6%, p = 0.03), while participants from Japan more frequently had atopic dermatitis (Poland: 12.8% vs. Japan: 35.2%, p < 0.001).

Study design

After collection of basic socio-demographic and clinical data, pruritic subjects were asked to rate their pruritus perceived within the previous 24 h using the horizontal VAS, followed by the vertical VAS, NRS and the 5-point VRS (no pruritus (0 points), mild pruritus (1 point), moderate pruritus (2 points), severe pruritus (3 points), very severe pruritus (4 points)). The VAS is a 10-cm long line (oriented horizontally or vertically), on which patients indicated the intensity of pruritus by crossing the line at the point that corresponded to their pruritus severity, being informed that the beginning of the scale refers to no pruritus (0 points) and the end to the most severe pruritus they can imagine (10 points). Using NRS, the patients assessed the intensity of pruritus verbally from 0 (no pruritus) to 10 (the most intensive pruritus they can imagine). As the results obtained from Asian patients were originally assumed to be only for defining the categories of VAS, this group completed the shorter form of the questionnaire including horizontal VAS and VRS for itch evaluation.

Test-retest comparison

Forty-nine randomly selected Caucasian patients were chosen for test-retest comparison. As pruritus intensity may vary significantly even within one day, we decided to perform the test-retest comparison during the same day with a 3-hour interval between the first and the second completion of the questionnaire in order to minimize circadian changes of the itching. Patients were asked to complete the questionnaire for the first time in the morning between 08:30 h and 10:30 h. The second questionnaire was distributed among same patients approximately 3 h later (i.e. between 11:30 h and 13:30 h). Both questionnaires were completed by patients in the same order. The test-retest reproducibility was based on the calculation of intraclass correlation coefficient (ICC) and Spearman’s rank correlation coefficient between first and second results.

Defining the categories of the visual analogue scale

To define the set of bands of VAS, we have chosen the results obtained using horizontal VAS. Determination of the VAS categories was performed similarly to Hongbo et al. (12). The mean, median and mode of the VRS scoring for each VAS score (rounded to the whole) were used to define cut-offs of the VAS scoring. Next, the κ coefficient of agreement was calculated for various sets of bands of the VAS scores as well as the correlation coefficient between the VRS scoring and individual sets of bands was determined. According to Landis & Koch (13) values of κ coefficient < 0 indicate no agreement, 0–0.2 slight, 0.21–0.4 fair, 0.41–0.6 moderate, 0.61–0.8 substantial, and 0.81–1 almost perfect agreement. Initially, the cut-offs were calculated in Caucasian subjects (14). To compare with other ethnic groups and to rule out that Asian subjects rate pruritus differently, we extended our study to the Japanese population as a representative of Asian subjects.

Statistical analysis

All data were analysed statistically using Microsoft Excel 2000 (Microsoft Corporation, Warsaw, Poland) and Statistica 7.0 Pl (Statsoft, Krakow, Poland). Mean values, standard deviations, minimal and maximal values, as well as frequencies, were calculated. Paired and unpaired Student’s t-test, χ2 test with Yates correction, analysis of variance (ANOVA) with Scheffe’s post hoc test and Spearman’s rank correlation test were used where appropriate. p-values < 0.05 were considered significant.

RESULTS

Pruritus severity

Based on the VRS, 36 (24.3%) Caucasian subjects had mild, 53 (35.8%) moderate, 38 (25.7%) severe and the remaining 21 (14.2%) very severe pruritus. Among 162 pruritic Asian subjects, 74 (45.7%) had mild, 53 (32.7%) moderate, 25 (15.4%) severe and 10 (6.2%) very severe itching. Caucasian subjects indicated significantly more commonly that they had severe or very severe itching, while Japanese individuals more frequently had mild pruritus (χ2 test: p < 0.001). Similarly, using horizontal VAS Caucasian subjects scored their pruritus significantly higher (mean: 5.3 ± 2.9 points) than Japanese subjects (mean: 4.1 ± 2.6 points; p < 0.001).

