A test-retest reliability study of human experimental models of histaminergic and non-histaminergic itch

Andersen, Hjalte Holm; Rosenkrans Sørensen, Anne-Kathrine; Nielsen, Gebbie A. R.; Sparso Mølgaard, Marianne; Stilling, Pernille; Boudreau, Shellie; Elberling, Jesper; Arendt-Nielsen, Lars

Published in:
Acta Dermato-Venereologica

DOI (link to publication from Publisher):
10.2340/00015555-2502

Publication date:
2017

Document Version
Publisher's PDF, also known as Version of record

Link to publication from Aalborg University

Citation for published version (APA):
INVESTIGATIVE REPORT

A Test–Retest Reliability Study of Human Experimental Models of Histaminergic and Non-histaminergic Itch

Hjalte H. ANDERSEN¹, Anne-Kathrine R. SØRENSEN¹, Gabriele A. R. NIELSEN¹, Marianne S. MØLGAARD¹, Pernille STILLING¹, Shellel BOUDREAU¹, Jesper ELBERLING² and Lars ARENDT-NIELSEN¹

¹SMi®, Department of Health science and Technology, Faculty of Medicine, Aalborg University, Aalborg, and ²The Allergy Clinic, Copenhagen University Hospital, Gentofte, Copenhagen, Denmark

Numerous exploratory, proof-of-concept and interventional studies have used histaminergic and non-histaminergic human models of itch. However, no reliability studies for such surrogate models have been conducted. This study investigated the test–retest reliability for the response to histamine- and cowhage- (5, 15, 25 spiculae) induced itch in healthy volunteers. Cowhage spiculae were individually applied with tweezers and 1% histamine was applied with a skin prick test (SPT) lancet, both on the volar forearm. The intensity of itch was recorded on a visual analogue scale and self-reported area of itch was assessed 5 and 10 min after itch provocation. Reliability of the evoked itch (area under the curve and peak intensity) was assessed by the coefficient of variation (CV), intra-class correlation coefficient (ICC), and sample size estimation for parallel and cross-over designs. Cowhage (ICC = 0.57–0.77, CV between = 97%, CV within = 41%) and histamine: (ICC = 0.83–0.93, CV between = 97%, CV within = 20%) exhibited moderate-to-excellent intra-individual reliability and moderate inter-individual reliability for the itch intensity. For a test–retest observation period of one week, SPT-delivered histamine and application of cowhage-spiculae are reproducible human models of itch. The high inter-individual and low intra-individual variability suggests cross-over designed studies when applicable.

Key words: itch; cowhage; histamine; reliability; itch models; test–retest.

Accepted Jul 4, 2016; Epub ahead of print Jul 5, 2016


Corr: Lars Arendt-Nielsen, Department of Health Science and Technology, Faculty of Medicine, Aalborg University, Fredrik Bajers Vej 703, DK-9220 Aalborg E, Denmark. E-mail: LAN@hst.aau.dk

Itch, also known as pruritus, is defined as “an unpleasant sensation evoking a desire to scratch” (1, 2). Chronic itch (>6 weeks) (3) is highly bothersome and is usually a manifestation of a disease, but can also occur without any obvious somatic causes. Chronic itch is often associated with decreased quality of life, insomnia, and attentional difficulties, and, in this respect, is similar to the well-known consequences of chronic pain (2, 4, 5). Neurophysiologically, itch is conveyed by at least 2 separate pathways; histaminergic itch is transmitted by nociceptive mechano-insensitive C-fibres (CMi) responsive to, for example, histamine, while non-histaminergic itch is transmitted by polymodal nociceptive mechano-sensitive C-fibres (PmC) responsive to, for example, mucunain from cowhage (6–8).

