

ItchApp[®]: An App-based eDiary for Assessment of Chronic Pruritus in Clinical Trials

Marvin GERNART¹, Athanasios TSIANAKAS¹, Claudia ZEIDLER¹, Claudia RIEPE¹, Nani OSADA¹, David PIHAN² and Sonja STÄNDER¹

¹Center for Chronic Pruritus, Department of Dermatology, University Hospital Münster, Germany, and ²Arone, Saint-Maur des Fosses, France

Performing a reliable assessment of chronic pruritus remains a challenge. Electronic diaries are often used, but many of the scales have not been validated. ItchApp[®] was developed for Android smartphones in order to address this lack. A total of 40 subjects with chronic pruritus completed questionnaires both on paper and with ItchApp[®] (verbal rating scale, numerical rating scale, dynamic pruritus score) in order to validate the software application. Strong correlations were found for test-retest reliability (intraclass correlation coefficient: 0.865–0.977) and convergent validity (Spearman’s r: 0.442–0.924). A feasibility questionnaire for ItchApp[®] revealed a high level of user friendliness and compliance. This was confirmed in a randomized controlled trial with 68 subjects, for which the clinically important difference in the numerical rating scale values for ItchApp[®] was calculated (2.61 points). In summary, ItchApp[®] is a recently developed eDiary that can provide experts with a reliable evaluation of patients with chronic pruritus. It will be made available for future clinical trials.

Key words: itch; visual analogue scale; numerical rating scale; minimal clinically important difference; psychometric.

Accepted Dec 15, 2016; Epub ahead of print Dec 16, 2016

Acta Derm Venereol 2017; 97: 601–606.

Corr: Sonja Ständer, Center of Chronic Pruritus, Department of Dermatology, University of Münster, Von-Esmarch-Str. 58, DE-48149 Münster, Germany. E-mail: sonja.staender@uni-muenster.de

Chronic pruritus (CP) is a frequent symptom with a high impact on patients’ quality of life (1–3). To date, there is almost no specific therapy available that successfully alleviates itching (4). New therapies are thus urgently needed. Many novel mechanisms of pruritus have been identified recently, allowing experts to establish therapy targets. Although many randomized controlled trials (RCT) have been initiated (5), it is difficult to assess pruritus because it is a subjective symptom (6). The study end-points, usually the determination of the course of the itch and quality of life, can be assessed by validated patient-reported outcomes (PRO). Several PRO questionnaires have been developed to better assist in evaluating CP in recent years, including the visual analogue scale (VAS), numerical rating scale (NRS) (7, 8), ItchyQoL (9), 5-D Itch Scale (10) and the pruritus-specific patient benefit index (11). The most common way to assess the course of pruritus is the NRS and VAS,

which serve in many RCTs as the primary endpoint (6). In such assessment using paper-based diaries, certain criteria cannot be regulated and properly documented, including the exact time at which a subject completed certain questions or the entirety of the questionnaire. Such hindrances have often proven to be great disadvantages in RCTs. A further issue is the quality of data in paper-based questionnaires. Subjects may forget to use a physical diary and attempt to complete it quickly prior to their follow-up visit. As a consequence, this retrospective data may not be entirely reliable, and the prospective course of the pruritus is thus difficult to estimate accurately. In order to address these issues, electronic diaries (eDiaries) have been adapted for use in RCTs. Several providers (e.g. PHT Corporation, Boston, USA; ERT, Philadelphia, USA) offer eDiaries with endpoints identified by sponsors. There are, however, various issues attributed to the use of eDiaries. The devices provided to subjects are often bulky and the questions may be non-validated or have been assessed with non-validated methods.

In order to address these issues we designed a novel software application (app) for use in clinical trials that is easy to use on Android smartphones (Appendix S1¹). The app contains multiple questions validated by experts (members of the Special Interest Group “Scoring Itch in Clinical Trials” of the IFSI) regarding the course of patients’ itch, which is intended for use on a daily basis. ItchApp[®] was developed in compliance with validation guidelines and security regulations (e.g. audit trails, electronic signatures, data encryption, etc.). A specific feature restricts subjects’ ability to answer questions retrospectively, thus ensuring the quality of the data entered. This assures subjects, physicians and researchers of accurate information. The aim of this study was to validate ItchApp[®] as a novel instrument.

