ORIGINAL ARTICLE

Prurigo Activity Score (PAS): validity and reliability of a new instrument to monitor chronic prurigo

J. Pölking,[1] C. Zeidler,[1] F. Schedel,[1] N. Osada,[1] M. Augustin,[2] D. Metze,[1] M. P. Pereira,[1] G. Yosipovitch,[3] J.D. Bernhard,[4] S. Ständer[1,*]

[1]Department of Dermatology, Center for Chronic Pruritus, University Hospital Münster, Münster, Germany
[2]Comprehensive Wound Center, Institute for Health Services Research in Dermatology and Nursing (IVDP), University Medical Center Hamburg-Eppendorf (UKE), Hamburg, Germany
[3]Miller School of Medicine, University of Miami, Miami, FL, USA
[4]University of Massachusetts Medical School, Worcester, MA, USA
*Correspondence: S. Ständer E-mail: sonja.staender@uni-muenster.de

Abstract

Background Currently available tools to monitor patients with chronic prurigo over time focus on pruritus and quality of life parameters, while no instrument objectively assessing the pruriginous lesions is yet available.

Objective The objective of this study was to develop a physician-assessed Prurigo Activity Score (PAS), a new tool to monitor the distribution and activity of chronic prurigo lesions and to evaluate its reliability and validity.

Methods The 7-item PAS questionnaire as well as validated pruritus intensity scales (VAS, NRS) and a skin-related quality of life score (DLQI) were completed for 264 patients (172 females, age 61 years) at least twice over a period of 2 years. In addition, a 60-min test–retest reliability test was performed by four experts for a random sample of 12 patients.

Results The PAS showed good test-retest reliability (Cohens $\kappa > 0.61$; Cronbach-alpha > 0.76), ordinal or metric items showed high inter-rater reliability (Kendalls > 0.61) and items recording the number of lesions correlated significantly to each other ($P < 0.001$). The highest correlation to external constructs was achieved with DLQI. The feasibility test conducted by four raters indicated the suitability of PAS for tracking chronic prurigo in the clinical setting.

Discussion The PAS is a useful tool to objectively monitor pruriginous lesions in chronic prurigo patients over time. The sensitivity of change in the PAS score should be analysed in future studies.

Received: 6 February 2018; Accepted: 13 April 2018

Conflict of interests

None.

Funding sources

Medical writing was supported financially by Trevi Therapeutics. We acknowledge the support by the German Federal Ministry of Education and Research (BMBF; No. 01KG1305 to SST) and the European Academy of Dermatology and Venereology (EADV, No. 2016-012 to MP).

Introduction

Chronic prurigo (CPG) is a highly pruritic, chronic skin disease emerging in patients with chronic pruritus and prolonged scratching over a long period of time.[1–3] CPG is characterised by the presence of symmetrically distributed, indurated papules, nodules or plaques.[2,3] According to the predominance of lesions, subtypes of CPG are defined as nodular (prurigo nodularis; PN), papular, plaque and umbilicated prurigo.[3] The vicious cycle of itching and scratching with visible skin lesions considerably affects the quality of life of these patients.[4]

The pathogenesis of CPG is not yet fully understood, and its chronicity and relapsing nature make its treatment challenging.[1,5,6] CPG lesions are frequently treatment resistant, healing slowly, if at all.[3] Currently available tools to monitor CPG patients during therapy and in clinical trials are focusing primarily on itch and its influence on the quality of life, sleep or psychological factors[7] while the structured assessment of the course of lesions remains neglected. However, the type of lesions (papules or nodules), scratching activity (monitored by excoriations or crusts on top of the pruriginous lesions pointing to scratching), distribution and number are indicative of the

severity and the progress of the disease.[2,8] Furthermore, although it is assumed that the healing of excoriations correlates with reduced scratching and disease improvement, so far it is not possible to assess objectively the scratching behaviour of the patient, unless devices such as actigraphy are used, which frequently do not allow inter-individual comparisons.[9]