Histamine, which is by far the most studied pruritogen, excites the subset of pruritceptive CMi-fibres expressing H1R and H4R. CMi-fibres are peptidergic, terminate within and below the stratum granulosum of the epidermis, and have large receptive fields. As a result the activation of CMi-fibres is usually associated with the neurogenic axon-reflex (flare) (1, 2, 4, 9). Histaminergic itch, e.g. in chronic urticaria, is generally responsive to antihistamines. However, a number of disorders causing chronic itch cannot be treated with antihistamines, underscoring the clinical importance of also understanding the non-histaminergic pathway of itch (7, 10–13). A distinct non-histaminergic pathway of itch is mediated by the protease-activated receptors (PAR2 and PAR4) and transmitted through PmC- and Aδ-fibres with smaller receptive fields than CMi-fibres (14). PAR2 is the prototypical human receptor associated with non-histaminergic itch (10, 15, 16). Upon binding of a ligand, either exogenously or endogenously, a cleavage of PAR2 occurs, as well as a co-activation of transient receptor potential anillin 1 (2, 10, 17, 18). The most frequently studied exogenous PAR2-agonist is mucunain, a cysteine protease found on the spiculae of the plant Mucuna pruriens, commonly known as cowhage. Cowhage spicula activate the subset of pruritceptive fibres leading to the perception of both itch and mild pricking/stinging pain (3, 8, 13). The co-activation of Aδ-fibres should conceivable result in better spatial resolution and localizability of the non-histaminergic itch compared with histaminergic itch (7, 19–23), but this has not been confirmed. This could theoretically parallel the findings in the area of pain, where Aδ-fibre-mediated pricking pain is well localized, whereas C-fibre-mediated burning pain is more diffuse (19–23). Pruritceptive primary afferents of both the histaminergic and non-histaminergic itch pathway are multimodal, in the sense that they are also capable of transmitting nociceptive signals in response to algogenic stimuli, such as capsaicin (24).

The methods for induction of histamine- and cowhage-induced itch have been based on numerous different application techniques (8, 15, 25), but very little is known about their test–retest reliability. Reliable, symptom-
and mechanism-specific human and animal surrogate models are increasingly applied in biomedical research for several different purposes: (i) they provide a convenient, mechanism-based approach to profiling novel drug candidates in early clinical development; (ii) they can be used as provocation modalities to probe relevant patient populations for explorative, segmentational or diagnostic purposes or to test potential peripheral and central sensitization phenomena; and (iii) they can be readily applied for testing neurophysiological aspects of particular sensory symptomatology in animals and humans (3–5, 13, 26–30). Human surrogate itch models can also be applied to effectively guide the choice of the clinical study population in clinical trials, increasing the likelihood of a successful drug development (3, 5, 31). The present study aimed to assess the test–retest reliability of histamine- and cowhage-induced itch, together with an assessment of potential differences in the spatial discrimination of itch evoked by cowhage and histamine.

MATERIALS AND METHODS

Two 1.5 h study sessions were performed, approximately one week apart, on the same group of male volunteers. The order of itch provocations with cowhage and histamine, as well as arm dominance was randomized. This order of itch provocation was established in the first experimental session and repeated in the second session. While the subjects were unaware of which itch provocations were being applied and were instructed to look away during the administration, unblinding could have occurred based on the different tactile sensations evoked during application. The regional ethics committee approved the experimental protocol (approval number N-20140078) and the experiments were conducted at the Laboratory of Experimental Cutaneous Pain, Center for Sensory-Motor Interaction, Aalborg University in accordance with the principles of the Declaration of Helsinki (32). All participants provided written and oral informed consent prior to study participation and were informed that they could withdraw from the study, at any time.

Subjects

Fifteen male caucasian subjects (age 24.07 ± 0.73 years) participated in the study. Inclusion criteria were healthy male subjects, in the age range 20–35 years. Exclusion criteria were: any history of neurological, musculoskeletal, mental, dermatological, or allergic disorders, as well as any other on-going itch or pain problems.

