

## Pruritus intensity scales across Europe: a prospective validation study







M. Storck, S. Sandmann, P. Bruland, M. P. Pereira, S. Steinke, C. Riepe, Iñaki Soto Rey, S. Garcovich, M. Augustin, C. Blome, S. Bobko, F. J. Legat, N. Potekaev, A. Lvov, L. Misery, W. Weger, A. Reich, E. Şavk, M. Streit, E. Serra Baldrich, J. C. Szepietowski, M. Dugas, S. Ständer, C. Zeidler

### Angaben zur Veröffentlichung / Publication details:

Storck, M., S. Sandmann, P. Bruland, M. P. Pereira, S. Steinke, C. Riepe, Iñaki Soto Rey, et al. 2021. "Pruritus intensity scales across Europe: a prospective validation study." *Journal of the European Academy of Dermatology and Venereology* 35 (5): 1176–85.  
<https://doi.org/10.1111/jdv.17111>.

## ORIGINAL ARTICLE

# Pruritus Intensity Scales across Europe: a prospective validation study

M. Storck,<sup>1,\*</sup>  S. Sandmann,<sup>1</sup> P. Bruland,<sup>1,2</sup> M.P. Pereira,<sup>3</sup>  S. Steinke,<sup>3</sup> C. Riepe,<sup>3</sup> I. Soto-Rey,<sup>1,4</sup> S. Garcovich,<sup>5</sup>  M. Augustin,<sup>6</sup>  C. Blome,<sup>6</sup> S. Bobko,<sup>7</sup> F.J. Legat,<sup>8,9</sup> N. Potekaevev,<sup>7</sup> A. Lvov,<sup>7</sup>  L. Misery,<sup>10</sup>  W. Weger,<sup>8</sup> A. Reich,<sup>11</sup> E. Şavk,<sup>12</sup> M. Streit,<sup>13</sup> E. Serra-Baldrich,<sup>14</sup> J.C. Szepietowski,<sup>15</sup> M. Dugas,<sup>1</sup> S. Ständer,<sup>3</sup> C. Zeidler<sup>3</sup>

<sup>1</sup>Institute of Medical Informatics, University of Münster, Münster, Germany

<sup>2</sup>inIT – Institute Industrial IT, Ostwestfalen-Lippe University of Applied Sciences, Lemgo, Germany

<sup>3</sup>Department of Dermatology and Center for Chronic Pruritus, University Hospital Münster, Münster, Germany

<sup>4</sup>IT-Infrastructure for Translational Medical Research, Faculty of Applied Computer Science, Faculty of Medicine, University of Augsburg, Augsburg, Germany

<sup>5</sup>Dermatology, F. Policlinico Universitario A. Gemelli IRCCS, Università Cattolica del Sacro Cuore, Rome, Italy

<sup>6</sup>Institute for Health Services Research in Dermatology and Nursing (IVDP), University Medical Center Hamburg-Eppendorf, Hamburg, Germany

<sup>7</sup>Moscow Scientific and Practical Center of Dermatovenerology and Cosmetology, Moscow, Russia

<sup>8</sup>Department of Dermatology, Medical University of Graz, Graz, Austria

<sup>9</sup>Das Kurhaus, Bad Gleichenberg, Austria

<sup>10</sup>Department of Dermatology, University Hospital of Brest, Brest, France

<sup>11</sup>Department of Dermatology, Institute of Medical Sciences, Medical College of Rzeszow University, Rzeszów, Poland

<sup>12</sup>Department of Dermatology, Aydin Adnan Menderes University, Aydin, Turkey

<sup>13</sup>Department of Dermatology, Kantonsspital Aarau, Aarau, Switzerland

<sup>14</sup>Cutaneous Allergy Unit, Department of Dermatology, Hospital Sant Pau, Universitat Autònoma, Barcelona, Barcelona, Spain

<sup>15</sup>Department of Dermatology, Venereology and Allergology, Wrocław Medical University, Wrocław, Poland

\*Correspondence: M. Storck. E-mail: michael.storck@uni-muenster.de

## Abstract

**Background** Chronic pruritus (CP) is a subjective symptom, and it is necessary to assess its intensity with validated patient-reported outcome tools in order to allow determination of the treatment course.

**Objectives** So far, the itch intensity scales were validated in small cohorts and in single languages. Here, we report the validation of the numerical rating scale, the verbal rating scale and the visual analogue scale for the worst and average pruritus intensity in the last 24h in several languages across Europe and across different pruritic dermatoses.

**Methods** After professional translation, the intensity scales were digitized for use as a tablet computer application. Validation was performed in clinics for Dermatology in Austria, France, Germany, Italy, Poland, Russia, Spain, Switzerland and Turkey.

**Results** A total of 547 patients with contact dermatitis, chronic nodular prurigo, psoriasis vulgaris, lichen planus or cutaneous T-cell lymphoma were included. The intensity scales showed a high level of reproducibility and inter-correlations with each other. The correlation with the Dermatology Life Quality Index was weak to strong in nearly all countries and dermatoses with the exception of France and patients with chronic nodular prurigo, for which no statistically significant correlations were found.

**Conclusions** The numerical rating scale, the verbal rating scale und the visual analogue scales are valid instruments with good reproducibility and internal consistency in German (Germany, Austria, Switzerland), French, Italian, Polish, Russian, Spanish and Turkish for different pruritic dermatoses. VAS worst was the best reproducible and consistent measuring instrument in all countries.

Received: 15 October 2020; Accepted: 1 December 2020

## Conflicts of interest

None.

## Funding sources

This article was supported by the European Academy of Dermatology and Venereology (EADV, No. 2014-022 to Sonja Ständer).