The absence of an instrument suitable to evaluate the CPG characteristics and its course leads to limited objectivity and comparability between treatments in clinical studies and in routine care. Previous studies tried to monitor the resolution of lesions[10–13]. In one study, the total number of palpable nodules was counted and the average diameter of three nodules at each leg was measured[10]. Two other studies selected a physician global assessment approach by semi-quantitative judgment of pruritus relief and reduction/flattening of nodules on a three-point Likert scale (successful/slight, improvement/unsuccessful)[11,12]. However, none of these scales were validated or used in subsequent trials. So far, we used preliminary versions of a prurigo score (PRUNOSI, Scratch Symptom Score (SSS)) in two clinical studies (ClinicalTrials.gov: NCT00507832, German Register Clinical Trials: DRKS00005284).[13,14] PRUNOSI evaluated the type (papules, nodules, lichenification, excoriations and crusts) and percentage of lesions in affected body areas versus the whole body area. In one of these studies, we compared atopic dermatitis to CPG and detected a significant difference in the score between these two groups.[14] In the other double-blind, randomised, controlled study, in which the antipruritic effect of topical pimecrolimus versus topical hydrocortisone in non-atopic CPG was evaluated,[13] PRUNOSI showed a significant score difference before and after the therapy.[13] The first signs of CPG improvement were healing of excoriations, crusts and erythema.[13] Based on these results, we developed the prurigo activity score (PAS). The PAS includes metrics for grading active pruriginous lesions and a method for measuring designated representative lesions, allowing the determination, classification and monitoring of CPG lesions in the clinical setting. Hence, the aim of this study was to assess the ctivity, reliability and validity of this new prurigo evaluation tool.

Methods

Patient population

Adult patients with clinical and histological diagnosed CPG of papular, nodular, plaque or ulcerating type were included in this prospective, longitudinal, uncontrolled, non-interventional study over a period of 2 years (Fig. S1). Pregnancy or lactation, drug abuse or presence of single excoriations without chronic scratch lesions were exclusion criteria. All patients gave written informed consent. The Ethics Committee of the University of Münster approved the study, which was registered at the German Clinical Trials Register (DRKS00005383).

Prurigo activity score

Basis for the development of PAS were the earlier applied PRUNOSI and SSS scores.[13,15] The PAS questionnaire version 0.9 consists of 7 items (Table S1). The items of the PAS evaluate the type (visible lesions: item 1a; predominant lesions: item1b), estimated number (item 2), distribution (item 3, 4) and size (biggest lesion: item 6a; representative lesion: item 6b) of pruriginous lesions, the representative body area and exact number of lesions (item 5), the activity in terms of percentage of pruriginous lesions with excoriations/crusts on top (reflecting active scratching; item 7a) and the percentage of healed pruriginous lesions (reflecting healing of CPG; item 7b).

Further assessment scales

To assess the validity of the PAS, patients were additionally asked to fill out electronically established scales at each visit (on the visual analogue scale (VAS): average pruritus of the past 24 h, worst and average pruritus of the past 4 weeks;[16,17] on the numeric rating scale (NRS): average pruritus of the past 24 h;[18] dermatology quality of life index (DLQI),[19] dynamic pruritus score (DPS)[20]). The participating dermatologists were asked to complete once a feasibility questionnaire on the purpose, comprehensibility, completeness and suitability of the PAS for CPG assessment.

Procedures

To assess the reliability and objectivity of the PAS, a sample of 12 patients with active CPG was randomly chosen from the total study population without considering their treatment regimen or VAS scores. Four experienced dermatologists ('raters') completed the PAS once for each patient and additionally the feasibility questionnaire. For test–retest reliability, the PAS procedure was repeated 1 h later.

Statistical analysis

Statistical analysis was performed with SPSS 23.0 for Windows. Inferential statistics were intended to be exploratory and no adjustment for multiplicity was made. The level of significance was defined at $P \leq 0.05$. Mean and medians, standard deviations, minimal and maximal values as well as frequencies were calculated. Sample differences in categorical variables were tested by Pearson-χ^2-test with significant values indicating a relation. The Wilcoxon-test was applied for PAS differences between the first and last visit and McNemar's test was used for paired binary data.