Elicitation of itch with cowhage and histamine

Itch was induced in predefined application areas of the volar forearms of each subject. Three different arrangements of cowhage, using 5, 15, and 25 spiculae, as well as an application of histamine (1%) (Default Preparation, Diagenics, Bletchley, UK), were tested and compared (Fig. 1A). Application of the cowhage spiculae and histamine was randomized between the right and left arm. The 4 areas used for cowhage and histamine application were next to the middle of the volar forearms measured from the wrist to the cubital fossa (Fig. 1B). Using this demarcation, one application area was marked 2.5 cm above the middle of the forearm and one area was marked 2.5 cm below. The 4 areas measured 1.5×1.5 cm and were marked with a metallic paint marker (054, Edding International GmbH, Ahrensburg, Germany). Cowhage spiculae were prepared immediately prior to application under a stereo-microscope (Seben Incognita, Seben GmbH, Berlin, Germany) using a negative-grip tweezer (Electron Microscopy Science, Dumont, Switzerland) to arrange and transfer spiculae to another set of up to 5 insertion tweezers (Wittex GmbH, Munich, Germany). Each of the prepared tweezers was loaded with 5 spiculae, approximately 1 mm apart, which permitted the insertion of 4 (±1) spiculae in the skin per tweezer. The spiculae were inserted in the skin (Fig. 1B) within a 30 s interval and at an angle of approximately 45°, as described previously (33). The subjects did not observe the application of cowhage spicules or the skin prick test (SPT) lancet introduction of histamine. After 15 min the spiculae were removed with tape. A microdermascope (DinoLite, AnMo Electronics Corporation, Naarden, The Netherlands) was used to capture magnified images of the cowhage spiculae applied to the skin, and to validate that an acceptable amount of histamine had been inserted successfully (±20% of intended spiculae). The images were analysed with the proprietary manufacturer software (DinoCapture 2.0, AnMo Electronics Corporation).

Histamine was applied with an SPT lancet (Allergopharma, Diagnostics, Bletchley, UK) and a pressure of 120 g, using a SEN-SEBox electronic von Frey transducer (SENSEBox EvF, Somedic AB, Höby, Sweden), with an attached custom-made prick lancet mount (Aalborg University, Aalborg, Denmark). The controlled pressure applied was aimed at increasing reliability and based on an initial pilot study showing that 120 g was sufficient to establish a wheal > 3 mm in diameter in all participants, while minimizing the occurrence of bleeding (34, 35). Histamine application using weight-controlled SPT was chosen as a comparator, instead of, for example, iontophoresis, intradermal injections or inactivated histamine-coated spicule for several reasons: (i) SPT application of histamine has been reported to produce a stronger and more lasting itch sensation than iontophoresis and is associated with less pain and discomfort than intradermal injections (36, 37); (ii) it is the most commonly studied approach, while being easy, safe and clinically applicable (e.g. compared with coated spiculae and intradermal injections) (38); (iii) it requires a minimum of equipment and thus the method can easily be applied in any setting in future studies; (iv) compared with intradermal injection and iontophoresis, SPT application resembles that of spicules, in that the applied substance is mainly brought into contact with the epidermis and the dermoepidermal junction in a very localized manner.
Outcome measures

Assessment of experimentally-induced itch and pain intensity. The intensity of the elicited itch was assessed every 20 s for 10 min, from the moment of application, using a computerized visual analogue scale (VAS). The VAS ranged from 0 = "no itch" to 10 = "worst imaginable itch" with a label at 50% of its length, labelled "strong urge to scratch". Similar VAS modifications were used in the studies by Mørch et al. (39) and Hartmann et al. (13). The subjects were instructed to rate the sensation of itch by using the 50% mark as a guideline. Simultaneously to itch rating, the pain intensity of 2 sensory descriptors "pricking/stinging" and "burning" were reported for 10 min using 2 similar VASs with a sampling rate of 1/min. The VAS intensity for "pricking/stinging" allowed ratings from 0 = "no pricking/stinging" to 10 = "worst imaginable pricking/stinging pain". Likewise, the VAS for the burning pain allowed ratings from 0 = "no burning" to 10 = "worst imaginable burning pain". Both scales for nociceptive sensory qualities had identical 50% marks labelled "strongly painful sensation" analogue to the 50% label on the scale applied for itch sensation.

Localization and spatial distribution of induced itch. Navigate Pain software (Aalborg University, Aalborg, Denmark) was used to record the location and area of itch experienced by each subject 5 and 10 min after each itch provocation. This method has previously been validated and comparatively analysed matched to traditional hard-copy pain drawings in pain patients and healthy controls (40). The subjects were specifically instructed to: (i) draw the area where they felt itch on a 3D body schema representing the left or right volar forearm similar to the arm schema (Fig. 1A), (ii) disregard the area where the spiculae or the histamine had been applied which was visually observable during the VAS recordings, and (iii) focus only on the perceived area of itch. The percentage of the drawn area vs. the total arm area was multiplied by 2 to account for the fact that the applied body schema showed both volar forearms. Finally, the spatial distribution was obtained by superimposing the perceived areas of itch for each of the 4 itch provocations, at 5 and 10 min.