METHODS

Validation study

Subjects (≥ 18 years old) with CP with a score of at least 2 points on the NRS were asked to use ItchApp[®] twice within 1 h (Fig. S1¹). This time period was selected according to a previous study (7) and the notion that it is acceptable in terms of a balance between the natural daily variabilities in itch, the minimized recall bias due to the number of questions, and distraction due to the

¹<https://doi.org/10.2340/00015555-2602>

doctor's visit between the first and second assessment. Subjects answered routine paper-based questionnaires only once prior to using ItchApp[®]. Immediately afterwards, the use of the app was explained to the subject, including the use of the demo tool (allowing the user to answer sample questions), and the patients then directly completed ItchApp[®]. To ensure data security and quality, the data were transferred immediately to the central database after inputting the subject's password. The database contains the date of data entry, with the exact times at which the entry was begun and completed. HTC One SV smartphones (4.3-inch, 480 × 800 pixel resolution, Wide Video Graphics Array (VGA)), Android 4.0.4 OS) were utilized for the study. After completing the app, subjects completed a feasibility questionnaire comprising 10 questions related to the use, simplicity of the content and convenience of ItchApp[®] for assessing their pruritus. Before any specific study procedures were initiated, all subjects gave written informed consent. The Münster University Ethics Committee approved this validation study (No. 2014-385-f-S), which was registered in the German Clinical Trials Register (DRKS00009902). The study was performed in accordance with the principles of the Declaration of Helsinki.

Use of ItchApp[®] in a clinical trial

ItchApp[®] on HTC One SV smartphones was distributed to 69 subjects in a randomized, double-blind, placebo-controlled trial of the neurokinin-1 receptor antagonist VLY-686 sponsored by Vanda Pharmaceuticals (protocol number VP-VLY-686-2101; ClinicalTrials.gov No. NCT02004041). To investigate the reliability of ItchApp[®], the data transfer (the date and time of expected and real data transfer) and subject compliance (data consistency, duration of data entry, number of recall entries) information were extracted from the central database and analysed. Fifty-three of the 69 subjects also completed a further feasibility questionnaire similar to that used in the validation study. The mean NRS and first band (slight improvement) of the Dynamic Pruritus Score (DPS) data were taken from the VP-VLY-686-2101 dataset and used to calculate the minimal clinically important difference (MCID).

Study outcomes

All questionnaires were provided in German. The paper-based itch intensity scales asked subjects to record their mean and worst itch on the VAS (10-cm line) within the past 4 weeks. In addition, the mean VAS, the mean verbal rating scale (VRS, range from 0: no pruritus to 4: very severe pruritus) and the mean NRS (full numbers from 0–10) within the past 24 h were assessed. The DPS (percentage change in pruritus intensity since the beginning of treatment: –100% to 0, 0 to 100%) was obtained (12). Questionnaires surveying the subjects' quality of life, such as the German language ItchyQoL (GerItchyQoL; 22 questions regarding the past week) and Dermatology Life Quality Index (DLQI; 10 questions considering the last 7 days) were also answered in their paper version. Subjects were asked to use ItchApp[®] to answer questions regarding the following: (i) VRS mean within the last 24 h (range 0–4), (ii) NRS mean and worst in the last 24 h (range 0–10), (iii) Global Pruritus Rating: Itching in the last 24 h? (yes or no), (iv) DPS – Score I: Change in itching since the beginning of the study (yes or no). Score II: If yes, an improvement or worsening? Score III (score bands): If improved: 1–30% (slight improvement); 31–50% (moderate improvement); 51–70% (good improvement); 71–100% (very good improvement).

Statistical analysis

All ItchApp[®] entries were extracted from Arone's database and imported into Microsoft Excel. This data was then imported into

SPSS and integrated with the data from the paper-based questionnaires that had already been entered manually into a separate SPSS document. All items were tested for correlations regarding test–retest reliability, convergent and inter-item validity. Cohen's kappa was used to determine the repeatability with 2 nominal variables. We then tested for Spearman's rank correlation coefficient and the intraclass correlation coefficient (ICC). The Mann–Whitney *U* test was used to determine the 2 independent subgroups of ItchApp[®] global scores (yes vs. no) and continuous or ordinal variables. Wilcoxon signed-rank test was used to compare the baseline NRS mean with the mean NRS of the last time a slight improvement was noticed by the patients. Thus we calculated the *p*-value of the MCID. A χ^2 test was used to determine the significance of subjects' age and its impact on their use of a smartphone app. Pearson's product-moment correlation coefficient was used to estimate the strength of the relationship between the time needed to complete the questionnaires and the age of the patients. The Mann–Whitney *U* test was used to survey the assessment times of male and female subjects. Statistical analyses were performed with IBM SPSS Statistics for Windows, Version 22.0 (Armonk, NY: IBM Corp.). The statistical significance was found to be $p \leq 0.05$ (2-sided).

