CLINICAL REPORT

Check for updates

Measurement of Nocturnal Scratching in Patients with Pruritus Using a Smartwatch: Initial Clinical Studies with the Itch Tracker App

Akihiko IKOMA[1,2], Toshiya EBATA[3], Laurent CHANTALAT[2], Kimitoshi TAKEMURA[2], Fabienne MIZZI[4], Michel PONCET[4] and Didier LECLERCQ[2]

[1]Department of Future Design, Maruho Co., Ltd, Osaka, Japan, [2]SHIELD Center, Nestlé Skin Health SA, USA, [3]Chitofuna Dermatology Clinic, Setagaya, Tokyo, Japan, and [4]Research and Development, Galderma SA, France

Three clinical studies were conducted to test a newly-developed app for smartwatches, which included an algorithm to measure nocturnal scratching using acceleration data. The first study in 5 patients with atopic dermatitis demonstrated high reliability of the app for measurement of scratching compared with video monitoring (positive predictive value 90.2±6.6%, sensitivity 84.6±10.2%, correlation of scratching duration per h r = 0.851–0.901, $p < 0.001$). The second study in 20 patients with atopic dermatitis and 10 healthy volunteers showed that total scratching duration in patients was significantly longer than in healthy volunteers and correlated positively with Eczema Area and Severity Index (EASI) scores. In the third study, conducted in an open-entry manner in which 201 evaluable participants measured nocturnal scratching, those who self-reported itch or pruritic diseases had a significantly longer duration of scratching than those who did not. In conclusion, this app has a high reliability and potential clinical usefulness for measurement of nocturnal scratching.

Key words: actigraphy; itch; scratching; atopic dermatitis; smartwatch.

Accepted Dec 6, 2018; E-published Dec 6, 2018

Acta Derm Venereol 2019; 99: 268–273.

Corr: Akihiko Ikoma, Department of Future Design, Maruho Co., Ltd., 1-5-22, Nakatsu, Kita-ku, Osaka, 531-0071, Japan. E-mail: ikoma_eom@mii.maruho.co.jp

SIGNIFICANCE

This paper reports 3 clinical studies to measure scratching in patients with chronic itch, using a new app developed for use in widely-used smartwatches. The results indicate that the smartwatch app is reliable and useful for measurement of scratching during sleep, enabling the indirect, but objective, evaluation of itch severity. The app has potential for use in daily clinical settings.

Itch is a major symptom associated with various dermatological, systemic and neurological conditions. Chronic itch seriously impairs the quality of patients' life. Itch is usually evaluated by patients using subjective scales, such as a visual analogue scale (VAS) in clinical settings. The lack of objective quantitative methods to evaluate itch makes it difficult to evaluate itch severity during sleep, especially in small children and elderly people with cognitive impairment, and hence difficult to precisely assess the efficacy of antipruritic therapies.

Since itch provokes the desire to scratch, measurement of scratching can be used to indirectly, but objectively, evaluate itch intensity. There have been attempts to measure scratching by various methods, such as video-monitoring, acoustic assessment and measurement of body movements (1–3). The most frequently applied method is the measurement of wrist movement, or actigraphy (4). A large variety of actigraphy devices has been used, some of which are marketed as products for sleep assessment (5–9). Their common feature is that acceleration signals produced by wrist movement are recorded by a wristwatch-like device equipped with a built-in accelerometer, followed by analysis of data to quantify scratching. Some recent studies have attempted to differentiate signals caused by scratching signals from those caused by other types of movement, using a specific algorithm to accurately quantify scratching activity (10–12). However, all of the above actigraphy devices have been produced for specific purposes of measurement and are not general smartwatches that are widely used, thus hampering the spread of this method for daily use.

In the studies reported here, a newly-developed application software (app) containing a unique algorithm, termed "Itch Tracker", was installed into an Apple Watch (Apple Inc., Cupertino, USA), one of the most widely-used smartwatches worldwide, and used for measurement of scratching during sleep. ResearchKit, an open-source framework offered by Apple Inc. for medical research using mobile devices, was used to collect data from a large number of participants.

METHODS

Study objectives

A total of 3 clinical studies were conducted. The first study (Study 1) was conducted to verify the reliability of the app for measurement of nocturnal scratching compared with video monitoring. The second study (Study 2) was conducted to compare patients with atopic dermatitis and healthy volunteers with regards to nocturnal scratching measured by the app. The third study (Study 3) was conducted in an open-entry manner, with anonymous users of smartwatches worldwide who downloaded the app from the service site (App Stores, Apple Inc. USA) and installed it on

doi: 10.2340/00015555-3105
Acta Derm Venereol 2019; 99: 268–273

This is an open access article under the CC BY-NC license. www.medicaljournals.se/acta
Journal Compilation © 2019 Acta Dermato-Venereologica.