Statistical analysis

The software SPSS Statistics version 22 (IBM corporation, Redmond, USA) was used to analyse all data and to assess statistical significance. A significance level at \( p \leq 0.05 \) was chosen, marked with (*) in the results section, while highly significant differences with \( p \leq 0.01 \) are marked with (**) for groups. Descriptive parameters; arithmetic mean and standard error of the mean (SEM) were calculated for all reported parameters and are shown in the results section unless specifically stated otherwise. Visual inspection and confirmatory Shapiro-Wilk’s tests were performed to ascertain the normality of the data. For all parameters recorded with the VAS, the area under the curve (AUC) was calculated and the peak itch intensity was noted for every subject in response to each of the 4 itch provocations (5, 15, 25 spiculae, and 1% histamine) and at each session. The itch AUC, the peak itch intensity and the perceived area of itch from the 2 sessions were tested by 2 repeated measures analysis of variance (ANOVAs) (exposure; 4 levels, outcome; 2 levels, session; 2 levels) followed by Fischer’s test of least significant difference for pairwise post hoc comparisons. The main effect of sessions was used to ascertain whether a bias was present between data-sets obtained from session 1 vs. session 2. Correlation analyses were conducted using the Pearson product-moment correlation coefficient between the reported itch areas at 5 or 10 min and the peak itch intensity/AUC and itch intensity at 5 min. Correlations were assessed within each itch provocation condition (i.e. \( n = 15 \)) and across all 4 conditions for each session (i.e. \( n = 60 \)).

Reliability statistics

To estimate the reliability between the 2 sessions the intra-class correlation coefficient (ICC) and the coefficient of variation (CV) were calculated. In addition, the estimated minimum sample sizes for parallel and cross-over studies were calculated and Bland-Altman plots were produced. To evaluate the relative reliability between the 2 sessions the ICC (model 2.1; randomly selected raters and single measurements) for absolute consistency was used. Further information is available on different ICC models elsewhere (41–43). The ICC was calculated for each of the 4 applied itch provocations for the parameters itch AUC and peak itch intensity. ICC values ranging from 0.5 to 0.75 were considered moderate, while values > 0.75 are considered excellent (44). In the calculations of CV and ICC values, subjects who rated zero in both sessions were excluded from reliability calculations pertaining to that particular condition (i.e. 3 subjects for the 5 spiculae condition and 1 subject for 15 spiculae condition). This was done to avoid artificially inflating the produced measures of intra-individual reliability in the lower stimulus intensity conditions, where zero was occasionally rated in both sessions. One significant outlier (subject number 6, 15 spiculae condition) was excluded from further reliability assessments. The data from subjects are included in all other analyses. CV was calculated both within- and between-sessions, signifying inter- and intra-individual measurements, respectively. The following equations were used (their respective theoretical frameworks are elaborated upon in (39, 45))

\[
\text{Within session CV} = \frac{SD}{\mu} \cdot 100
\]

\[
\text{Between session CV} = \frac{SD_{\text{difference}}}{\mu} \cdot 100
\]

where SD represents the standard deviation and \( \mu \) the mean for each session. For potential future use in clinical or exploratory studies, minimally required sample size estimations were performed using the data obtained from each of the 4 itch provocations (5, 15, 25 spiculae, and 1% histamine). Using the following equations, the minimal amounts of participants in future parallel and cross-over studies were calculated. The power (1–b) was set to 0.8 and the significance level (a) at 0.05, using the standard textbook value (\( 2\times7.8=15.6 \)) at the relevant \( \alpha \) and \( \beta \)-levels (46, 47). The relevant detectable effect (E) was estimated as a 50% reduction of the mean itch intensity (itch AUC or peak itch intensity) averaged between sessions 1 and 2.

\[
\eta_{\text{parallel}} = \frac{15.6 \times SD^2}{E^2}
\]

\[
N_{\text{minimal}} = \frac{15.6 \times SD^2 \times (1-ICC)}{E^2}
\]

Bland-Altman-plots were produced to illustrate possible variations between the sessions, variations between the 4 itch provocations, outliers, and potential systematic biases. The difference, as calculated by subtraction of session 1 from session 2 for each parameter, was compared with the mean.