RESULTS

Validation study

Forty subjects with CP were recruited for the validation study during their visit to the Center for Chronic Pruritus, University Hospital Münster, Germany, for either inpatient ($n=14$) or outpatient ($n=26$) treatment. The difference in itch intensities between the underlying disease categories was not significant. No correlation was found between itch intensity and duration of CP (Spearman's $r=0.04$; $p=0.78$). The results of all parameters, including those from ItchApp[®] data, are shown in **Table I**.

Test–retest reliability. Excellent and very strong correlations were found for the itch intensity scales in the test–retest analysis regarding the use of ItchApp[®] at 2 different time points (T1 and T2) (**Table II**). Cohen's kappa was calculated with the Global Pruritus Rating and the DPS derived from ItchApp[®], for which the DPS Score I was found to be of significance (yes/no; 0.894; $p < 0.001$). The significance was found to be very high in all other tests conducted, with $p < 0.001$ (Table II).

Inter-item validity. Items from ItchApp[®] were correlated among each another (**Table III**). Except for data from one subject, all data taken from the Global Pruritus Rating (itch, yes or no?) and intensity scales (rated as 0 or >0) was consistent. Inter-item correlation tests proved a strong correlation between the VRS and NRS scores ($p < 0.001$; Spearman's correlation). As expected, the strongest coefficients were identified among the mean NRS and worst NRS scores (Spearman's correlation, $r=0.80$).

Convergent validity. The itch intensities recorded by ItchApp[®] were tested against those documented by the paper-based questionnaires (Table III). Correlation analyses confirmed an excellent correlation between most instruments. A comparison of both the electronic (range 0–4) and paper-based (range 0–3) VRS proved a modera-

Table I. Validation of ItchApp®: subject demographics and itch intensity scores

	Total cohort (n = 40)	Male (n = 17; 42.5%)	Female (n = 23; 57.5%)
Age, years, mean ± SD	57.9 ± 15.2	60.7 ± 15.7	55.8 ± 14.7
Clinical IFSI Group			
CP on inflamed skin, n (%)	12 (30)	5 (29.4)	7 (30.4)
CP on non-inflamed, n (%)	16 (40)	7 (41.2)	9 (39.1)
Chronic scratch lesions, n (%)	12 (30)	5 (29.4)	7 (30.4)
IFSI category of underlying disease			
Dermatological, n (%)	14 (35)	6 (35.3)	8 (34.8)
Systemic, n (%)	1 (2.5)	0	1 (4.3)
Neurological, n (%)	3 (7.5)	1 (5.9)	2 (8.7)
Psychiatric, n (%)	1 (2.5)	0	1 (4.3)
Multifactorial, n (%)	19 (50)	8 (47.1)	10 (43.5)
Unknown origin, n (%)	2 (5)	2 (11.8)	0
Duration of chronic pruritus			
≤1 years	6 (15)	2 (11.8)	4 (17.4)
1–5 years	16 (40)	6 (35.3)	10 (43.5)
6–10 years	7 (17.5)	4 (23.5)	3 (13.0)
> 10 years	10 (97.5)	5 (29.4)	5 (21.7)
Duration, NA	1 (2.5)	0	1 (4.3)
Itch Intensity – Paper			
VAS mean, 4 weeks, mean ± SD (range)	4.98 ± 2.40 (1–9.5)	4.79 ± 2.44 (1.0–9.0)	5.13 ± 2.40 (1.5–9.5)
VAS worst, 4 weeks, mean ± SD (range)	6.61 ± 2.68 (0.5–10)	6.35 ± 2.74 (1.0–10.0)	6.80 ± 2.69 (0.5–10.0)
VAS mean, 24 h, mean ± SD (range)	4.26 ± 2.62 (0.5–9)	4.23 ± 2.10 (1.0–8.0)	4.28 ± 2.99 (0.5–9.0)
VRS mean, 24 h, mean ± SD (range)	1.6 ± 0.67 (0–3)	1.53 ± 0.62 (0–2)	1.65 ± 0.71 (1–3)
NRS mean, 24 h, mean ± SD (range)	4.73 ± 2.54 (2–10)	4.35 ± 2.31 (2–9)	5.0 ± 2.71 (2–10)
DPS, mean ± SD (range)	34.24 ± 35.34 (–76.54–79.01)	32.35 ± 32.79 (0.0–75)	35.64 ± 37.78 (–76.54–79.01)
Itch Intensity – App			
VRS mean, 24 h ItchApp®, mean ± SD (range)	2.0 ± 0.87 (0–4)	2.18 ± 0.72 (0–3)	1.87 ± 0.96 (0–4)
NRS mean, 24 h ItchApp®, mean ± SD (range)	5.10 ± 2.43 (2–10)	5.29 ± 2.31 (2–9)	4.96 ± 2.56 (2–10)
NRS worst, 24 h ItchApp®, mean ± SD (range)	6.15 ± 2.42 (2–10)	6.35 ± 2.29 (2–9)	6.00 ± 2.55 (2–10)
Global Pruritus Rating, mean ± SD (range)	0.98 ± 0.15 (0/1)	1 ± 0 (1)	0.96 ± 0.20 (0/1)
DPS, Score I (Yes/No), mean ± SD (range)	0.65 ± 0.48 (0/1)	0.59 ± 0.50 (0/1)	0.70 ± 0.47 (0/1)
DPS, Score II (Improved/Worsened), mean ± SD (range)	0.92 ± 0.27 (0/1)	0.90 ± 0.31 (0/1)	0.94 ± 0.25 (0/1)
DPS, Score III (Bands), mean ± SD (range)	2.63 ± 0.92 (1–4)	2.56 ± 1.13 (1–4)	2.67 ± 0.81 (2–4)
Quality of life			
DLQI, mean ± SD (range)	7.93 ± 5.59 (1–24)	7.94 ± 5.48 (1–23)	7.91 ± 5.79 (1–24)
ItchyQol total score, mean ± SD (range)	2.71 ± 0.85 (1.3–4.7)	2.67 ± 0.97 (1.3–4.7)	2.75 ± 0.76 (1.4–4.4)
ItchyQol symptom, mean ± SD (range)	2.96 ± 0.88 (1.2–5.0)	2.93 ± 0.98 (1.2–4.7)	2.98 ± 0.82 (1.7–5.0)
ItchyQol function, mean ± SD (range)	2.85 ± 0.93 (1.0–4.9)	2.80 ± 1.05 (1.0–4.9)	2.88 ± 0.86 (1.0–4.7)
ItchyQol emotion, mean ± SD (range)	2.46 ± 1.07 (1.0–5.0)	2.40 ± 1.20 (1.0–5.0)	2.50 ± 0.99 (1.0–4.4)