Test-retest reliability For every metric item mean, standard deviation, median, minimal and maximal values were calculated. Cronbachs-alpha or interclass-coefficient (ICC) was determined for continuous variables and Cohen's-Kappa for ordinal or nominal values.

Inter-rater-reliability ICC was calculated for metric and Fleiss' Kappa for nominal variables. Inter-rater-reliability of ordinal-

scaled items were determined by Kendall's coefficient of concordance.

Inter-item-correlation Spearman's correlation coefficients were calculated between ordinal-scaled items. A positive correlation (>0) meant that the items were related, while a coefficient of 0 indicated that there was no relation.[21] Differences in the values of dichotomous items regarding ordinal or metric scaled items were tested by the non-parametric Mann–Whitney *U* test. The difference of the dichotomous item regarding the metric item was also calculated. Dichotomous values were pairwise cross-tabulated and their correlation was assessed by χ^2-test.

Convergent validity Spearman's correlation coefficient was calculated between ordinal or metric PAS items and the external constructs VAS-today, VAS average, VAS worst, NRS and DLQI.

Concurrent validity For concurrent validity testing, the DLQI scores were first grouped into bandings according to Hongbo *et al.*[22] The Kruskal–Wallis test was used to assess ordinal and metric scaled PAS item differences in between bandings, and for dichotomous items, the Mann–Whitney *U* test was performed.

Risk analysis The Mann–Whitney U test was used to evaluate differences among genders regarding ordinal or metric scaled items and the χ^2-test was used to assess dichotomous items. Correlations between the age and ordinal or metric scaled items were calculated with Spearman's rank correlation coefficient. Differences among dichotomous items regarding age were determined via the Mann–Whitney U test.

Development of the score A multivariate regression model, which contained items 1b, 2, 3, 7a and 7b, was adapted. The items (1a, 4, 5, 6) are descriptive and were not included to the scoring system. For each item, an estimated model parameter was calculated, which determined the influence of the feature on the DLQI, NRS and VAS.

Results

Between 01/2013 and 01/2015, 264 patients (172 females (65.2%); mean age at first visit: 61.0 ± 15.2 years) were enrolled. Average VAS pruritus intensity in the 4 weeks before enrolment was 5.8 ± 2.9 and DLQI score 9.8 ± 6.7 (Table 1).

Table 1 Demographics and baseline characteristics

		Total population	Reliability random sample
Number	*N*	264	12
Female	Percentage	65.2	58.3
Age	Mean (SD)	60.98 (15.2)	63.0 (11.2)
CPG underlying cause (%)	Dermatosis	23.1	8.3
	Neurologic	5.7	0
	Systemic	4.6	16.6
	Psychic	1.1	8.3
	Multifactorial	58.7	58.3
	Unclear	3.4	8.3
	No answer	3.4	
VAS average (past 24 h)	*N*	238	11
	Mean (SD)	4.90 (3.2)	4.55 (2.7)
VAS worst (past 4 weeks)	*N*	238	11
	Mean (SD)	6.79 (3.0)	7.34 (3.0)
VAS average (past 4 weeks)	*N*	237	11
	Mean (SD)	5.75 (2.9)	6.05 (2.9)
NRS average (past 24 h)	*N*	240	11
	Mean (SD)	5.17 (2.8)	5.18 (2.7)
DLQI (past 7 days)	*N*	232	12
	Mean (SD)	9.79 (6.7)	11.08 (6.7)
DPS	*N*	144	10
	Mean (SD)	51.05 (37.0)	34.20 (56.8)

CPG, chronic prurigo; DLQI, dermatologic life quality index; DPS, dynamic pruritus score; NRS, numeric rating scale; VAS, visual analogue scale; SD, standard deviation.