RESULTS

All of the 15 male subjects enrolled in this study completed both of the experimental sessions. The 2 sessions were conducted with \( 7.73 \pm 0.83 \) days in between. No significant main effect of session was found pertaining to induced itch: \( F(1, 14=0.073, p=0.791) \) or itch areas: \( F(1, 14=1.54, p=0.235) \), indicating no significant bias between experimental sessions.
Itch intensity and pain sensations

No statistical differences in itch AUC or peak itch intensity were detected between the 2 sessions for neither histamine nor cowhage. Distinct 10 min temporal itch intensity profiles were recorded in response to the 4 itch provocations in each of the 2 sessions (Fig. 2). Peak itch intensity was achieved 2–4 min after both cowhage and histamine applications with the histamine-induced itch exhibiting a slower decline than cowhage. The 3 applications of cowhage (5, 15, and 25 spiculae) exhibited dose-responsiveness, both in terms of itch intensity AUC and itch peak intensity, with the 5 spiculae and the 25 spiculae applications being significantly different (p<0.01 and p<0.05, respectively; Fig. 2E). While a significant difference was found between the 5 spiculae condition and 1% histamine for the peak itch intensity (p<0.05), the 1% histamine and the 25 spiculae, produced similar individual itch peak intensities at 4.09 ± 0.76 and 4.07 ± 0.77, respectively (p>0.9) (Fig. 2F). The histamine resulted in significantly higher itch AUC values than any of the cowhage conditions (p<0.05).

The recorded pricking/stinging peak VAS intensities were in mean 1.07 ± 0.29 for all 3 cowhage applications and 1.44 ± 0.37 for histamine averaged between sessions. The temporal VAS profiles are shown in Fig. 2. Not all subjects experienced the itch provocations as pricking/stinging and the frequency of reports differed between conditions, e.g. in session 1 after application of 25 spiculae, 14/15 subjects reported some presence of pricking/stinging, while after application of 5 spicules only 6/15 subjects reported these sensations. For burning, the intensities reported were very low: 0.38 ± 0.26 for cowhage and 0.42 ± 0.24 for histamine (data not plotted), and only 6 of 15 subjects employed the “burning” quality during the course of the study. Lastly, several participants spontaneously reported that cowhage-elicited itch and pain sensations occurred in intense “on/off” bouts; however, this was unquantifiable in the present study design.

Spatial discrimination of cowhage- and histamine-induced itch

The size of the reported perceived areas of itch exhibited high intra- and inter-individual variation. No significant differences were found for the perceived areas of itch between the sessions for any of the itch provocations with or without pooling data from 5 and 10 min drawings and sessions. The area of perceived itch to cowhage was significantly larger at 5 min compared with 10 min after application (for pooled sessions and conditions) (p<0.05), while no differences were observed for the size of the histamine-induced areas of itch at 5 min compared with 10 min after application for unpooled and pooled sessions (Fig. 3A). Moreover, no significant differences were found between the spatial distribution of cowhage- and histamine-evoked itch for any of the 4 applications including at similar peak intensities, i.e. between the 25 spiculae and the histamine applications (Fig. 4B).

Overlays of the spatial distributions of perceived itch are presented for 5, 15, 25 spiculae, and 1% histamine, with the 2 sessions being pooled (Fig. 4). As is apparent from Fig. 4A, the perception of low-intensity itch (5 spiculae) was difficult for the subjects to localize. No significant correlations were found between the perceived areas of itch and the rated itch intensity (AUC, peak or intensity at 5 min) within the individual provocations. Analyses across all 4 conditions in each of the 2 sessions yielded significant positive correlations between area size and itch intensity at 5 min in session 1: ρ=0.27, p<0.05, n=60, and in session 2: ρ=0.42, p<0.01, n=60.