CP: chronic pruritus; DPS: Dynamic Pruritus Score; IFSI: International Forum for the Study of Itch; NRS: numerical rating scale; VAS: visual analogue scale; VRS: verbal rating scale; NA: not available; SD: standard deviation.

te–positive correlation (Spearman's correlation, $r=0.61$; $p<0.001$) despite the range of the questions differs. The

Table II. Analysis of ItchApp® test–retest reliability in the validation study: correlation, coefficients and subject assessed results

Number	Item	First use vs. Second use correlation
1	VRS 24 h average	0.865 ² T1. 2.00 ± 0.87 (median: 2.00), T2. 2.10 ± 0.84 (median: 2.00)
2	NRS 24 h average	0.977 ² T1. 5.10 ± 2.43 (median: 5.00), T2. 5.07 ± 2.43 (median: 5.00)
3	NRS 24 h worst	0.948 ² T1. 6.15 ± 2.42 (median: 6.50), T2. 5.95 ± 2.36 (median: 6.50)
4	Global Pruritus Rating	0.481 ¹ ($p=0.075$), pruritus 24 h, Yes (T1=T2): 37/40 (92.5%); No (T1=T2): 1/40 (2.5%); T1≠T2 2/40 (5%)
5	DPS, Score 1	0.894 ¹ ($p<0.001$)
6	DPS, Score 2	0.647 ¹ ($p=0.083$)
7	DPS, Score 3	0.896 ² T1. 2.64 ± 0.953 (median: 3), T2. 2.63 ± 0.924 (median: 2.50)

VRS: verbal rating scale; NRS: numerical rating scale; VAS: visual analogue scale; Global Pruritus Rating (Itching in the last 24 h? Yes or no), DPS Score 1 (Change of itching since the beginning of the study; yes or no), DPS Score 2 (If yes, an improvement or worsening?), DPS Score 3 (score bands): If improved: 1–30% (slight improvement); 31–50% (moderate improvement); 51–70% (good improvement); 71–100% (very good improvement). T1: Timepoint 1; T2: Timepoint 2 – after 60 min; ¹k Coefficient; ²ICC (intraclass correlation coefficient).

design of ItchApp®'s DPS scale differs greatly from that included on the paper-based version. The horizontal scale included in paper-based questionnaires, on which patients are asked to mark their itch intensity, was converted into check boxes. In this case, a Spearman's correlation exhibited moderate scores of 0.46 and $p=0.02$.