## Introduction

Chronic pruritus (CP; lasting for six weeks or longer) is a common symptom in dermatology. Estimates are that more than one-third of dermatological patients suffer from CP.<sup>1</sup> CP can arise from various dermatoses, as well as from internal, neurological and mental diseases.<sup>2</sup> It has a high humanistic burden<sup>3</sup> and affects the quality of life negatively.<sup>4</sup> Due to its subjectivity, it is necessary to use a valid patient-reported measuring instrument for recording the intensity of pruritus in order to assess

the course of pruritus and to evaluate the efficacy of the treatment. The most frequently used instruments to assess itch intensity are the numerical rating scale (NRS, range: 0–10), visual analogue scale (VAS, range: 0–10/0–100) and the verbal rating scale (VRS, range: 0–4). These scales were originally developed to assess the intensity of pain and have been validated successfully in chronic pain in several studies.<sup>5,6</sup> Analogous to pain these scales proved to be a valid instrument for the assessment of CP intensity in various pruritic conditions in German- and Polish-speaking patients<sup>7,8</sup> and in a cohort of CP patients in the Republic of Korea,<sup>9</sup> Japan<sup>10</sup> and Sweden. Using the VAS and NRS, the Special Interest Group of the International Forum for the Study of Itch categorized the severity of pruritus according to the following cut-offs: 0: no pruritus, 1–<3: mild pruritus, 3–<7: moderate pruritus, 7–<9: severe pruritus and 9–10: very severe pruritus.<sup>11</sup> In previous validation studies, the assessment of the pruritus intensity via NRS was validated further for atopic dermatitis<sup>12</sup> and psoriasis vulgaris<sup>13,14</sup> in the USA and for chronic nodular prurigo in Germany.<sup>15</sup> Data on the validity of the intensity scales in other pruritic diseases and in different languages are still lacking.

In order to validate and harmonize pruritus assessment tools for both routine care and randomized clinical trials (RCTs) across Europe, the European Network on Assessment of Severity and Burden of Pruritus (PruNet) was founded in 2014.<sup>16</sup> As a result of PruNet's efforts, three pruritus intensity scales (VAS, NRS and VRS) were validated European-wide and in several

**Figure 1** Visual analogue scale (worst) (a), verbal rating scale (worst) (b) and numerical rating scale (worst) (c) within the mobile patient questionnaire application MoPat.

**Table 1** Description of the study population (AD: atopic dermatitis; CD: contact dermatitis; CNPG: chronic nodular prurigo; Pso: psoriasis; LP: lichen planus; CTCL: cutaneous T-cell lymphoma)

		Austria	France	Germany	Italy	Poland	Russia	Switzerland	Spain	Turkey	All
<b>Total, n</b>		60	54	68	71	68	51	53	55	67	547
<b>Gender, n</b>	Female	35	24	40	42	29	34	28	41	40	313
	Male	25	30	28	29	39	17	25	14	27	234
<b>Age in years</b>	Min	18	18	22	25	20	18	18	20	18	18
	Max	83	98	81	80	91	87	90	90	81	98
	Mean	46.65	51.26	54.49	53.04	50.49	50.45	58.28	53.64	46.61	51.56
	SD (+/)	15.86	21.93	15.25	15.53	17.64	18.96	19.93	20.18	16.91	18.14
	Median	47	50.5	56	54	52	50	60	55	48	52
<b>Diagnosis, n (%)</b>	AD	23 (38)	10 (19)	20 (29)	25 (35)	12 (18)	18 (35)	25 (47)	27 (49)	2 (3)	162 (30)
	CD	1 (2)	14 (26)	1 (1)	11 (15)	6 (9)	18 (35)	7 (13)	17 (31)	19 (28)	94 (17)
	CNPG	7 (12)	5 (9)	12 (18)	8 (11)	8 (12)	3 (6)	8 (15)	7 (13)	11 (16)	69 (13)
	Pso	28 (47)	20 (37)	28 (41)	21 (30)	34 (50)	8 (16)	10 (19)	1 (2)	30 (45)	180 (33)
	LP	1 (2)	3 (6)	3 (4)	5 (7)	2 (3)	3 (6)	2 (4)	3 (5)	5 (7)	27 (5)
	CTCL	0 (0)	2 (4)	4 (6)	1 (1)	6 (9)	1 (2)	1 (2)	0 (0)	0 (0)	15 (3)
<b>Duration (n)</b>	<6 weeks	0	4	1	5	7	9	4	3	9	42
	6 weeks to 6 months	0	6	6	14	10	11	6	6	8	67
	6 to 12 months	5	7	2	12	6	4	6	7	7	56
	1 to 10 years	12	19	25	32	23	14	22	24	32	203
	>10 years	43	18	34	8	22	13	15	15	11	179

**Table 2** Descriptive statistics for the intensity scales overall and segregated for the participating sites

		Austria	France	Germany	Italy	Poland	Russia	Switzerland	Spain	Turkey	All
<b>VAS average</b>	Mean	5	5.74	4.29	5.44	5.9	4.24	4.77	5.18	6.6	5.27
	Median	5	6	4	5	6	4	5	5	6	5
	Range	1–10	1–10	0–10	1–10	2–10	1–9	1–9	0–10	2–10	0
	1. Quartile	3.75	4	3	4	4	2	3	3	4.5	3
	3. Quartile	7	7	6	7	8	6.5	7	6	10	7
<b>VAS worst</b>	Mean	5.58	6.43	5.29	6.7	6.96	4.96	6.51	5.87	7.46	6.24
	Median	6	7	5	7	7	5	7	7	8	7
	Range	0–10	0–10	0–10	0–10	2–10	1–9	0–10	0–10	1–10	0
	1. Quartile	3.75	5	3	4.25	5.75	2.5	3.5	5	5	4
	3. Quartile	8	8	7	8.75	9	7	8	9	10	8
<b>NRS average</b>	Mean	5.42	5.54	4.81	6.01	6.37	5.37	5.04	5.73	7	5.73
	Median	6	5	5	6	6	5	5	6	7	6
	Range	1–10	0–9	1–10	3–10	2–10	1–10	0–10	0–10	2–10	0
	1. Quartile	3.75	4	3	4	5	3	4	4	5	4
	3. Quartile	7	7	6	8	8	7	8	6	10	7
<b>NRS worst</b>	Mean	6.15	6.59	5.88	7.06	7.3	6.29	6.81	6.76	7.76	6.75
	Median	6	7.5	6	7	8	7	7.5	7	8.5	7
	Range	1–10	0–10	1–10	1–10	3–10	2–10	0–10	0–10	3–10	0
	1. Quartile	4	5	4.75	6	6	4	5	5	5.25	5
	3. Quartile	8	8	8	9	9	8	9	9	10	9
<b>VRS average</b>	Mean	1.95	2.24	1.82	2.28	2.56	2.06	2.18	1.92	2.65	2.20
	Median	2	2	2	2	2.5	2	2	2	2	2
	Range	1–4	0–4	0–4	0–4	1–4	1–4	0–4	0–4	1–4	0
	1. Quartile	1	2	1	2	2	2	2	1	2	2
	3. Quartile	2.25	3	2	3	3	2	3	2	4	3
<b>VRS worst</b>	Mean	2.23	2.59	2.18	2.62	2.83	2.25	2.56	2.31	2.71	2.49
	Median	2	3	2	3	3	2	3	2	3	2
	Range	1–4	0–4	0–4	0–4	1–4	1–4	0–4	0–4	0–4	0
	1. Quartile	1.75	2	2	2	2	2	2	2	2	2
	3. Quartile	3	3	3	3	3	3	3	3	4	3

pruritic dermatoses. We present and discuss the results of these validation studies in this article.