ActaDV Acta Dermato-Venereologica ActaDV Advances in dermatology and venereology

ActaDV Acta Dermato-Venereologica ActaDV Advances in dermatology and venereology

their smartwatches, to investigate the correlation between scratching duration during sleep and other factors. All 3 studies were conducted in accordance with the Declaration of Helsinki 1964 and its later amendments, and in accordance with ethics committee approval (Studies 1 and 2; OPHAC Hospital No. 977, Study 3; Yoyogi Mental Clinic No. NSH-901). All study subjects gave informed consent before participation.

Itch Tracker app

An app named "Itch Tracker", incorporating an algorithm that analyses acceleration data, detects scratching and differentiates it from other types of movement, was developed so that a smartwatch in which the app was installed could be used to measure scratching. Acceleration data in x-, y-, and z-dimensions were recorded at 50 Hz during the whole period of measurement and, once the measurement was terminated, transferred from the smartwatch to the phone device, where the processor scanned all changes in acceleration with the algorithm for analysis. The algorithm was created on the basis of pre-studies that revealed the pattern of acceleration characteristics in scratching movements of the wrist, differentiated from other types of movement. The algorithm recognizes scratching as starting when the change in acceleration in any of the 3 dimensions continuously meets all of the following conditions for 3 s or longer, and recognizes scratching to to end when at least one of the conditions is no longer met during a predetermined minimum number of consecutive time-periods in the same dimension (**Fig. 1**a):

- the length of a time-period (T_X) does not exceed a predetermined upper limit. A time-period is defined as a period between the time when the acceleration measured by the actigraph sensor is zero and the subsequent time when the acceleration measured by the actigraph sensor is again zero;
- the difference in the length of a time period (T_X) and the length of the next time period (T_{X+1}) does not exceed a predetermined upper limit;
- the maximum absolute value of the acceleration during a time period (M_X) is not less than a predetermined lower limit;
- the difference in the maximum absolute value of acceleration during a time-period (M_X) and the maximum absolute value of acceleration during the next time-period (M_{X+1}) does not exceed a predetermined upper limit.
- the error rate does not exceed a predetermined upper limit, with the error rate being defined as the percentage of time periods that fail to meet any of the above conditions.

An example of acceleration patterns that were recognized as scratching is shown in Fig. 1b. Since the acceleration pattern is the only information that the algorithm uses to recognize scratching, some activities, such as tooth-brushing and neurological disorders, such as Parkinson's disease, which lead to similar acceleration patterns of wrist movement, may cause false-positive results.

Study participants and methods

Study 1. Five male Japanese patients with atopic dermatitis (age 36.2 ± 9.3 years, mean ± standard deviation (SD)) participated in the study. The subjects underwent a physical examination at the study site by a dermatologist on the day of enrollment. The severity of atopic dermatitis was evaluated as moderate in all subjects (Eczema Area and Severity Index (EASI): 15.9 ± 1.8). Self-evaluated itch severity for the previous 7 days and on the day of assessment was 6.6 ± 1.0 and 7.2 ± 0.7, respectively, on a 0–10 numerical scale (0: no itch, 10: worst imaginable itch). The patients were hospitalized on the same day for one night and video-monitored while asleep, sleeping on a bed in an air-conditioned room with a thin blanket to cover the body, so that the evaluator could easily recognize their body motions when analysing the video. Each subject wore 2 identical smartwatches (Apple Watch Sport 38 mm, Apple Inc., USA), in which the app was installed, with one watch on each wrist. The evaluator, who was blinded to the smartwatch data, watched the video of the subjects and recorded the timing and duration of all body motions during sleep. If the motions were judged as scratching by the evaluator, they were recorded together with the information on which arm was used and categorized into 3 grades depending on the intensity of scratching motions; "weak" if only fingers were used to scratch without visible movement of the elbow, "moderate" if the elbow visibly moved, and "strong" if both of the elbow and the shoulder visibly moved. Scratching motions with intervals of less than 3 s were considered as a single scratching episode. All motion episodes that were recorded as scratching by the watch and those that were recorded as scratching with the same arm by the evaluator were compared. If a watch-recorded scratching episode was partially or completely overlapped in terms of timing with any evaluator-recorded scratching episode with the same arm, both episodes were considered as consistent. The percentage of consistent episodes out of all scratching episodes recorded by the watch was defined as the positive predictive value. The percentage of consistent episodes out of all scratching episodes recorded by the evaluator, except for the episodes in which the duration was less than 3 s or the intensity was "weak", was defined as the sensitivity. The total evaluator-recorded scratching duration during sleep was compared between dominant and non-dominant arms in each subject. The sleep duration was equally divided into multiple sections in each subject, so that each section of sleep would be approximately 1 h (between 1 h and