Quality of Life: ItchyQol and DLQI. The paper-based assessed DLQI and ItchyQol scores were tested against the itch intensities recorded by ItchApp® (Table III), revealing a moderate correlation with the VRS (Spearman's correlation, $r=0.422/0.429$) and mean NRS score ($r=0.422/0.476$), and NRS worst, which was higher ($r=0.510/0.552$; $p<0.001$). Weak correlations were found in tests including the ItchyQol subscores symptom and function; only emotions displayed a relatively moderate correlation (Spearman's correlation, $r=0.468$ for NRS mean, 0.521 for worst NRS). To determine whether any of the 22 questions of the ItchyQol correlate with ItchApp® questions, Spearman's correlation was again computed, which resulted in a moderate level of correlation found in 8 of 22 questions.

Table III. Inter-item and convergent validity of the ItchApp®

	VRS mean last 24 h, ItchApp®	NRS mean last 24 h, ItchApp®	NRS worst last 24 h, ItchApp®
Itch Intensity assessed via ItchApp®			
VRS	n.a.	0.785 ^a (p < 0.001)	0.685 ^a (p < 0.001)
NRS mean		n.a.	0.808 ^a (p < 0.001)
Itch Intensity assessed via paper			
VAS mean, 4 weeks	0.513 ^a (p < 0.001)	0.671 ^a (p < 0.001)	0.685 ^a (p < 0.001)
VAS worst, 4 weeks	0.442 ^a (p < 0.004)	0.578 ^a (p < 0.001)	0.716 ^a (p < 0.001)
VAS mean, 24 h	0.713 ^a (p < 0.001)	0.879 ^a (p < 0.001)	0.747 ^a (p < 0.001)
VRS mean, 24 h	0.61 ^a (p < 0.001)	0.788 ^a (p < 0.001)	0.764 ^a (p < 0.001)
NRS mean, 24 h	0.613 ^a (p < 0.001)	0.924 ^a (p < 0.001)	0.736 ^a (p < 0.001)
Quality of life assessed via paper			
ItchyQol total score	0.429 ^a (p = 0.006)	0.476 ^a (p = 0.002)	0.552 ^a (p < 0.001)
ItchyQol symptom	0.306 ^a (p = 0.054)	0.299 ^a (p = 0.06)	0.373 ^a (p = 0.018)
ItchyQol function	0.342 ^a (p = 0.03)	0.395 ^a (p = 0.012)	0.492 ^a (p = 0.001)
ItchyQol emotion	0.369 ^a (p = 0.019)	0.468 ^a (p = 0.002)	0.521 ^a (p = 0.001)
DLQI score	0.426 ^a (p = 0.006)	0.422 ^a (p = 0.007)	0.510 ^a (p = 0.001)

Inter-Item validity was calculated of ItchApp® intensities timepoint 1 vs. ItchApp® intensities timepoint 2. Convergent validity was calculated of ItchApp® intensities vs. paper intensities and paper Quality of Life scores. All data were obtained in the validation study. ^aSpearman's rank correlation coefficient r. NRS: numerical rating scale, VAS: visual analogue scale, VRS: verbal rating scale; DLQI: Dermatology Life Quality Index.

Use of ItchApp® in a clinical trial

The RCT was performed in 2 centres located in Münster and Düsseldorf, Germany and information can be found elsewhere (ClinicalTrial.gov No. NCT02004041). ItchApp® use analysis showed good data integrity and technical and compliance properties (Table SI¹, Fig. S2¹). Completion time was short (2.3 min); the technical use was influenced by the subject's age and sex. The first band (slight improvement) of ItchApp's® DPS (anchor of MCID calculation) was selected by 46 of 60 subjects at day 27.6 ± 14.1 (range 2–49) of the study. The mean NRS value was 6.65 ± 1.98 (median 7.0, range 1–10) at baseline in this group. On the final day, on which subjects reported a slight improvement, the NRS value was 4.04 ± 1.95 (median 4.0, range 0–8), resulting in an NRS-MCID of 2.61 (*p* < 0.001; Wilcoxon signed-rank test). To achieve a minimal clinically important improvement, it was thus necessary to adjust the mean NRS by 2.61 points.