Neither gender, nor age significantly influenced the pruritus intensity scoring assessed with VAS (males vs. females: 4.6 ± 2.6 points vs. 4.8 ± 3.0 points, respectively, Student’s t-test: p = 0.58; age: Spearman’s rank correlation test: R = −0.04, p = 0.4) or VRS (males vs. females: 2.1 ± 1.0 points vs. 2.0 ± 0.9 points, respectively, Student’s t-test: p = 0.26; age: Spearman’s rank correlation test: R = 0.005, p = 0.93). No significant differences were observed between patients taking anti-pruritic medicines and those who did not (VAS: 4.8 ± 2.8 points vs. 4.4 ± 2.8 points, respectively, Student’s t-test: p = 0.2; VRS: 2.1 ± 1.0 points vs. 1.9 ± 1.0 points, respectively, Student’s t-test: p = 0.06). We did not find any significant differences between various diseases included in the study regarding the itch scoring (VAS: p = 0.22; VRS: p = 0.5; results based on ANOVA). Similar results were found using NRS and vertical VAS in the Caucasian population (data not shown).

Comparison of various scales of pruritus assessment

There was no statistical difference between the horizontal and vertical VAS scoring (mean: 5.3 ± 2.9 points (range 0.4–10.0 points) vs. 5.3 ± 3.0 points...
Second pruritus scoring was noted in all scales (Table I). Interestingly, a slight, although significant, decrease in the pruritus was evaluated by horizontally-oriented VAS (Table I).

All scales showed a very good reproducibility in the defined study settings (Table I). Among the tested scales, the highest ICC value (ICC = 0.88) was observed when pruritus was evaluated by horizontally-oriented VAS (Table I). Interestingly, a slight, although significant, decrease in the second pruritus scoring was noted in all scales (Table I).

Fig. 1. Severity of pruritus based on horizontal visual analogue scale (VAS), vertical VAS and numeric rating scale (NRS) (results presented as means and standard deviations) (n = 148). The significance level indicate difference between NRS and each of the two VAS values (ANOVA).

Test-retest comparison

Comparing VRS with the other scales, we found that all other scales significantly correlated with VRS, and the highest correlation was observed between VRS and NRS (Spearman’s rank correlation test: R = 0.82, p < 0.001), followed by horizontal VAS (R = 0.75, p < 0.001) and vertical VAS (R = 0.74, p < 0.001). Results obtained with vertical VAS were very similar to results of horizontal VAS (R = 0.95, p < 0.001). Results obtained with NRS also correlated significantly (Student’s t-test: p < 0.001 for horizontally-oriented VAS and p = 0.001 for vertically-oriented VAS) (Fig. 1). Similarly to the Caucasian population, the results of VAS assessment in Japanese patients showed a very good correlation with VRS (R = 0.82, p < 0.001).

Defining the bands of the visual analogue scale

Each category of VRS differs significantly from the others regarding the VAS scoring (Fig. 2). After calculation of means, medians and modes of VRS scoring for each VAS point, we concluded that the limit between mild and moderate pruritus should be between 3 or 4 points, the limit between moderate and severe pruritus should be 6 or 7 points, and the limit between severe and very severe pruritus should be approximately 9 points of VAS (Table II). These data were confirmed by the calculation of the κ coefficient of agreement (Table SI; available from http://www.medicaljournals.se/acta/content/?doi=10.2340/00015555-1265). However, we have observed some differences between Caucasian and Asian subjects regarding VAS categorization. For Caucasian subjects the most suitable set of bands were estimated as follows: mild pruritus ≥ 0 but < 3 points, moderate pruritus ≥ 3 but < 7 points, severe pruritus ≥ 7 but < 9 points, and very severe pruritus ≥ 9 points. In the Japanese population the limit between mild and moderate pruritus was higher (4 points), while the limit for moderate/severe pruritus was slightly lower (6 points). However, based on the analysis of all patients together, we have proposed a final set of bands for VAS (κ = 0.53) (Table III).

DISCUSSION

VAS is a simple and reproducible tool for the assessment of pain severity (5–11). Despite some limitations, this scale provides physicians with valid and reliable estimates of pain (15). This scale has also been used widely for the assessment of pruritus intensity; however, the instrument has never been validated in pruritic subjects before and, until now, the results obtained with VAS in these populations of patients had to be interpreted with some caution. Therefore, in the current study we performed psychometric assessments of VAS as a method of pruritus measurement.