## Material and methods

### Patients in PruNet

The data collection was part of the non-interventional prospective cohort study performed in the PruNet project. Data were collected in dermatological centres in Austria, France, Germany, Italy, Poland, Russia, Spain, Switzerland and Turkey.

Adult CP patients with atopic dermatitis (AD), contact dermatitis (CD), chronic nodular prurigo (CNP), psoriasis vulgaris (Pso), lichen planus (LP) or cutaneous T-cell lymphoma (CTCL) were included in this study. Patients without sufficient understanding of the local language were excluded.

### Study design

In February 2015, a consensus meeting with 28 pruritus experts from 15 EU countries (21 dermatologists, 5 medical computer

scientists, 2 psychologists) took place to identify the most common itch assessment tools. In a Delphi process, the instruments VAS, NRS und VRS were chosen to be validated European wide.<sup>16</sup> The recall period of these three intensity scales was determined to be 24 h; both the worst pruritus intensity and the average pruritus intensity were recorded with each of the three scales. All pruritus intensity scales were validly translated (forward/backward translation)<sup>17</sup> into the respective languages by a professional translation office. Data collection was performed using the mobile patient survey application MoPat<sup>18</sup> on an iPad, and entered data were transferred into the electronic study database (x4T-EDC) located in Münster, Germany. MoPat is a validated system for the collection of electronic patient-reported outcome.<sup>19</sup> Fig. 1 depicts the pruritus intensity scales within MoPat.

Subjects were recruited by the local principal investigators and associated physicians. A trained medical student supported data collection and administrative processes of each site for a period of three to four weeks. After signing an informed consent,

**Table 3** Descriptive statistics for the intensity scales overall and segregated for the dermatoses (AD: atopic dermatitis; CD: contact dermatitis; CNPG: chronic nodular prurigo; Pso: psoriasis; LP: lichen planus; CTCL: cutaneous T-cell lymphoma)

		AD	CD	CNPG	Pso	LP	CTCL	All
<b>VAS average</b>	Mean	4.93	5.59	6.22	5.14	4.96	4.87	5.27
	Median	5	5	6	5	5	5	5
	Range	0–10	0–10	1–10	0–10	1–10	2–9	0
	1. Quartile	3	3	5	3	3,5	3	3
	3. Quartile	7	8	8	7	6	6	7
<b>VAS worst</b>	Mean	6.05	6.46	7.25	5.89	6.26	6.4	6.24
	Median	7	7	7	6	6	6	7
	Range	0–10	0–10	1–10	0–10	1–10	1–10	0
	1. Quartile	4	4	5.75	4	5	5	4
	3. Quartile	8	9.75	10	8	8	9	8
<b>NRS average</b>	Mean	5.41	6.09	6.54	5.53	5.93	5.33	5.73
	Median	5	6	6	5	6	5	6
	Range	0–10	0–10	3–10	0–10	3–10	2–10	0
	1. Quartile	4	4.25	5	4	4	3,5	4
	3. Quartile	7	8	8	7	7,5	6,5	7
<b>NRS worst</b>	Mean	6.69	7.13	7.65	6.32	6.48	6.47	6.75
	Median	7	7.5	8	7	7	7	7
	Range	0–10	0–10	3–10	0–10	3–10	2–10	0
	1. Quartile	5	5	6	4	5	5	5
	3. Quartile	9	9.75	9	8	8	9,5	9
<b>VRS average</b>	Mean	2.12	2.31	2.43	2.1	2.19	2.47	2.2
	Median	2	2	2	2	2	3	2
	Range	0–4	0–4	1–4	0–4	0–4	1–4	0
	1. Quartile	2	2	2	2	2	2	2
	3. Quartile	3	3	3	3	3	3	3
<b>VRS worst</b>	Mean	2.52	2.5	2.7	2.36	2.56	2.53	2.49
	Median	2	2	3	2	2	3	2
	Range	0–4	0–4	1–4	0–4	1–4	1–4	0
	1. Quartile	2	2	2	2	2	2	2
	3. Quartile	3	3	3	3	3	3	3

subjects were registered to the study database and basic demographic data and disease-related parameters were entered.<sup>20</sup> Afterwards, the subject filled in following patient-reported outcome tools: for testing test–retest reliability, all subjects recorded their average and worst pruritus intensity in the previous 24 h by completing the NRS, VAS and VRS twice after 30–60 min. Additionally, subjects filled in the Dermatology Life Quality Index (DLQI).<sup>21</sup>

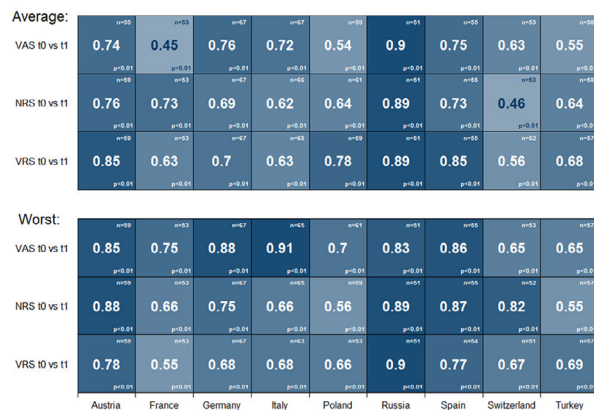
The study was approved by the leading ethic committee in Münster (2015-171-f-S) and by each ethic committee of the participating sites. The study was registered at the German Clinical Trial Register (DRKS00007958).

### Statistical analysis

The statistical analysis was performed using R (version 3.6.0).<sup>22</sup> Subgroup analysis was performed for each pruritic condition (AD, CD, CNPG, Pso, LP and CTCL) and recruiting centre (Austria, France, Germany, Italy, Poland, Russia,

Spain, Switzerland and Turkey). For each pruritus intensity scales (VAS, NRS and VRS), we evaluated (1) re-test reliability, (2) correlation among the intensity scales and (3) correlation with the impairment of quality of life measured by the DLQI.