Integrity and transmission of data in the clinical trial

Sixty of 69 randomized subjects completed the study and a total of 2,497 total entries were saved in the database. Of these entries, the majority (2,419 days; 96.9%) was entered within the requested time period of up to 3 days (Fig. S2¹). Among these, 70.5% of the subjects always made the entry on the same day. The exact entry time could not be saved in 3.1% of cases due to technical interferences (e.g. an audit trail missing from the data packets, etc.), resulting in empty data lines in the table. For example, there were 4 technical errors identified (0.16%) related to entries within 4 days, and 1 in 11 days (0.04%) showing that there were data synchronization issues. In addition, some subjects forgot to finish completing the questionnaires within the given time period. Of the total entries 1.2% (29) were completed the day after they were begun, leading to the assumption that this small percentage of subjects did not save some entries or forgot to complete them.

Feasibility questionnaire

The validation study subjects (*n* = 40; “validation subjects”) as well as the majority from the clinical trial (*n* = 53/60, “trial subjects”) were asked to complete a paper-based feasibility questionnaire based on their experience with the smartphone and ItchApp®. The validation subjects can be considered as short-term users (ItchApp® use during a single visit), and the trial subjects as long-term users. This method of evaluation allows for a comprehensive assessment of how users navigate the application. In summary, the majority of subjects were able to use the smartphone for data entry without issue and ItchApp® was described as simple to use and easily understood (Table SI¹). Interestingly, the patients who had difficulties with the devices were, in mean, 17.6 years older (*n* = 10; age 63.2 ± 8.2 years; *p* = 0.001) than those who were able to use them with ease (33.4 ± 9.9 years). The same was true for the touch screens (13/15 subjects older than the median of 49 years in the collective; *p* = 0.001). In interviews, some subjects stated that they had never used a smartphone before, but could imagine themselves using one if needed. However, the majority of subjects could imagine themselves once again participating in a study in which smartphones are used. Subjects were asked if they considered ItchApp® to be viable for measuring pruritus (face validity). Thirty-nine (97.5%) of the validation subjects and all trial subjects agreed that it fulfils its purpose. In summary, our data suggests that, regarding sex and age, there are no severe limitations to using the app.

Completion time

Subjects required a short amount of time to complete ItchApp®. Asked for the estimated time which was needed to complete ItchApp®, validation subjects reported a mean ± SD of 3.97 ± 3.30 min (median: 3; range: 0.5–20 min). With information from the database, it was found that the validation subjects overestimated the time

they needed for data entry: a mean \pm SD of 2.3 ± 1.2 min (median 2.1; range 0.53–7.41 min) was required during the first session. The time range was even shorter during the second session (1.4 ± 0.54 min; median: 1.2; range: 0.31–3.46 min). During the clinical trial, the mean \pm SD was 2.3 ± 5.3 min (median 0.8; range 0.39–30.71 min). A few patients had interrupted their entry at some point (maximum 30 min for data entry). We found a low negative correlation (-0.09 , Pearson) in the entry times needed for long-term users in younger patients who were analysed for age and sex differences. This group was found to need less time for data entry. These findings, however, are insignificant ($p=0.49$). Male subjects needed a mean of 1.9 min less than female subjects.

DISCUSSION

ItchApp[®] was developed and validated for use in clinical trials according to the COSMIN (<http://www.cosmin.nl/>) recommendations for health measurement instruments (13). ItchApp[®] demonstrated high quality and integrity of data assessment, patient compliance and easy data management. This was demonstrated in a validation study and a RCT. We used validated scales (7, 8, 12) and easy-to-use smartphone controls. Development and validation of the instrument was performed according to rules regarding ICH-GCP regulations in terms of validation and security (e.g. data encryption). ItchApp[®] thus provides many benefits that could be of use in future clinical trials. Furthermore, the system provides scalable statistics and can be adapted easily for use in new studies by modifying the study information, the sponsor's information and the questionnaire content, if necessary.

Equivalence

When developing ItchApp[®], it was necessary to make some minor adjustments to the original paper-based questionnaires. This study, for example, abstained from using a digital VAS due to technical difficulties (e.g. smartphone screen sensitivity). The DPS was also adjusted to better suit the smartphone screen. These changes were validated due to the possibility that they could affect subjects' responses (14). The test–retest reliability method provided excellent results (ICC 0.86–0.97), thus proving the scales' consistency (15). Convergent validity tests revealed high correlations between the paper-based and digital intensity scales. These positive response rates correspond to the results of a meta-analytic review that used correlation tests to confirm the high compatibility of instruments of both eDiaries and paper-based diaries (16). Ninety-four percent of the studies analysed for the review had, according to correlation and repeatability tests, values of ≥ 0.75 .