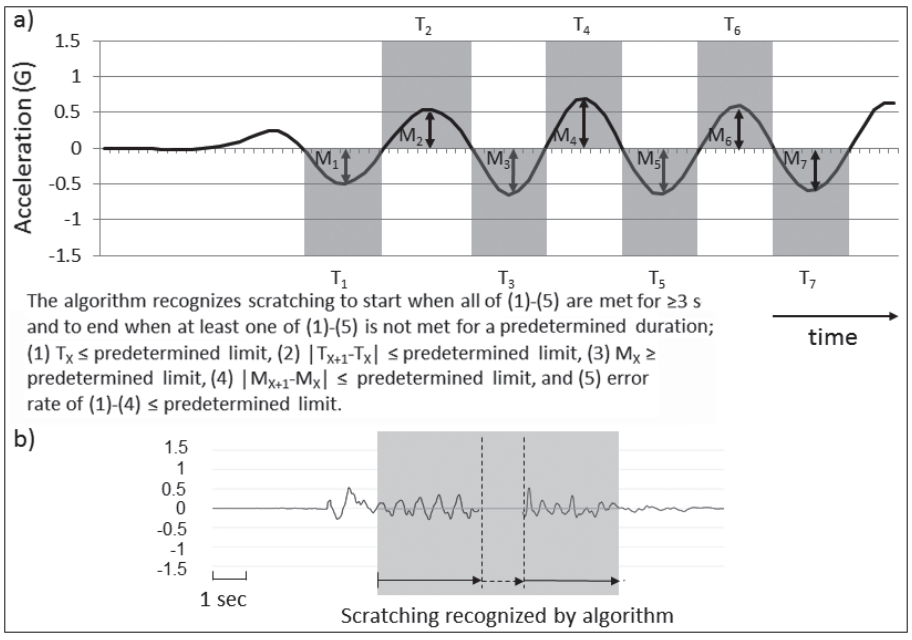

Fig. 1. (a) Algorithm of Itch Tracker app to measure scratching. (b) Example of actual acceleration patterns.

ActaDV Acta Dermato-Venereologica ActaDV Advances in dermatology and venereology

1 h 10 min) and the total watch-recorded scratching duration with 1 arm in each hourly section of sleep was compared with the total evaluator-recorded scratching duration in the same hourly section of sleep (scratching with the same arm except for short or weak scratching, scratching with the same arm including short and weak scratching, or scratching with both arms including short and weak scratching).

Study 2. A total of 20 Japanese patients with atopic dermatitis (11 males and 9 females, aged 38.8 ± 9.3 years), and 10 Japanese healthy volunteers (5 males and 5 females, aged 35.5 ± 5.3 years) participated as the study subjects. The subjects underwent a physical examination by a dermatologist at the study site on the day of enrollment and the severity of atopic dermatitis in the patients was evaluated as mild to severe (EASI 10.4 ± 6.3). The subjects wore one smartwatch (Apple Watch Sport 38 mm, Apple Inc., USA), in which the app was installed, while asleep at home for 7 consecutive days starting from the day of enrollment. They wore the smartwatch on the same arm on which they usually wore their own watch. The scratching duration was recorded by the smartwatch and the severity of itch and daytime life and sleep disturbance due to itch were self-evaluated using a 0–10 numerical scale (0: no itch or disturbance, 10: worst imaginable itch or disturbance) and recorded by the subjects every day.