Test-retest reliability in the random sample of 12 patients (Table S2)

Excellent congruence was shown for items 2 (0.901), 3 (1.000), 5 regarding the number of lesions (0.907) and 6a for the biggest monitor lesion (elevation: 0.885; longitudinal: 0.764; crosswise: 0.893). Good congruence was demonstrated for item 1a 'plaques' (0.823), for item 6b regarding the representative lesion (elevation: 0.764; cross: 0.853; across: 0.770), for items 7a (0.885) and 7b (0.838) and for item 4 in the characteristics 'trunk ventral' (0.719 P), 'scalp' (0.860) and 'face' (0.855). For all other item, 4 variables congruence was moderate. The congruence for item 1b was questionable (0.695) and for item 1a 'ulcers' low (0.253). In item 5, the chosen representative area was not always the same at test and retest (0.566).

Inter-rater reliability (Table S2)

Very good congruence among the raters was demonstrated for the number of lesions in the representative area in item 5 (0.894), the elevation of the biggest lesion (0.814) and item 2 (0.806), good for the elevation and diameter of the representative lesion (0.707–0.733) and for items7a (0.734) and 7b (0.727), moderate for item 1b (0.462) and for the expression 'whole body except head' (0.438) in item 4, slight for 'umbilicated ulcers' (0.365) and 'nodules' (0.333) of item 1a, and for 'whole body head included' (0.295) and 'trunk ventral' (0.210) of item 4 and slight for item 3 (0.062).

Inter-item correlation (Table S3)

Of the 289 (17 × 17 items) tested correlations, 72 were significant (Table S3). High associations were found among the type of lesions themselves (item 1a and 1b), distribution (item 3 and 4: $P < 0.001$; 3 and 5: $P = 0.006$; 4 and 5: $P = 0.044$), lesion size

(item 6a and 6b: $P < 0.001$) and disease activity (item 7a and 7b: $P < 0.001$). Interestingly, the estimated number of lesions (item 2) was highly associated with all other items (item 1a (nodules), 3–7: $P < 0.001$; item 1 b: $P = 0.003$; item 1a plaques: $P < 0.046$) and seems to represent a good severity and activity parameter of prurigo. Likewise, item 7a (scratching activity) showed also a large number of associations to most other items (except item 1a papules and maculae, 1b, 6a elevation, and 6b).

Convergent validity (Table 2)

The PAS items 1b (predominant lesion type), 2 (estimated number) and 7a (excoriations/crusts) showed good correlations with mostly high significances to all used instruments such as the VAS and NRS scores and DLQI (Table 2). Item 3 (distribution) and 7b (healed lesions) correlated to DLQI, NRS and some VAS scores.

Concurrent validity (Table S4)

With increasing DLQI scores (reflected by DLQI bands), the percentage of patients with papules (item 1a; $P = 0.003$) or healed lesions (item 7b; $P = 0.176$) decreased. Additionally, the percentage of patients with item 1a ulcers ($P = 0.029$), or plaques($P = 0.004$), the estimated number of lesions (item 2; $P < 0.001$), the percentage of patients with disseminated lesions (item 3; $P < 0.001$), or affected body areas (item 4; $P = 0.145$), the percentage of excoriations and crusts (item 7a; $P < 0.001$) and the size of the biggest and representative lesion (item 6a, 6b) increased. This was also the case for the correlation of DLQI with item 1b (predominant prurigo; $P = 0.011$). There were no significant differences among the DLQI bandings regarding nodules ($P = 0.270$), maculae ($P = 0.549$) and number of lesions in the representative area ($P = 0.332$).

Risk analysis

Neither sex nor age correlated significantly with any PAS item.

Feasibility

All four raters completing the feasibility questionnaire considered the instructions to use the PAS to be clear and its readability to be good. The item related questions were easy to understand, all important points were addressed and overall the PAS was considered an useful instrument. One rater considered the PAS too long (average completion time 2.6 ± 1.0 min), two raters expressed decision difficulties for item 6 and one for item 5.

Scoring system

The DLQI achieved the best predictive and adaptive quality compared to NRS and VAS (Fig. S2). Item 2 and 7a had a significant effect on the DLQI (Table 3). However, trends indicating that advanced prurigo lesions (plaque-type or ulcerative prurigo predominant) and disseminated distribution produced a higher score. A higher percentage of healed lesions than 0%–24% resulted in a reduction of the PAS total score (Table 3). The score is calculated via summation of the scoring values, added up with 123 and afterwards divided by 10. This results in a range of values from 1.3 to 21.3.