Compliance

PROs are well-known endpoints used in diaries of RCTs that are absolutely vital to various therapeutic areas (17). However, it is well known that many patients are non-compliant regarding paper-based diaries, in contrast to the use of eDiaries (18). The level of compliance with paper-based diaries was found to be 11% in a study of subjects with chronic pain (18). In contrast, the eDiary in this study achieved a 94% level of compliance. Accordingly, several features were added to ItchApp[®] in order to enhance patient compliance and ensure the integrity and correctness of the data. These include, for example, a feature preventing subjects from skipping questions, a summary of the answers that allows for changes to be made, and an alarm, which can be individually activated to remind the subject to complete their daily entry. Low compliance may be due to subjects' unfamiliarity with itch intensity scales and training sessions are advocated (7). Subjects' understanding of ItchApp[®] scales was, however, ensured by the demo's features, possibly contributing to the documented high level of compliance. Used in the RCT, 87% of subjects reported that they answered the questions daily; however, the real proportion (without retrospective entries) was 70.5%. This difference of 16.5% is considerably low compared with previous studies (18). In addition, 96.9% of entries were made within the required 3 days, allowing for a conclusion to be drawn as to the course of the pruritus. The mean time needed for data entry was relatively short (2.3 min), which also has a positive impact on subject compliance, as the diary entry can be completed quickly.

Administration and data management

Several points regarding the administration and management of data have to be considered, for example, completion of data and management of missing data, accurate entry times, data confidentiality, information provided to subjects (including emergency contacts, study design) and long-term data storage (19). ItchApp[®] addressed all of these areas. Nearly all ItchApp[®] entries were complete. In 3.1% of cases, due to a technical interference, the exact input time was not saved, although this problem was resolved quickly. However, the invalid data could be identified easily, thus lowering the burden on the study associates, as invalid data can be sorted automatically.

Both health authorities and patients are increasingly aware of the necessity for data confidentiality in contemporary society. Each subject was assigned an identification code and their personal information was anonymized and data automatically encrypted once they had completed the questionnaires. Data cannot be modified following encryption, ensuring its secure, long-term storage in the central database. After the trial is concluded, the collected data can be analysed immediately using Excel or

SPSS without the need for manual data entry. The server maintenance and database management was assigned to an IT company (Arone), leaving only the download of the data to the protected server website to the clinician. In summary, ItchApp® provides the opportunity to decrease administrative efforts and to simplify the data collection process. As mentioned above, the data generated by ItchApp® allows for easy calculation of the MCID by using the first category of the DPS as the anchor. Our NRS-MCID of 2.6 can be compared with the paper-based NRS-MCID reported recently (20). In a study by Reich et al., the MCID of the NRS values for the past 24 h was found to be 2.7 (20). The MCID for ItchApp®'s NRS is very similar and serves to further emphasize the successful conversion of paper-based questionnaires into electronic questionnaires.

Future prospects

The app was used not only in the validation study and RCT, but also in other recent clinical trials performed in Germany. There are currently plans to validate other various translated versions of ItchApp® that will enable use of the app in multilingual studies. At present, ItchApp® is available in German and English, of which the former is validated. This app is designed for daily use in RCTs by a large number of patients. In 2015 44 million people in Germany used smartphones, indicating the practicability of an eDiary (21). The app runs on the Android OS, which has the biggest market share, both in Germany (77.7% in July 2016) and globally, therefore many patients will be familiar with the interface (22), thus enabling a broad use in RCTs.

ACKNOWLEDGEMENTS

The authors thank Emily R. Burnett for her assistance in the preparation of this manuscript. The authors would also like to extend their gratitude to DermLink and Vanda Pharmaceuticals for their advice and technical support. This study was supported by the German Federal Ministry of Education and Research (BMBF; FKZ 01KG1305).

Conflict of interest: ItchApp® is licensed to S. Ständer and D. Phan. The other authors declare no conflicts of interest.

LICENSE

ItchApp® is licensed to S. Ständer and Arone. To request use of ItchApp® in clinical trials and other research, contact sonja.staender@uni-muenster.de or david.pihan@arone.com.