For ordinal variables, we used the weighted Kappa statistic  $\kappa_w$ , as a meaningful measure of agreement. The statistic measures the proportion of weighted agreement corrected for chance agreement.<sup>23</sup> We evaluated re-test reliability using (equal-spacing) weighted Cohen's Kappa  $\kappa_w$  and confusion-matrix plots with weights according to the fraction of column and row population size for ordinal data.<sup>23</sup> Cohen's Kappa  $\kappa_w$  coefficients above 0.8 are considered as almost perfect, and between 0.61 and 0.8 are considered as substantial.<sup>24</sup> Correlation analyses across pruritus intensity scales and between these scales and DLQI were performed using Spearman's rank correlation coefficient.<sup>25</sup> All reported p values were corrected for multiple testing applying Bonferroni correction.



**Figure 2** Correlation between the repeatedly stated intensity scales at two points of time t0 and t1 segregated for the participating sites using the weighted Kappa  $\kappa_W$  (<0 Less than chance agreement, 0.01–0.20 Slight agreement, 0.21–0.40 Fair agreement, 0.41–0.60 Moderate agreement, 0.61–0.80 Substantial agreement, 0.81–0.99 Almost perfect agreement). The number of cases is shown in the upper right corner, and the adjusted p-value is shown in the lower right corner for the particular correlation.

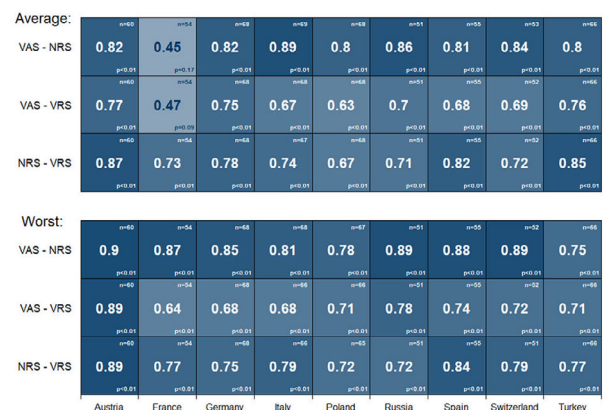


**Figure 3** Correlation between the repeatedly stated intensity scales at two points of time t0 and t1 segregated for the (AD: atopic dermatitis; CD: contact dermatitis; CNPG: chronic nodular prurigo; Pso: psoriasis; LP: lichen planus; CTCL: cutaneous T-cell lymphoma) using the weighted Kappa  $\kappa_W$  (<0 Less than chance agreement, 0.01–0.20 Slight agreement, 0.21–0.40 Fair agreement, 0.41–0.60 Moderate agreement, 0.61–0.80 Substantial agreement, 0.81–0.99 Almost perfect agreement). The number of cases is shown in the upper right corner, and the adjusted p-value is shown in the lower right corner for the particular correlation.

## Results

### Study population

From August 2015 until June 2016, a total of 547 (57.2% female) subjects were enrolled (Table 1). The participants were aged between 18 and 98 (mean 51.6, SD  $\pm$  18.1, median 52.0) years.



**Figure 4** Correlation between the different intensity scales segregated for the participating sites using the Spearman correlation coefficient (0.00–0.30 negligible correlation, 0.30–0.50 Low correlation, 0.50–0.70 Moderate correlation, 0.70–0.90 High correlation, 0.90–1.00 Very high correlation). The number of cases is shown in the upper right corner, and the adjusted p-value is shown in the lower right corner for the particular correlation.



**Figure 5** Correlation between the different intensity scales segregated for the dermatoses (AD: atopic dermatitis; CD: contact dermatitis; CNPG: chronic nodular prurigo; Pso: psoriasis; LP: Lichen planus; CTCL: cutaneous T-cell lymphoma) using the Spearman correlation coefficient (0.00–0.30 negligible correlation, 0.30–0.50 Low correlation, 0.50–0.70 Moderate correlation, 0.70–0.90 High correlation, 0.90–1.00 Very high correlation). The number of cases is shown in the upper right corner and the adjusted P-value is shown in the lower right corner for the particular correlation.

The two most frequent dermatoses were psoriasis vulgaris (33.0%) and atopic dermatitis (30.0%). Most of the participants had been suffering from CP for one to ten years (37.1%) or even more than ten years (32.7%). There were no significant



DLQI									
Average:									
VAS	0.21 n=59 p=1	0.17 n=54 p=1	0.45 n=67 p=0.04	0.53 n=61 p=0.01	0.4 n=67 p=0.13	0.57 n=61 p=0.01	0.29 n=55 p=1	0.35 n=63 p=1	0.31 n=63 p=1
NRS	0.42 n=59 p=0.24	0.28 n=54 p=1	0.51 n=67 p=0.01	0.52 n=61 p=0.01	0.37 n=67 p=0.04	0.62 n=61 p=0.01	0.39 n=55 p=0.88	0.4 n=63 p=0.75	0.36 n=63 p=1
VRS	0.42 n=59 p=0.21	0.3 n=54 p=1	0.45 n=67 p=0.04	0.48 n=61 p=0.01	0.51 n=67 p=0.01	0.55 n=61 p=0.01	0.42 n=55 p=0.34	0.41 n=63 p=0.61	0.44 n=63 p=0.1
Worst:									
VAS	0.42 n=59 p=0.21	0.04 n=54 p=1	0.48 n=67 p=0.01	0.52 n=61 p=0.01	0.36 n=67 p=0.01	0.56 n=61 p=0.01	0.29 n=55 p=1	0.42 n=63 p=0.53	0.22 n=63 p=1
NRS	0.46 n=59 p=0.07	0.09 n=54 p=1	0.54 n=67 p=0.01	0.54 n=61 p=0.01	0.3 n=67 p=0.01	0.53 n=61 p=0.01	0.31 n=55 p=1	0.41 n=63 p=0.77	0.26 n=63 p=1
VRS	0.46 n=59 p=0.06	0.22 n=54 p=1	0.48 n=67 p=0.01	0.46 n=61 p=0.01	0.42 n=67 p=0.01	0.49 n=61 p=0.01	0.29 n=55 p=1	0.29 n=63 p=0.62	0.44 n=63 p=0.01
Austria France Germany Italy Poland Russia Spain Switzerland Turkey									

**Figure 6** Correlation between the different intensity scales and the DLQI segregated for the participating sites using the Spearman correlation coefficient (0.00–0.30 negligible correlation, 0.30–0.50 Low correlation, 0.50–0.70 Moderate correlation, 0.70–0.90 High correlation, 0.90–1.00 Very high correlation). The number of cases is shown in the upper right corner, and the adjusted *P*-value is shown in the lower right corner for the particular correlation.