Reliability of human models of itch

Table I shows the values of within-sessions CV, between-sessions CV, ICC values between sessions and estimated minimum sample sizes for the induced itch, measured as itch AUC and peak itch intensity. The inter-individual variation is generally high for induced itch, whereas the

![Fig. 2. Comparative figures of the temporal visual analogue scale (VAS) intensity for itch and “pricking/stinging” pain for each of the 4 itch provocations in session 1 and 2 (A: 1% histamine, B: 25 spiculae, C: 15 spiculae and D: 5 spiculae). VAS recordings of “burning” pain intensity are not included. (E) Mean of session 1 and 2 for AUC of itch intensity for each itch provocation. (F) Average of session 1 and 2 for VAS-peak of itch intensity for each itch provocation. Mean± standard error of the mean (SEM), *p<0.05, **p<0.01.](image-url)
intra-individual variation is relatively low. All tested itch provocations exhibited moderate-to-excellent reliability estimated by ICC calculation. For the 5, 15 and 25 spiculae applications ICC-values ranged from 0.63 to 0.77 for itch AUC, and 0.57 to 0.72 for peak itch intensity, without coherence in terms of the influence of the number of spiculae inserted. However, the itch induced by 1% histamine yielded better ICC values than for the cowhage applications, i.e. 0.83 for itch AUC and 0.93 for peak itch intensity. Parameter comparison between itch AUC and peak itch intensity from 10 min recordings showed a better performance for achieving low and intra-individual variability by measuring individual peak itch intensity over itch AUC.

The within-sessions CV values showed stimulus-intensity-dependent improvements in absolute reliability, i.e. lower CV values for itch provocations associated with more intense itch (such as 25 spiculae or 1% histamine). A similar trend was observed for between-sessions CV values, indicating that increasing the stimulation intensity can reduce both intra- and inter-individual variability (Table I). For absolute reliability measured by CV within both sessions, the peak itch intensity parameter was superior to the itch AUC, with a CV across all itch provocations, which was, in mean, 22.24 ± 3.77% lower. This resulted in proportionally lower minimum required sample sizes for cross-over studies compared with parallel designs.

The Bland-Altman difference plots (Fig. 5) illustrate the level of agreement between the 2 sessions for itch AUC and peak itch intensity. The difference plots showed no evidence of a fixed or proportional bias. The Bland-Altman plots for peak itch intensity showed a smaller mean difference than for itch AUC for all itch provocations, indicating a better between-session consistency for peak itch intensity, especially between 25 spiculae and 1% histamine (Fig. 5).

DISCUSSION

The present study showed that insertion of 5, 15, or 25 cowhage spiculae and intra-epidermal delivered 1% histamine resulted in moderate-to-excellent relative reliability and moderate absolute reliability of the evoked itch responses. The overall reliability was best for the provocations causing higher itch intensities. A total of 25 spiculae evoked itch with a peak itch intensity comparable to that of 1% histamine. No differences were detected in the sizes of the itch areas reported for any of the itch provocations. The study provides the basis for calculating sample sizes for parallel or cross-over studies employing human models of itch. Based on the reliability assessed here, cross-over designs are recommended whenever possible.

Fig. 4. Overlays for the overall spatial distribution and localization of itch for each of the applied itch provocations. A: 5 spiculae, B: 15 spiculae, C: 25 spiculae, and D: 1% histamine. Sessions 1 and 2 are pooled.
Sensory responses to histamine and cowhage

The temporal itch intensity profiles, evoked in response to the 4 itch models, are well aligned with previous research in terms of peak itch intensity and time course (8, 11, 13, 31). The peak itch intensity was generally reached within 2–4 min after administration and cowhage-induced itch exhibited a more rapid decrease than that of histamine, also in accordance with some of the previous findings (8, 48, 49). However, it should be noted that some previous studies have utilized a method of rubbing 40–45 spicules onto a larger skin area (11, 48, 50), and that, in some of these studies, this method is reported to produce a more intense itch (peak VAS ≈ 6–7) with less latency, which is probably associated with the application technique (11, 50). These discrepancies in the literature regarding the magnitude and quality of cowhage-evoked psychophysical responses highlight the problem of comparability between studies utilizing the cowhage model and standardizability of the model itself. Importantly, the model is reliant on diverse, untested plant material sourced from very different conditions across the world (9).

The 3 cowhage itch provocations caused stimulus-intensity dependent responses measured by itch AUC and peak itch intensity, whilst 25 spiculae and 1% histamine elicited very similar peak itch intensity. The SPT-delivered 1% histamine gave rise to a much longer-lasting itch than any of the cowhage applications, also in accordance with previous observations (11, 31).

These characteristic time/intensity profiles of itch induced by histamine and cowhage closely mimic the temporal profile of discharge rates recorded from histamine (53) or cowhage-responsive spinothalamic tract (STT) units in monkeys when their receptive fields were probed with histamine or cowhage, respectively (54).