Discussion

Our aim was to assess the feasibility and validity of the PAS, a novel instrument to monitor CPG. We hypothesised that PAS describes accurately the clinical features of CPG and its activity. Based on the demographic characteristics, the reliability sample was representative of the total study population. Results can thus be extrapolated to the total population.

Table 2 Convergent validity in the study population at baseline ($n = 264$)

	Item 1a: Papules	Item 1a: Nodules	Item 1a: Plaques	Item 1a: Ulcer	Item 1a: Macule	Item 1b	Item 2	Item 3	Item 4
VAS 24 h	Diff. = 0.526	Diff. = −0.016	Diff. = −0.472	Diff. = −1.158	Diff. = 1.071	r = **0.192****	r = **0.219****	Diff. = 0.976	r = 0.183
VAS worst	Diff. = 0.531	Diff. = −0.002	Diff. = −0.646	Diff. = **−1.268***	Diff. = **1.423***	r = **0.193****	r = **0.285*****	Diff. = **1.381*****	r = 0.210
VAS average	Diff. = 0.368	Diff. = 0.098	Diff. = −0.22	Diff. = **−1.782***	Diff. = **1.534***	r = **0.150***	r = **0.265****	Diff. = **1.26****	r = 0.205
NRS	Diff. = 0.22	Diff. = −0.013	Diff. = −0.48	Diff. = −0.8	Diff. = **1.39***	r = **0.177***	r = **0.225****	Diff. = **1.2****	r = 0.170
DLQI	Diff. = 2.95**	Diff. = −1.01	Diff. = **−2.94****	Diff. = **−3.5***	Diff. = 0.76	r = **0.229*****	r = **0.370*****	Diff. = **4.44*****	r = **0.299****

	Item 5	Item 6a: Elevation	Item 6a: Cross	Item 6a: Across	Item 6b: Elevation	Item 6b: Cross	Item 6b: Across	Item 7a	Item 7b
VAS 24 h	r = 0.119	r = 0.004	r = 0.093	r = 0.072	r = −0.060	r = 0.019	r = −0.011	r = **0.203****	r = 0.108
VAS worst	r = 0.158	r = −0.063	r = 0.031	r = 0.030	r = −0.077	r = 0.005	r = −0.009	r = **0.267*****	r = 0.130
VAS average	r = 0.151	r = −0.011	r = 0.059	r = 0.065	r = −0.081	r = 0.028	r = 0.006	r = **0.241*****	r = **0.136***
NRS	r = 0.097	r = 0.040	r = 0.067	r = 0.054	r = −0.033	r = 0.005	r = −0.017	r = **0.248*****	r = **0.143***
DLQI	r = 0.035	r = −0.012	r = **0.200****	r = **0.180***	r = −0.118	r = **0.154***	r = 0.121	r = **0.323*****	r = **0.154***

*$P \leq 0.05$; **$P \leq 0.01$; ***$P \leq 0.001$.

DLQI, dermatologic life quality index; DPS, dynamic pruritus score; NRS, numeric rating scale; VAS, visual analogue scale.

Table 3 Score development using a multivariate regression model and estimated model parameter determining the influence of the PAS items on DLQI

Item	Characteristics	Estimate	SD	*P*-value	Scoring (Estimate ×10)
1b (predominantlesions)	Maculae	Reference			0
	Papules	−0.951	4.413	0.830	−10
	Nodules	−1.020	4.414	0.818	−10
	Plaques	4.853	4.563	0.289	49
	Ulcer	2.594	4.847	0.593	26
2 (estimated number)	>100	Reference			0
	0–19	−4.731	1.880	0.013	−47
	20–100	−1.610	1.382	0.246	−16
3 (distribution)	Disseminated	Reference			0
	Localised	−1.877	1.326	0.159	−19
7a (reflecting active scratching)	0%–25%	Reference			0
	26%–50%	0.683	1.608	0.671	7
	51%–75%	4.136	1.959	0.037	41
	76%–100%	2.754	2.238	0.221	28
7b (reflecting healing of CPG)	75%–100%	−1.197	1.889	0.527	−12
	50%–74%	−1.031	1.786	0.565	−10
	25%–49%	−3.434	1.883	0.070	−34
	0%–24%	Reference			0

SD, standard deviation.