DLQI						
Average:						
VAS	0.46 n=162 p=0.01	0.32 n=84 p=0.46	0.12 n=69 p=1	0.38 n=180 p=0.01	0.26 n=27 p=1	0.78 n=15 p=0.01
NRS	0.5 n=161 p=0.01	0.47 n=84 p=0.01	0.14 n=68 p=1	0.4 n=179 p=0.01	0.32 n=27 p=1	0.82 n=15 p=0.01
VRS	0.46 n=161 p=0.01	0.44 n=84 p=0.01	0.22 n=68 p=1	0.44 n=179 p=0.01	0.54 n=27 p=1	0.74 n=15 p=0.01
Worst:						
VAS	0.45 n=162 p=0.01	0.32 n=84 p=0.01	0.18 n=68 p=1	0.3 n=180 p=0.01	0.46 n=27 p=1	0.66 n=15 p=0.01
NRS	0.43 n=161 p=0.01	0.4 n=84 p=0.01	0.17 n=68 p=1	0.32 n=179 p=0.01	0.4 n=27 p=1	0.83 n=15 p=0.01
VRS	0.36 n=161 p=0.01	0.48 n=84 p=0.01	0.19 n=68 p=1	0.43 n=179 p=0.01	0.45 n=27 p=1	0.59 n=15 p=0.01
AD CD CNPG Pso LP CTCL						

**Figure 7** Correlation between the different intensity scales and the DLQI segregated for the dermatoses (AD: atopic dermatitis; CD: contact dermatitis; CNPG: chronic nodular prurigo; Pso: psoriasis; LP: lichen planus; CTCL: cutaneous T-cell lymphoma) using the Spearman correlation coefficient (0.00–0.30 negligible correlation, 0.30–0.50 Low correlation, 0.50–0.70 Moderate correlation, 0.70–0.90 High correlation, 0.90–1.00 Very high correlation). The number of cases is shown in the upper right corner, and the adjusted *p*-value is shown in the lower right corner for the particular correlation.

differences in itch intensity between males and females in any of the six measures, and there was no significant correlation between itch intensity and age and gender (Figure S2).

### Pruritus intensity scales

For all pruritus intensity scales, observed scores ranged from their minimum, 0 to their maximum, 4 (VRS) or 10 (VAS, NRS)

(Tables 2 and 3). The average and worst intensity measured by VAS were highest in Turkey with a mean of 6.6 and 7.46 and a median of 6 and 8, respectively. Accordingly, the NRS was rated in Turkey with a mean and median of 7 for average and with 7.76 and 8.50 for the worst intensity. While in Poland the VAS and NRS were slightly lower, the NRS was similar to Turkey. The lowest intensities were found in Germany and Russia. While the reported VAS and VRS were quite similar in Germany and Russia, the NRS was slightly lower in Germany (average mean: 4.81 to 5.37 and worst mean: 5.88 to 6.29). Regarding the considered diagnoses, the highest intensities of pruritus were found in the CNPG group. The mean of the VAS average and VAS worst was 6.22 and 7.25 and for the NRS 6.54 and 7.65. In contrast, the lowest average intensities were measured for the atopic dermatitis and cutaneous T-cell lymphoma, while the lowest worst intensities were found in psoriasis vulgaris.

### Reproducibility

Testing for re-test reliability, mostly substantial agreement was found ( $0.61 \leq \kappa_w \leq 0.8$ ), while 8 correlation coefficients had only moderate agreement ( $0.41 \leq \kappa_w \leq 0.6$ ) (Figs 2 and 3). The best agreement was observed in Russia with consistently almost perfect agreement ( $0.81 \leq \kappa_w \leq 0.99$ ). For some intensity scales, only moderate agreement ( $0.41 \leq \kappa_w \leq 0.6$ ) was obtained, for example for the VAS average in France. The correlation segregated by dermatosis was even more consistent. Only two correlations were slightly below (VAS average for cutaneous T-cell lymphoma) or above (VAS worst for atopic dermatitis) the substantial agreement, while all other correlations had substantial agreement.