Based on the results achieved, it could be concluded that VAS may indeed serve as a reliable method of pruritus assessment. However, considering significant differences between Japanese and Polish patients with pruritic skin diseases regarding VAS scoring, we cannot exclude that pruritus rating could be influenced by some cultural and ethnic aspects that have not yet been determined.

Table I. Test-retest reproducibility of the visual analogue scale (VAS), numeric rating scale (NRS) and verbal rating scale (VRS) in the assessment of pruritus (n = 49)

<table>
<thead>
<tr>
<th></th>
<th>Horizontal VAS</th>
<th>Vertical VAS</th>
<th>NRS</th>
<th>VRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>First assessment</td>
<td>4.8 ± 2.6</td>
<td>5.1 ± 2.8</td>
<td>5.3 ± 2.5</td>
<td>2.2 ± 0.9</td>
</tr>
<tr>
<td>Second assessment</td>
<td>4.4 ± 2.4</td>
<td>4.7 ± 2.6</td>
<td>4.8 ± 2.2</td>
<td>2.0 ± 0.8</td>
</tr>
<tr>
<td>Student’s t-test</td>
<td>t = 2.1 (p &lt; 0.05)</td>
<td>t = 1.9 (p = 0.06)</td>
<td>t = 2.5 (p = 0.02)</td>
<td>t = 2.0 (p = 0.05)</td>
</tr>
<tr>
<td>Intraclass correlation coefficient</td>
<td>0.88</td>
<td>0.87</td>
<td>0.83</td>
<td>0.85</td>
</tr>
<tr>
<td>Spearman’s rank correlation coefficient</td>
<td>0.88 (p &lt; 0.001)</td>
<td>0.87 (p &lt; 0.001)</td>
<td>0.83 (p &lt; 0.001)</td>
<td>0.8 (p &lt; 0.001)</td>
</tr>
</tbody>
</table>
Although this observation might be related to significant differences in the age of, and some variations regarding the dermatological conditions in, both populations, such explanation is rather unlikely, as neither age nor diagnosis have been found to significantly influence the VAS scoring. Furthermore, we could not exclude inter-rater bias, as both parts of the study were performed by various research groups. Thus, further research is needed to better characterize these differences in VAS scores, including other populations, especially Black Africans.

Importantly, the positioning of VAS (vertical vs. horizontal) does not have a significant impact on the pruritus scoring and it appears that both versions can be used interchangeably. Similar findings were observed previously with pain assessment (5, 10). However, we suggest that horizontally-oriented VAS may be the preferred method of pruritus assessment, as it is the most commonly used version of VAS. Interestingly, the rating of pruritus with NRS, a very similar instrument to VAS, does not provide the same results as VAS. Usually patients rated pruritus higher with NRS than with VAS, a phenomenon also found by other authors during pain assessment (11, 16–18). Therefore, NRS should not be considered as a verbal version of VAS and the results obtained with NRS cannot be directly compared with VAS scoring.

Both VAS variants showed a very good correlation with both NRS and VRS, indicating a good convergent validity of VAS in the pruritus assessment, a result also demonstrated in another study (19). Significant differences in VAS scoring between various itch categories of VRS confirmed a content validity of VAS. However, for unknown reasons seven (7.2%) patients, who stated that they did not have pruritus crossed the VAS line at the point other than 0, indicating that some patients may have difficulty understanding the VAS. On the other hand, none of the group confirming pruritus indicated 0 on the VAS scale.

VAS also showed a good test-retest reproducibility (ICC = 0.88); however, a significantly lower pruritus scoring during the second assessment of itch intensity needs some caution. We suppose that this phenomenon might reflect circadian variation in pruritus intensity. However, this observation should be interpreted with care, and requires further investigation.