All items, except 1a, 4 and 5 exerted an at least acceptable (Cronbachs-Alpha > 0.7[23]), respectively, good (Cohens κ 0.61-08[24]) test–retest reliability. To reduce the risk of erroneously high coefficients due to memory effect, physicians had to first examine all patients before retesting. The coexistence of 'ulcers' and 'umbilicated ulcers' could be the reason for the low outcome of item 1a. 'Upper arm right', had the lowest coefficient in item 4. A possible reason could be the high number of choices in this item. Therefore, it was suggested to delete the expressions 'whole body including' or 'excluding head'. For item 5, the low congruence could possibly be explained by the free choice of the representative area.

The high objectivity of the either ordinal or metric scaled items 2, 6a, 6b, 7a and 7b was shown by the high congruence (Kendall's coefficient resp. ICC > 0.61) among the raters. Items with nominal variables, such as 1a, 1b, 3 and 4, achieved less reliable results. For item 1a, the difficult differentiation of lesions could be an explanation for the low result. As this item is only qualitative, it was proposed to use it in the future for documentation only. The two characteristics 'ulcerated' and 'umbilicated prurigo' should be combined in the future, as only two physicians have considered them. Here again, the low inter-rater reliability of item 4 indicates that the expressions 'whole body', ' with or without head' should be deleted from the item.

As the PAS consists of metric, ordinal and nominal scaled items, it was not expected to observe high inter-item correlations between all items. Significant was the correlation of the items describing the type of lesions, namely 1a with 1b. Items recording the number of lesions that is 2, 3, 4 and 5 correlated also significantly to each other. The strong correlation of item 2 and 5, indicates that the number of pruriginous lesions in the representative area is representative of the number on the whole body, solving the difficulty of nodule counting described previously.[2] The items 7a (scratching activity) and 7b (healed lesions), which depict the prurigo activity, showed the highest inter-item correlation ($r = 0.675$; $P < 0.001$). Thus, the expectation that items covering similar content are consistent could be confirmed. Nevertheless, also relationships between items asking different contents could be observed, as for example the significant correlations between the estimated number of lesions and the prurigo activity. High scratch activity (item 7a) resulted in high number of lesions (item 2, $r = 0.401$), which in turn correlated highly significantly with a low number of healed lesions ($r = 0.251$; $P < 0.001$) as previously described.[25–27] In addition, intensive scratching increases not only the number of lesions but also the distribution of the lesions throughout the body (high correlation between items 7a and 7b with 4). In the absence of nodules, plaques and ulcers (item 1a) the number of crusts was lower, which is in agreement with the observation that crusts decrease with lesion healing.[2] The 11, 12 and 10, respectively, out of 17 possible significant correlations among items 1b, 2 and 7a (Table S3), demonstrated the importance of the predominant prurigo type, of the number of lesions and of their activity for capturing prurigo as already presented before.[2] More, item 2 (estimated number of lesions) was highly associated with all other

items and seems to represent a good severity and activity parameter of prurigo.