### Intraclass Correlation between the Intensity Scales

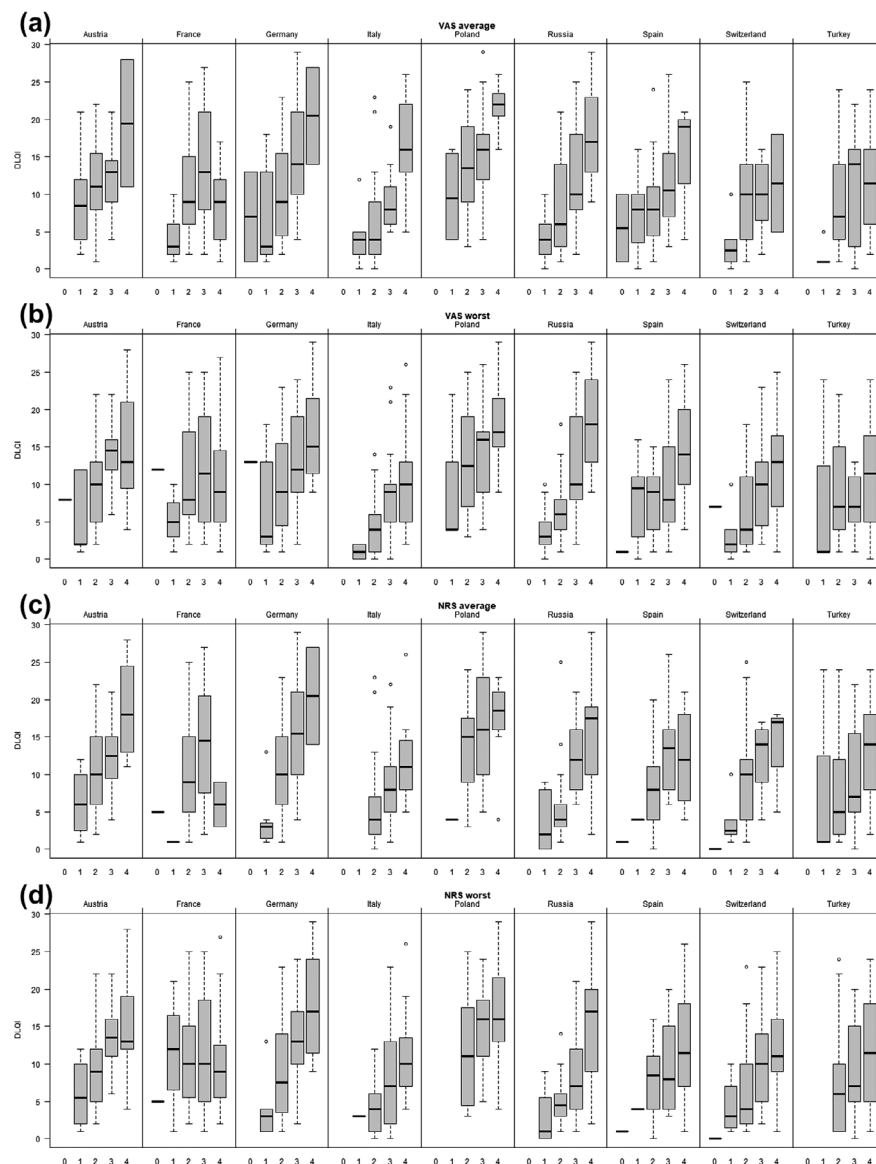
The intraclass correlation was high for most scales and sites (Fig. 4). One exception was recorded in France, in which the correlation between the VAS and the NRS and VRS was low. Considering the pruritus intensity scales for each dermatosis (Fig. 5), the intraclass correlation coefficient was also high, while in 13.89% of the cases a moderate correlation was shown.

### Convergent validity

The correlation of the pruritus intensity scales with the DLQI was low. The lowest correlation was measured at the French study site (0.04–0.30), while the highest correlation was attained in Russia (0.49–0.62) (Fig. 6). With regard to the dermatoses (Fig. 7), the correlations were very similar. Mostly, the analysis resulted in a low correlation, but the lowest correlation was found for CNPG (0.12 up to 0.22) and the highest for CTCL (0.58–0.83).

### Concurrent validity

An overview of the correlation analyses between the VAS and NRS categories (no pruritus, mild, moderate, severe, very severe



**Figure 8** Visualization of the correlation between the different grouped (0  $\triangleq$  No pruritus; 1  $\triangleq$  Mild pruritus; 2  $\triangleq$  Moderate pruritus; 3  $\triangleq$  Severe pruritus; 4  $\triangleq$  Very severe pruritus) intensity scales ((a) VAS average; (b) VAS worst; (c) NRS average; (d) NRS worst) and the DLQI segregated for the dermatoses (AD: atopic dermatitis; CD: contact dermatitis; CNPG: chronic nodular prurigo; Pso: psoriasis; LP: lichen planus; CTCL: cutaneous T-cell lymphoma).

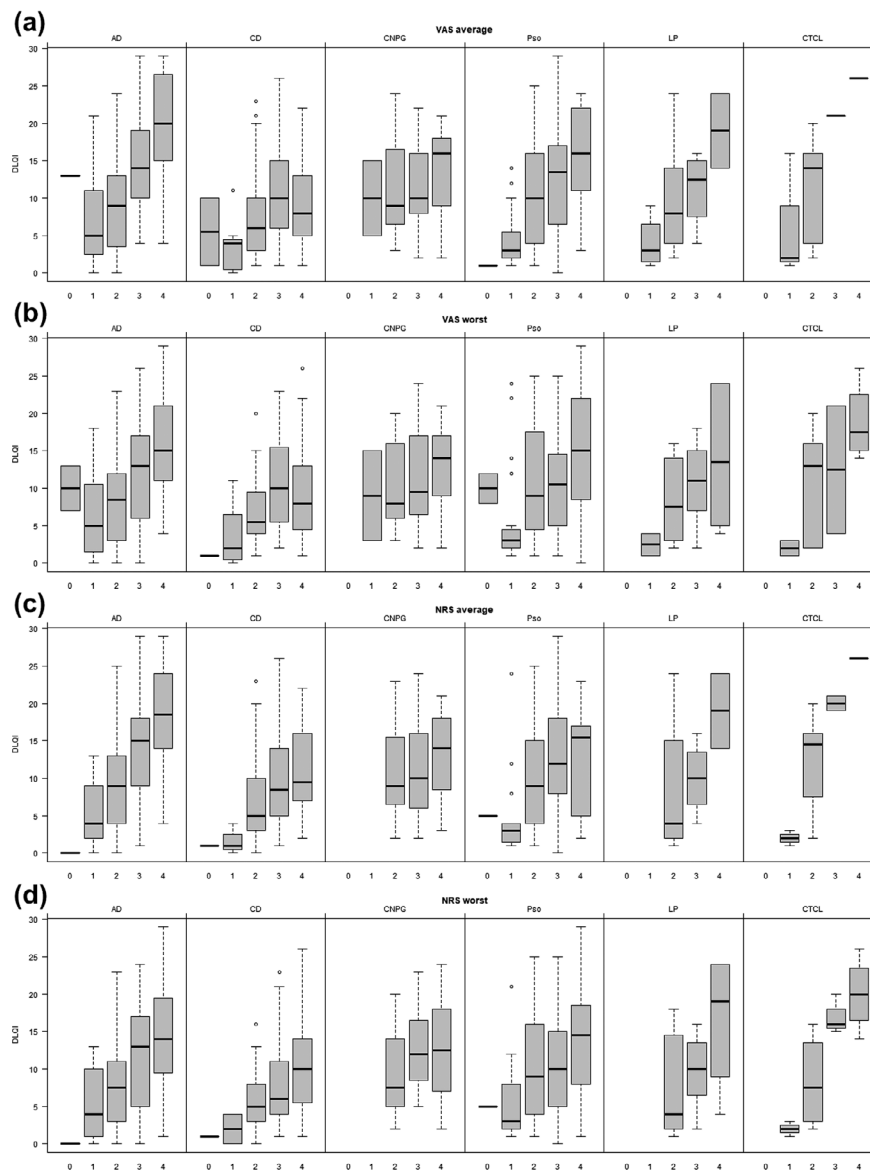
pruritus) and the DLQI can be seen in Fig. 8 (segregated by country) and Fig. 9 (segregated by dermatosis). Generally, analysis in patients with no, mild, moderate, severe and very severe pruritus for VAS and NRS showed increasing DLQI scores with stronger pruritus intensity.