Study 3. In study 3, the app was developed under an open-source framework for medical research targeting mobile phone users (ResearchKit, Apple Inc., USA) and published as an app for open access on an app-downloading service site (App Stores, Apple Inc. USA). The study subjects downloaded and installed the app to their own smartwatch as well as the paired mobile phone and gave their informed consent on a screen in the app. Besides the function to measure scratching, the app also contained a questionnaire with several questions for the study subjects to answer regarding demography (age range, sex, and ethnicity), the medical background related to 8 diseases (atopic dermatitis, dry skin, chronic urticaria, psoriasis, other skin diseases, liver dysfunction, renal dysfunction, and diabetes) and self-evaluated severity of itch (itch in the past 7 days and today's itch) and today's daytime life and sleep disturbance due to itch on a numerical rating scale (0–10). Measurement of scratching, together with the self-evaluation of the severity of today's itch and today's daytime life and sleep disturbance due to itch, was repeatable as many times as the subject wished. All data of scratching measurement and the answered questionnaire were transferred to the investigator's data server with no individual-identifying information. In order to exclude the cases that apparently deviated from the protocol, the data were considered evaluable only if the sleep started in a 10-h time slot between 19.00 h and 05.00 h and the total sleep duration was between 4 and 10 h.

A total of 262 subjects participated in the study, of which 201 subjects (180 males, 21 females, age; mainly 30–39 and 40–49 years, ethnicity; 181 Asian, 19 White and 1 other) provided evaluable data. The evaluable subjects included 49 patients with atopic dermatitis, 74 with dry skin, 11 with chronic urticaria (urticaria lasting for 6 weeks or longer), 12 with psoriasis, 41 with other skin problems, 1 with renal dysfunction, 9 with liver dysfunction, 7 with diabetes and 81 with none of these diseases, according to self-reported diagnoses (**Fig. 2**). If a single subject repeated scratching measurement multiple times, the first evaluable data of the subject was adopted for analysis. Referring to the interpretation of itch severity on a numerical rating scale proposed by the International Forum for the Study of Itch (13), the itch rating was graded into 3 categories; none for 0, mild for 1–2, moderate for 3–6, severe/very severe for 7–10. The same grading system was also applied to the severity of daytime life and sleep disturbance due to itch on a numerical rating scale.

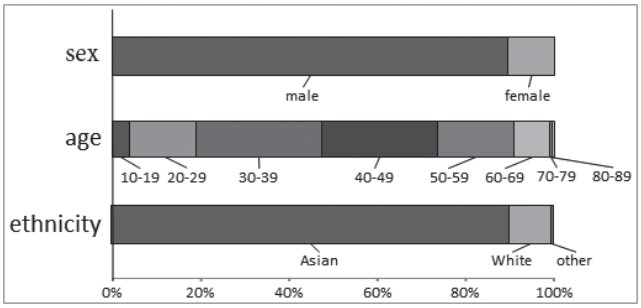

Fig. 2. Demography in Study 3 ($n = 201$).

RESULTS

Study 1

The positive predictive value was 90.2 ± 6.6%, i.e. 90.2 ± 6.6% of the watch-detected scratching episodes were consistent with the episodes recorded as scratching by the evaluator. The sensitivity was 84.6 ± 10.2%, i.e. 84.6 ± 10.2% of the evaluator-recorded scratching episodes, except for the episodes that were weak scratching or lasted for less than 3 s, were also detected as scratching by the watch. There was no significant difference in the scratching duration (in s, and % of total sleep duration) between dominant and non-dominant arms (296 ± 149 s and 0.90 ± 0.47% vs. 388 ± 210 s and 1.17 ± 0.64%). The total watch-recorded scratching duration in each hourly section of sleep was positively correlated in a statistically significant manner with the total evaluator-recorded scratching duration in the same section of sleep with the same arm excluding short and weak scratching