To test if it accurately covers the desired aspects of CPG, the PAS questionnaire was correlated with established instruments (convergent validity).[15,28,29] As there is no validated instrument to assess CPG lesions,[15] we have chosen instruments assessing pruritus intensity (VAS, NRS) and quality of life (DLQI). Interestingly, the highest correlation was achieved with DLQI, demonstrating that CPG progress affects considerably the quality of life. High significances were detected specifically for item 1b (predominant type), item 2(estimated number), item 3 (distribution) and item 7a (scratch activity). These items also correlated best with the intensity scales both with long (VAS 4 weeks) and short-term (VAS, NRS 24 h) assessments. Accordingly, these items seem to be the most representative items for CPG. Furthermore, item 2 (estimated number) seems to be of particular interest as it shows the highest inter-item correlation and high to very high significant correlations to external constructs. Future studies on sensitivity of change in PAS will aim to test this. Item 7b showed only good relation to the external constructs and inter-item correlations. This might reflect the difficulty to determine how many lesions are already healed when the patient is assessed for the first time without previous photos or visits. Usually, CPG heals leaving a scar at the previously affected area. In addition, it might be difficult to assess the exact extent of previous skin damage in case of concomitant presence of a dermatosis or of scars of other origin. Therefore, and due to the relatively high number of lesions, items 2 and 7 refer to estimations and not to exact counting of lesions. Future digital imaging measuring applications of pruriginous lesions may solve this problem. The lacking significance for the representative area (item 5) is possibly due to different areas chosen. Two raters expressed difficulties in choosing a representative area, as a clear demarcation of the areas was not always obvious. Moreover, the number of lesions in the representative area is probably not sufficient to relate to the quality of life of the patient, something that should be further investigated. However, overall the raters appraised the PAS as a suitable CPG assessment instrument.

In this study, it was possible to validate an instrument for the clinical assessment of the manifold aspects of CPG. The acceptable to good reliability and the moderate to very good objectivity of the majority of the items allow the conclusion that PAS can be regarded as a reliable and objective tool. The above-mentioned changes should be implemented and the reliability test should be repeated to get a stable PAS. The weak correlation of pruritus intensity with the percentage of healed lesions showed that pruritus intensity is not sufficiently representing the number of pruriginous lesions. This underlines that the PAS is required as an independent instrument for recording CPG lesion. The estimated number of lesions and the prurigo activity are especially relevant parameters. They not only react sensitively to pruritus decrease, they provide also information on patients' quality of life.

Acknowledgments

We thank Charlotte Schürmann for the contribution to the recruitment, Laura Kerschke for helping with statistics and Helena Karajiannis for her assistance in the preparation of the manuscript.

References

1 Tsianakas A, Zeidler C, Stander S. Prurigo nodularis management. *Curr Probl Dermatol* 2016; **50**: 94–101.
2 Schedel F, Schurmann C, Metze D, Stander S. Prurigo. Clinical definition and classification. *Hautarzt* 2014; **65**: 684–690.
3 Pereira MP, Steinke S, Zeidler C *et al.* European academy of dermatology and venereology European prurigo project: expert consensus on the definition, classification and terminology of chronic prurigo. *J Eur Acad Dermatol Venereol* 2017. http://doi.org/10.1111/jdv.14570.
4 Zeidler C, Stander S. The pathogenesis of Prurigo nodularis–'Super-Itch' in exploration. *Eur J Pain* 2016; **20**: 37–40.
5 Fostini AC, Girolomoni G, Tessari G. Prurigo nodularis: an update on etiopathogenesis and therapy. *J Dermatological Treat* 2013; **24**: 458–462.
6 Vaidya DC, Schwartz RA. Prurigo nodularis: a benign dermatosis derived from a persistent pruritus. *Acta Dermatovenerol Croat* 2008; **16**: 38–44.
7 Pereira MP, Stander S. Assessment of severity and burden of pruritus. *Allergol Int* 2017; **66**: 3–7.
8 Schedel F, Schürmann C, Augustin M, Ständer S. Prurigo nodularis: introduction of a re-defined classification and prurigo activity score (PAS). *Acta Derm Venereol* 2013; **93**: 610.
9 Silverberg JI. Practice gaps in pruritus. *Dermatol Clin* 2016; **34**: 257–261.
10 Wong SS, Goh CL. Double-blind, right/left comparison of calcipotriol ointment and betamethasone ointment in the treatment of Prurigo nodularis. *Arch Dermatol* 2000; **136**: 807–808.
11 Mazza M, Guerriero G, Marano G, Janiri L, Bria P, Mazza S. Treatment of prurigo nodularis with pregabalin. *J Clin Pharm Ther* 2013; **38**: 16–18.
12 Taefehnorooz H, Truchetet F, Barbaud A, Schmutz JL, Bursztejn AC. Efficacy of thalidomide in the treatment of prurigo nodularis. *Acta Derm Venereol* 2011; **91**: 344–345.
13 Siepmann D, Lotts T, Blome C *et al.* Evaluation of the antipruritic effects of topical pimecrolimus in non-atopic prurigo nodularis: results of a randomized, hydrocortisone-controlled, double-blind phase II trial. *Dermatology* 2013; **227**: 353–360.
14 Stander S, Pogatzki-Zahn E, Stumpf A *et al.* Facing the challenges of chronic pruritus: a report from a multi-disciplinary medical itch centre in Germany. *Acta Derm Venereol* 2015; **95**: 266–271.
15 Stander S, Blome C, Breil B *et al.* Assessment of pruritus - current standards and implications for clinical practice : consensus paper of the Action Group Pruritus Parameter of the International Working Group on Pruritus Research (AGP). *Hautarzt* 2012; **63**: 4–31.
16 Phan NQ, Blome C, Fritz F *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 patients with chronic pruritus. *Acta Derm Venereol* 2012; **92**: 502–507.
17 Reich A, Heisig M, Phan NQ *et al.* Visual analogue scale: evaluation of the instrument for the assessment of pruritus. *Acta Derm Venereol* 2012; **92**: 497–501.
18 Jenkins HH, Spencer ED, Weissgerber AJ, Osborne LA, Pellegrini JE. Correlating an 11-point verbal numeric rating scale to a 4-point verbal rating scale in the measurement of pruritus. *J Perianesth Nurs* 2009; **24**: 152–155.