## Discussion

In this study, the most commonly used patient-reported instruments for measuring pruritus intensity were validated in

different languages and in different dermatoses.<sup>26</sup> With a few exceptions, the reproducibility showed a substantial agreement in all countries. A moderate agreement between the scores of the two measurement time points were found in France VAS average and VRS worst, in Poland VAS average, in Switzerland NRS average and VRS average, and in Turkey VAS average and NRS worst. VAS worst showed a good result for the reproducibility in all countries. The intraclass correlation coefficients indicated with at least moderate correlation between the measures – with two





**Figure 9** Visualization of the correlation between the different grouped (0  $\hat{=}$  No pruritus; 1  $\hat{=}$  Mild pruritus; 2  $\hat{=}$  Moderate pruritus; 3  $\hat{=}$  Severe pruritus; 4  $\hat{=}$  Very severe pruritus) intensity scales ((a) VAS average; (b) VAS worst; (c) NRS average; (d) NRS worst) and the DLQI segregated for the dermatoses (AD: atopic dermatitis; CD: contact dermatitis; CNPG: chronic nodular prurigo; Pso: psoriasis; LP: lichen planus; CTCL: cutaneous T-cell lymphoma).

exceptions: in France, VAS average correlated only weakly with both NRS and VRS.

As a second objective, the validity of the intensity scales was tested in various pruritic dermatoses. Here, test–retest showed a substantial agreement and an at least moderate correlation between the scales showing a good temporal stability and inter-item correlation.

Intensity of pruritus correlated significantly with QoL in Italy, Poland, Switzerland, Turkey, Spain and with one exception in

Germany and Russia and also in pruritic dermatosis like atopic dermatitis, contact dermatitis, psoriasis vulgaris and cutaneous T-cell lymphoma. The low-to-moderate correlation between the intensity scales and the DLQI might be a sign of a low validity; otherwise, it can be assumed that the impairment of QoL with subdimensions of function, emotion and symptoms is more independent of intensity of pruritus. The highest (moderate) correlation to the DLQI was found in Germany, Italy, Poland and Russia. In France, there was only a significant correlation

with NRS average und VRS average. An explanation for the small proportion of statistically significant correlations between the intensity scales and the DLQI could be that in France more patients were male and the population was the third youngest in this investigation. It could be assumed that younger men rate their impairment of QoL lower even though pruritus intensity is high. This is confirmed by the observation that males under 65 years of age with CP achieve lower DLQI scores, both as compared to females of the same age and to males over 65 years of age.<sup>27</sup>

Regarding dermatoses, there was a high correlation of the DLQI Score with the intensity of pruritus in cutaneous T-cell lymphoma, which suggests that – if present – pruritus has a great influence on the QoL in this dermatosis,<sup>28</sup> whereas there was a low correlation in the other diseases. It is particularly striking here that the high pruritus intensity of CNPG showed no significant influence on QoL measured by the DLQI. Previous research has shown that patients with CNPG have a severe impairment in quality of life<sup>29</sup>; here, in addition to the intensity of pruritus, other factors like the frequency of pruritus and the presence of pruriginous lesions may appear to play also a negative role in influencing QoL.

The concurrent validity showed, with the exception of France and patients with CNPG, a higher impairment of QoL with increasing severity of CP. In France, there was no further increase in the DLQI in the group of very severe pruritus (NRS, VAS = 9–10).

The median pruritus intensity measured by the VAS and NRS values were 5–6. The highest scores were found in Turkey, the lowest in Russia and Germany. Measuring the itch intensity by the VRS, the median of the VRS was 2; patients in Italy, Poland, Turkey and Spain scored slightly higher. Neither age, gender, diagnosis nor duration of pruritus in these countries is particularly different from the others. One can assume that this could be a culture effect since the humanistic burden in patients with CP in inflammatory dermatoses is also strongly influenced by cross-cultural factors.<sup>3</sup> However, we did not correct for ongoing therapy, what might be an additional factor.

In the dermatoses, the highest pruritus intensity was found for CNPG in almost all intensity scales. This confirms other studies, in which the intensity of CNPG was higher than in other dermatoses.<sup>30</sup> Peripheral sensitization processes have been demonstrated in CPG.<sup>31</sup> These can lead to chronicity of the symptom, which results not only in the form of allodynia but also in the presence of constant pruritus with a high intensity.<sup>29</sup>

In summary, the itch intensity scales had good intraclass correlations, fulfilled the validity criteria ‘reproducibility’ and were independent from age and gender. The ‘convergent’ and ‘concurrent validity’ were good in Italy, Poland, Switzerland, Turkey, Spain and with one exception in Germany and Russia and in atopic dermatitis, contact dermatitis, psoriasis vulgaris, lichen

planus and cutaneous T-cell lymphoma. Country-specific, the VAS worst was the best reproducible and consistent measuring instrument in all countries. The NRS worst also showed a good correlation to the other intensity scales, but only a moderate agreement between the test–retest scores in Poland and Turkey. According to this analyse, the VAS worst might be superior to the NRS worst. However, in previous work VAS worst showed more missing values than the NRS worst.<sup>7</sup> Based on this investigation, the sensitivity to change cannot be illustrated. One can assume that the VAS worst and the NRS worst do not improve significantly and very fast under treatment, although it is effective. A suitable instrument for detecting an antipruritic effect would be one that measures the intensity, but also the duration and frequency of pruritus in the last 24h, as itch itself is known as a fluctuating symptom.

### Limitations

In this investigation, we tested the validity of pruritus intensity scales in the mentioned dermatoses and chronic prurigo at one day. The test for responsiveness is missing. Future studies should also test the validity in chronic pruritus in non-lesional skin like neuropathic, systemic or psychogenic itch.

### Acknowledgement

We thank all members in the PruNet group especially Anna Post, Juliane Mozar, Tanja Winter, Friederike Elisabeth Vietmeier and Julia Funke for helping collecting the data.