REFERENCES

1. Ständer S, Weisshaar E, Mettang T, Szepietowski JC, Carstens E, Ikoma A, et al. Clinical classification of itch: a position paper of the International Forum for the Study of Itch. *Acta Derm Venereol* 2007; 87: 291–294.
2. Mattered U, Apfelbacher CJ, Loerbroks A, Schwarzer T, Büttner M, Ofenloch R, et al. Prevalence, correlates and characteristics of chronic pruritus: a population-based cross-sectional study. *Acta Derm Venereol* 2011; 91: 674–679.
3. Ständer S, Schäfer I, Phan NQ, Blome C, Herberger K, Heigel H, et al. Prevalence of chronic pruritus in Germany: results of a cross-sectional study in a sample working population of 11,730. *Dermatology* 2010; 221: 229–235.
4. Siepmann D, Weishaupt C, Luger TA, Ständer S. Evaluation of the German Guideline for chronic pruritus: results of a retrospective study on 385 patients. *Dermatology* 2011; 223: 374–380.
5. Ständer S, Weisshaar E, Raap U. Emerging drugs for the treatment of pruritus. *Expert Opin Emerg Drugs* 2015; 20: 515–521.
6. Ständer S, Augustin M, Reich A, Blome C, Ebata T, Phan NQ, et al. Pruritus assessment in clinical trials: a consensus paper of the special interest group scoring itch in clinical trials of the International Forum for the Study of Itch (IFSI). *Acta Derm Venereol* 2013; 93: 509–514.
7. Phan NQ, Blome C, Fritz F, Gerss J, Reich A, Ebata T, et al. Assessment of pruritus intensity: prospective study on validity and reliability of the visual analogue scale, numerical rating scale and verbal rating scale in 471 patients with chronic pruritus. *Acta Derm Venereol* 2012; 92: 502–507.
8. Reich A, Heisig M, Phan NQ, Taneda K, Takamori K, Takeuchi S, et al. Visual analogue scale: evaluation of the instrument for the assessment of pruritus. *Acta Derm Venereol* 2012; 92: 497–501.
9. Krause K, Kessler B, Weller K, Veidt J, Chen SC, Martus P, et al. German version of ItchyQoL: validation and initial clinical findings. *Acta Derm Venereol* 2013; 93: 562–568.
10. Elman S, Hynan LS, Gabriel V, Mayo MJ. The 5-D itch scale: a new measure of pruritus. *Br J Dermatol* 2010; 162: 587–593.
11. Blome C, Augustin M, Siepmann D, Phan NQ, Rustenbach SJ, Ständer S. Measuring patient-relevant benefits in pruritus treatment: development and validation of a specific outcomes tool. *Br J Dermatol* 2009; 161: 1143–1148.
12. Ständer S, Blome C, Anastasiadou Z, Zeidler C, Jung KA, Tsianakas A, 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; 230–234.
13. Mokkink, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: A clarification of its content. *BMC Med Res Methodol* 2010; 10: 22.
14. U.S. Department of Health and Human Services. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims. 2009 [cited 2016 Jul 22]. Available from: <http://www.fda.gov/downloads/Drugs/.../Guidances/UCM193282.pdf>.
15. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychol Methods* 1996; 1: 30–46.
16. Gwaltney CJ, Shields AL, Shiffman S. Equivalence of electronic and paper-and-pencil administration of patient-reported outcome measures: a meta-analytic review. *Value in Health* 2008; 11: 322–333.
17. Willke RJ, Burke LB, Erickson P. Measuring treatment impact: a review of patient-reported outcomes and other efficacy endpoints in approved product labels. *Controlled Clinical Trials* 2004; 25: 535–552.
18. Stone AA, Shiffman S, Schwartz JE, Broderick JE, Hufford MR. Patient non-compliance with paper diaries. *BMJ* 2002; 324: 1193–1194.
19. Wehrlen L, Krumlauf M, Ness E, Maloof D, Bevans M. Systematic collection of patient reported outcome research data: a checklist for clinical research professionals. *Contemp Clin Trials* 2016; 48: 21–29.
20. Reich A, Riepe C, Anastasiadou Z, Mędrak K, Augustin M, Szepietowski JC, et al. Itch assessment with visual analogue scale and numerical rating scale: determination of minimal clinically important difference in chronic itch. *Acta Derm Venereol* 2016; 96: 9–11.
21. Bitkom. 44 Millionen Deutsche nutzen ein Smartphone. Mar 2016. Available from: <https://www.bitkom.org/Presse/Presseinformation/44-Millionen-Deutsche-nutzen-ein-Smartphone.html>.
22. Kantar. Smartphones OS market share Jul 2016. Available from: <http://www.kantarworldpanel.com/global/smartphone-os-market-share/>.