19 Finlay AY, Khan GK. Dermatology Life Quality Index (DLQI)–a simple practical measure for routine clinical use. *Clin Exp Dermatol* 1994; **19**: 210–216.
20 Stander S, Blome C, Anastasiadou Z *et al.* Dynamic pruritus score: evaluation of the validity and reliability of a new instrument to assess the course of pruritus. *Acta Derm Venereol* 2017; **97**: 230–234.
21 Eckstein P. Angewandte Statistik mit SPSS. Springer Gabler, Wiesbaden, 2016.
22 Hongbo Y, Thomas CL, Harrison MA, Salek MS, Finlay AY. Translating the science of quality of life into practice: what do dermatology life quality index scores mean? *J Invest Dermatol* 2005; **125**: 659–664.
23 George D, Mallery P. SPSS for Windows Step by Step: A Simple Guide and Reference. 11.0 Update ed.Allyn and Bacon, Boston, 2003.
24 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; **33**: 159–174.
25 Schneider G, Hockmann J, Stumpf A. Psychosomatic aspects of prurigo nodularis. *Hautarzt* 2014; **65**: 704–708.
26 Stander S, Darsow U, Mettang T *et al.* S2k guideline–Chronic Pruritus. *J Dtsch Dermatol Ges* 2012; **10**(Suppl 4): S1–S27.
27 Zeidler C, Stander S. Therapy of prurigo nodularis. *Hautarzt* 2014; **65**: 709–713.
28 Gernart M, Tsianakas A, Zeidler C *et al.* ItchApp(c): an App-based eDiary for assessment of chronic pruritus in clinical trials. *Acta Derm Venereol* 2017; **97**: 601–606.
29 Weisshaar E, Stander S, Gieler U, Matterne U, Darsow U. Development of a German language questionnaire for assessing chronic pruritus (AGP-questionnaire): background and first results. *Hautarzt* 2011; **62**: 914–927.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Figure S1. Study design and use of instruments every 3 months at the study visits.
Figure S2. Graphical representation of the predicted values for the PAS and the actual values based on VAS (a) NRS (b) and DLQI (c).
Table S1. Prurigo Activity Score (PAS).
Table S2. Test-retest reliability and Inter-rater reliability in the random sample of 12 patients (CI, confidence interval; E, elevation; L, longitudinal; C, crosswise).
Table S3. Inter-item correlation in the whole prurigo population (n = 264).
Table S4. Concurrent validity in the whole prurigo population (n = 264).