### References

- 1 Kopyciok MER, Ständer HF, Osada N *et al*. Prevalence and characteristics of pruritus: a one-week cross-sectional study in a German dermatology practice. *Acta Derm Venereol* 2016; **96**: 50–55.
- 2 Ständer S. Classification of Itch. *Curr Probl Dermatol* 2016; **50**: 1–4.
- 3 Steinke S, Zeidler C, Riepe C *et al*. Humanistic burden of chronic pruritus in patients with inflammatory dermatoses: Results of the European Academy of Dermatology and Venereology Network on Assessment of Severity and Burden of Pruritus (PruNet) cross-sectional trial. *J Am Acad Dermatol* 2018; **79**: 457–463.e5.
- 4 Leader B, Carr CW, Chen SC. Pruritus epidemiology and quality of life. *Handb Exp Pharmacol* 2015; **226**: 15–38.
- 5 Ferreira-Valente MA, Pais-Ribeiro JL, Jensen MP. Validity of four pain intensity rating scales. *Pain* 2011; **152**: 2399–2404.
- 6 Ferraz MB, Quaresma MR, Aquino LR *et al*. Reliability of pain scales in the assessment of literate and illiterate patients with rheumatoid arthritis. *J Rheumatol* 1990; **17**: 1022–1024.
- 7 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.
- 8 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.
- 9 Jang YH, Kim SM, Eun DH *et al*. Validity and reliability of itch assessment scales for chronic pruritus in adults: A prospective multicenter study. *J Am Acad Dermatol* 2020; **82**: 80–86.
- 10 Kido-Nakahara M, Katoh N, Saeki H *et al*. Comparative cut-off value setting of pruritus intensity in visual analogue scale and verbal rating scale. *Acta Derm Venereol* 2015; **95**: 345–346.

- 11 Reich A, Chatzigeorgidis E, Zeidler C *et al*. Tailoring the cut-off values of the visual analogue scale and numeric rating scale in itch assessment. *Acta Derm Venereol* 2017; **97**: 759–760.
- 12 Yosipovitch G, Reaney M, Mastey V *et al*. Peak Pruritus Numerical Rating Scale: psychometric validation and responder definition for assessing itch in moderate-to-severe atopic dermatitis. *Br J Dermatol* 2019; **181**: 761–769.
- 13 Naegeli AN, Flood E, Tucker J *et al*. The Worst Itch Numeric Rating Scale for patients with moderate to severe plaque psoriasis or psoriatic arthritis. *Int J Dermatol* 2015; **54**: 715–722.
- 14 Kimball AB, Naegeli AN, Edson-Heredia E *et al*. Psychometric properties of the Itch Numeric Rating Scale in patients with moderate-to-severe plaque psoriasis. *Br J Dermatol* 2016; **175**: 157–162.
- 15 Kimel M, Zeidler C, Kwon P *et al*. Validation of psychometric properties of the itch numeric rating scale for pruritus associated with Prurigo Nodularis: a secondary analysis of a randomized clinical trial. *JAMA Dermatol* 2020; **156**: 1354.
- 16 Ständer S, Zeidler C, Riepe C *et al*. European EADV network on assessment of severity and burden of Pruritus (PruNet): first meeting on outcome tools. *J Eur Acad Dermatol Venereol* 2016; **30**: 1144–1147.
- 17 Ofenloch RF, Diepgen TL, Popielnicki A *et al*. Severity and functional disability of patients with occupational contact dermatitis: validation of the German version of the Occupational Contact Dermatitis Disease Severity Index. *Contact Dermatitis* 2015; **72**: 84–89.
- 18 Soto-Rey I, Dugas M, Storck M. Implementation of an ODM and HL7 compliant electronic patient-reported outcome system. *Stud Health Technol Inform* 2016; **228**: 421–425.
- 19 Storck M, Zeidler C, Rehr M *et al*. Validation of pruritus measures gathered with the electronic patient-reported outcome system MoPat. *Acta Derm Venereol* 2018; **98**: 38–43.
- 20 Zeidler C, Steinke S, Riepe C *et al*. Cross-European validation of the Itch-yQoL in pruritic dermatoses. *J Eur Acad Dermatol Venereol* 2019; **33**: 391–397.
- 21 Basra MKA, Fenech R, Gatt RM *et al*. The Dermatology Life Quality Index 1994–2007: a comprehensive review of validation data and clinical results. *Br J Dermatol* 2008; **159**: 997–1035.
- 22 R Core Team. R. A Language and Environment for Statistical Computing, 2019: <https://www.R-project.org/>.
- 23 Cohen J. A Coefficient of agreement for nominal scales. *Educ Psychol Measure* 1960; **20**: 37–46.
- 24 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; **33**: 159.
- 25 Mukaka MM. A guide to appropriate use of Correlation coefficient in medical research. *Malawi Med J* 2012; **24**: 69–71.
- 26 Der Titel "" kann nicht dargestellt werden. Die Vorlage "Literaturverzeichnis - Internetdokument - (Standardvorlage)" beinhaltet nur Felder, welche bei diesem Titel leer sind.
- 27 Warlich B, Fritz F, Osada N *et al*. Health-related quality of life in chronic pruritus: an analysis related to disease etiology, clinical skin conditions and itch intensity. *Dermatology* 2015; **231**: 253–259.
- 28 Holahan HM, Farah RS, Fitz S *et al*. Health-related quality of life in patients with cutaneous T-cell lymphoma? *Int J Dermatol* 2018; **57**: 1314–1319.
- 29 Pereira MP, Hoffmann V, Weisshaar E *et al*. Chronic nodular prurigo: clinical profile and burden. A European cross-sectional study. *J Eur Acad Dermatol Venereol* 2020; **34**: 2373–2383.
- 30 Schut C, Dalgard FJ, Halvorsen JA *et al*. Occurrence, chronicity and intensity of itch in a clinical consecutive sample of patients with skin diseases: a multi-centre study in 13 European Countries. *Acta Derm Venereol* 2019; **99**: 146–151.
- 31 Pogatzki-Zahn EM, Pereira MP, Cremer A *et al*. Peripheral sensitization and loss of descending inhibition is a hallmark of chronic pruritus. *J Invest Dermatol* 2020; **140**: 203–211.e4.

## Supporting information

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

**Figure S1.** Distribution of the different intensity scales grouped by gender (F: female; M: male).

**Figure S2.** Correlation of the different intensity scales with the age of the participants.