Previous studies on cowhage- and histamine-induced itch have reported that the sensation of itch frequently occurs together with pricking/stinging and burning pain.

### Table I. Overview of reliability metrics for cowhage- and histamine-induced itch

<table>
<thead>
<tr>
<th></th>
<th>Absolute reliability</th>
<th></th>
<th>Relative reliability</th>
<th>Minimum estimated sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Within session 1 coefficient of variation %</td>
<td>Within session 2 coefficient of variation %</td>
<td>Coefficient of variation between sessions %</td>
<td>Intra-class correlation coefficient (2,1)</td>
</tr>
<tr>
<td>AUC of visual analogue scale itch intensity</td>
<td>5 spiculae 114.99</td>
<td>141.65</td>
<td>53.61</td>
<td>0.63</td>
</tr>
<tr>
<td></td>
<td>15 spiculae 129.32</td>
<td>96.17</td>
<td>42.03</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td>25 spiculae 84.15</td>
<td>89.68</td>
<td>42.58</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>1 histamine 86.94</td>
<td>95.55</td>
<td>25.43</td>
<td>0.83</td>
</tr>
<tr>
<td>Visual analogue scale itch peak intensity</td>
<td>5 spiculae 107.16</td>
<td>120.48</td>
<td>40.25</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>15 spiculae 89.41</td>
<td>81.64</td>
<td>36.18</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td>25 spiculae 48.26</td>
<td>67.70</td>
<td>33.79</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>1% histamine 68.56</td>
<td>77.31</td>
<td>15</td>
<td>0.93</td>
</tr>
</tbody>
</table>

*It may not be advisable to use there herein applied methods for elicitation of itch with n < 10 subjects regardless of the study design.

Absolute reliability, relative reliability and estimated sample size for itch area under the curve (AUC) and peak itch intensity. Notice that the minimum estimated sample size is a theoretical value based on calculations explained in methods section.

Fig. 5. Bland-Altman difference plots. (A) Itch visual analogue scale area under the curve (VAS AUC) and (B) Itch VAS peak of itch intensity for each of the applied itch provocations (5, 15, 25 spiculae and 1% histamine), containing limits of agreement (mean difference ± 1.96 SD difference). Solid line represents mean difference between sessions and the dashed lines the limits of agreement.
In the present study, the subjects did not observe the application of cowhage spicules or the SPT lancet introduction of histamine, both procedures are unavoidably associated with brief tactile/painful stimulation of the insertion area. Repeated vs. single insertion of spiculae, i.e. in the 5 vs. 25 spiculae conditions, may have produced the tendency towards more accurate localization of the areas of itch induced by 25 spiculae on the basis of more tactile noise rather than itch sensation even though the area drawings were recorded 5 min after itch provocations.

Sensory spatial characteristics and acuity of experimentally evoked itch and chronic itch occurring in patients (59) is a sparsely studied subject and it methodologically difficult to assess without introducing biasing tactile stimuli. A study by Wahlgren & Ekblom (60) explored the 2-point discrimination of itch stimuli using histaminergic itch provocations and found exceedingly poor spatial acuity compared with what is established for tactile and, more recently, also for pain stimuli (60, 61).

Reliability of surrogate models of itch
Test–retest reliability studies have previously been performed for different human surrogate models of pain and hyperalgesia, e.g. the ultraviolet B (UVB)-model of cutaneous hyperalgesia, the capsaicin model of secondary hyperalgesia or the L-menthol model of cold allodynia (29, 39, 62–64). Previously, the methods for induction of histaminergic and non-histaminergic itch have been based on a variety of different approaches to application (8, 15, 25), and have not been assessed for test–retest reliability. The lack of knowledge on test–retest performance decreases: (i) accurate power/sample-size calculation, (ii) usefulness of surrogate models of itch for antipruritic drug candidate screening, and (iii) comparability between studies applying models of itch.