One of the major achievements of this study is the categorization of VAS. To date, the VAS results were demonstrated as points, providing little information about the meaning of the scoring. Some authors previously proposed categorization of VAS, but without giving any psychometric background for such banding (20, 21). Therefore, it is difficult to state whether these categories are valid. In the current study we defined VAS scoring of mild, moderate, severe and very severe pruritus. We observed some differences between Caucasian and Asian subjects. Therefore, our categories need further confirmation, and it is also possible that, in certain populations, these categories could differ from our cut-offs. Moreover, a limitation of our study could be the fact that the categorization of VAS was based on VRS, because both scales are subjective. It would be highly recommended to confirm the defined set of bands in the future using an objective measurement of pruritus intensity, e.g. by counting the scratch episodes, although a study by Murray & Rees (22) has shown that there is a poor correlation between VAS and actigraphic measures of scratch. Furthermore, studies are needed to determine the responsiveness of VAS and other methods of pruritus assessment to change in the severity of itching. To date, no such studies have been undertaken. More work is also needed in defining the minimal clinically important

---

**Table II. Relationship between the horizontal visual analogue scale (VAS) and verbal rating scale (VRS)**

<table>
<thead>
<tr>
<th>VAS scoring</th>
<th>VRS scoring</th>
<th>Mean</th>
<th>Mode</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 (0–0.5 cm)</td>
<td>0.1 ± 0.3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1 (0.6–1.5 cm)</td>
<td>1.3 ± 0.7</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2 (1.6–2.5 cm)</td>
<td>1.2 ± 0.4</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3 (2.6–3.5 cm)</td>
<td>1.6 ± 0.6</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>4 (3.6–4.5 cm)</td>
<td>1.8 ± 0.6</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>5 (4.6–5.5 cm)</td>
<td>2.0 ± 0.6</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>6 (5.6–6.5 cm)</td>
<td>2.4 ± 0.5</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>7 (6.6–7.5 cm)</td>
<td>2.7 ± 0.6</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>8 (7.6–8.5 cm)</td>
<td>3.2 ± 0.7</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>9 (8.6–9.5 cm)</td>
<td>3.2 ± 0.9</td>
<td>3</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>10 (9.6–10 cm)</td>
<td>3.6 ± 0.6</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

**Table III. Set of bands for visual analogue scale (VAS) (κ = 0.53) in the assessment of pruritus**

<table>
<thead>
<tr>
<th>VAS scoring</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 points</td>
<td>No pruritus</td>
</tr>
<tr>
<td>&gt; 0 points but &lt; 4 points</td>
<td>Mild pruritus</td>
</tr>
<tr>
<td>≥ 4 points but &lt; 7 points</td>
<td>Moderate pruritus</td>
</tr>
<tr>
<td>≥ 7 points but &lt; 9 points</td>
<td>Severe pruritus</td>
</tr>
<tr>
<td>≥ 9 points</td>
<td>Very severe pruritus</td>
</tr>
</tbody>
</table>
difference for VAS as a tool for pruritus measurement. Despite these limitations, we believe that our results are of importance and that we have provided sufficient data to support the defined VAS categories.

Although VAS appears to be a valuable method of pruritus measurement, some limitations of this instrument must be mentioned. The most important issue is the fact that VAS is not suitable for people with motor or cognitive problems that preclude understanding the scale or marking the line with a pen. This limitation may be particularly important for elderly people and young children (6, 7). Furthermore, VAS only provides information about itch intensity, thus multidimensional assessment tools could be more suitable, if more detailed pruritus evaluation is needed, as they can provide comprehensive information about various aspects of itching. However, these instruments are usually considered to be too lengthy and not suited for repeated assessments, especially in clinical settings. Moreover, they usually require some psychometric expertise and time for proper interpretation (15, 23).

In conclusion, VAS seems to be a valuable method of pruritus assessment. Using VAS various levels of pruritus can be defined, similarly to VRS, but, in addition, better discrimination of pruritus severity can be carried out, a property that makes this scale highly suitable for clinical studies evaluating various antipruritic regimens.

ACKNOWLEDGEMENTS

This study was supported by a research fellowship within the “Development program of Wroclaw Medical University”, funded by the European Social Fund, Human Capital, National Cohesion Strategy” (contract no. UDA-POKL.04.01.01.-00-010/08-01).

The authors declare no conflicts of interest.

REFERENCES