In line with the present study, human surrogate models of itch, in general, exhibit high inter-individual variability (9, 13, 65). Moreover, in accordance with previous studies (33, 51), it was found that larger amounts of cowhage spiculae inserted into the skin resulted in more intense itch. Not previously explored, the present study revealed that itch induced with either cowhage or histamine was characterized by high test–retest reliability within individual subjects and that itch provocations evoking higher itch peak intensities were characterized by better test–retest reliability. It is plausible that application of even more spiculae, e.g. 40–45 or even 70–100 as...
used in some previous studies, would further improve reliability outcomes (5, 11, 48, 50). Along these lines, administration of a range of concentrations of histamine, e.g. by the use of intra-dermal injections, would have clarified whether histamine-induced itch also exhibits better test–retest reliability when more vigorous responses are evoked.

The high variability in itch intensity ratings between subjects is conceivably attributable to the following factors: (i) central or peripheral individual differences in the neurophysiological processing of itch, i.e. a result of different physiological susceptibility to itch induced by the utilized substances (33, 51); and (ii) methodological variations in relation to the amount of inserted spiculae, variations in skin thickness (resulting in different penetration depth of the spiculae) or personality traits (9, 51). As with all psychophysical studies relying on self-reporting of sensory experiences, variation could, in part, be a result of diverse interpretations of theVAS, thereby ensuing in dissimilar ratings of itch between subjects resulting in large inter-subject, but not intra-subject, variation. Although the amount of spiculae applied to the skin was tightly controlled by using a digital micro-dermascope, it was not possible to accurately assess the penetration depth and angle in the epidermis. Previous studies have asserted that cowhage spiculae should be inserted near the basal membrane of the epidermis or at least into the stratum granulosum in order to evoke a sensory response (66). In addition, it has been suggested that in order to elicit a neuronal excitation of pruriceptive fibres the spiculae must be inserted adjacent to a nerve terminal (8, 57), which is less likely to occur when applying very few spiculae (33). On the other hand, a previous study found that insertion of a single cowhage spicule produced itch in 80% of all subjects (10), suggesting that the cowhage spiculae application is generally a robust approach. Lastly, 2 findings counter the notion of spiculae insertion variability as a major issue for model reliability: (i) a considerable effect should also be observable for measures of relative reliability measures, which is not the case; (ii) a detectable, albeit small, micro-flare reaction was generally observed post-application for all subjects, even those reporting no or very low itch in response to cowhage, indicating adequate insertion of physiologically active spiculae. Thus, individual neurophysiological or psychophysical differences should explain the majority of the observed variability. In the present study a highly homogenous group of young healthy male caucasian subjects was recruited to match typical conditions of phase I proof-of-concept studies in early drug development. It is highly conceivable that at least higher inter-individual disagreement would be present in a more diverse sample composition, i.e. on parameters such as gender, age, ethnicity and morbidity.

Another finding pertaining to the reliability assessment was that peak itch intensity generally performed better than itch AUC, both in terms of absolute and relative reliability. These are both very common outcome measures for quantifying itch in response to a pruritogenic stimulus and/or a potential intervention (11, 50, 65, 67, 68), and both measures can easily be assessed in most study designs. However, since peak itch intensity does not convey information about the temporal profile of the perceived itch (which is important in a clinical context), it may be less relevant as a primary outcome measure and could be exchanged for itch AUC. Moreover, based on the herein proposed sample size calculations adjusted by the obtained ICC values it would not drastically increase the minimum required sample size to use itch AUC instead of peak itch intensity in studies where a cross-over design is possible.

**Conclusion**

The present study is the first to assess and compare the test–retest reliability of human surrogate models of itch. Cowhage application using negative-grip tweezers and application of histamine with an SPT lancet and a pressure of 120 g showed moderate-to-excellent intra-individual test–retest reliability and moderate inter-individual reliability. Itch provocation with 1% histamine and, to a lesser extent, insertion of 25 spiculae proved to have the best intra-individual test–retest reliability in a 1-week observation period. Moderate inter-subject reliability and moderate-to-excellent intra-subject reliability favour cross-over study designs. The extent of the perceived area of itch did not differ between cowhage- and histamine-induced itch and, generally, the variability for this parameter was high.

**ACKNOWLEDGEMENTS**

The authors thank the volunteers for participating in this study and Frank Birkhøy Schultz Andersen for technical assistance and development of the computerized VAS. Cowhage spicules were kindly supplied by Professor Robert LaMotte, Yale University.

*The authors have no conflicts of interest to declare.*

**REFERENCES**

Reliability of human models of itch