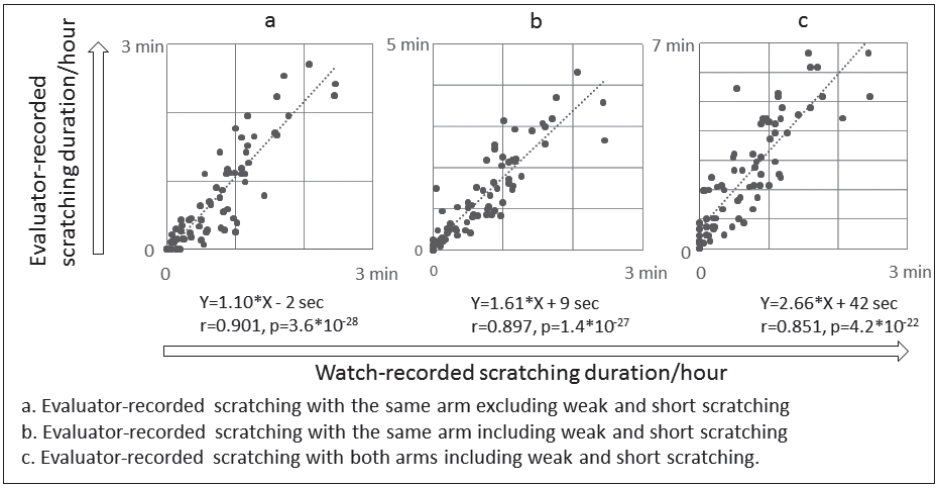

Fig. 3. Correlation of watch-recorded scratching duration/hour with evaluator-recorded scratching duration/hour in Study 1 ($n = 75$). Correlation analysis with Pearson's correlation coefficient.

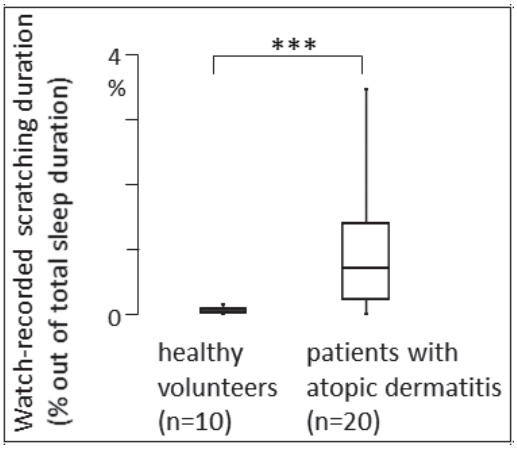

Fig. 4. Comparison of watch-recorded scratching duration (% out of total sleep duration) between healthy volunteers and patients with atopic dermatitis in Study 2. ***$p < 0.001$, Wilcoxon rank-sum test.

(r = 0.901, slope = 1.10, $p = 3.6*10^{-28}$), with the same arm including short and weak scratching (r = 0.897, slope = 1.61, $p = 1.4*10^{-27}$) and with both arms including short and weak scratching (r = 0.851, slope = 2.66, $p = 4.2*10^{-22}$) (**Fig. 3**).

Study 2

The mean 7-day of total watch-recorded scratching duration (in s, and as a % of total sleep duration) in the 20 patients with atopic dermatitis was significantly longer than in the 10 healthy volunteers (median 151 s and 0.72% vs. 7 s and 0.04%, $p = 0.000184$) (**Fig. 4**), and was positively correlated with the EASI score (r = 0.60, $p = 0.00841$) (**Fig. 5**). On the other hand, the self-evaluated severity of itch, daytime life and sleep disturbance due to itch were not significantly correlated with the EASI score or with the total duration of watch-recorded scratching.

Study 3

The subjects who self-reported itch had a significantly longer duration of watch-recorded scratching than those who did not (**Fig. 6**). The participants who had atopic dermatitis, dry skin, chronic urticaria or psoriasis had a significantly longer watch-recorded scratching duration (in s, and as a % of total sleep duration) compared with those who had none of the 8 diseases; e.g., 274 ± 313 s and 1.22 ± 1.34% in those with atopic dermatitis vs. 41 ± 52 s and 0.18 ± 0.22% in the control (**Fig. 7**).

DISCUSSION

The studies reported here provide evidence that the newly-developed app, Itch Tracker, in combination with Apple Watch enables measurement of nocturnal scratching in a simple and reliable manner. The algorithm used is different from other algorithms and systems used in previous studies (10–12). The validated high performance of the app shown by the high sensitivity and high positive predictive values fits the clinical use of this method.

One of the limitations of this method is that short (less than 3 s) and weak (using only fingers) scratching cannot be detected. The limitation in detecting short scratching is specific to the applied algorithm, for which the minimum length of movement for assessment was set as 3 s so that the accuracy of the scratching detection by the algorithm, represented by positive predictive values in the present studies, would not be deteriorated. The 3-s duration has often been considered the minimum duration of physically-meaningful scratching, as some previous studies of scratching also counted movements as scratching only when they were 3 s or longer (3, 14, 15). The limitation in the detection of weak scratching in which only fingers are used is another common feature with the wrist actigraphy, when it is used for the measurement of scratching. In practice, the method in the present studies is capable of detecting harmful scratching, i.e. scratching that can damage skin, since it is usually long-lasting and moderate or strong from the perspective of strength associated with a visible movement of arms. The correlation of evaluator-recorded duration of scratching, per h, that lasted 3 s or longer and was moderate or strong vs. watch-recorded duration of scratching per

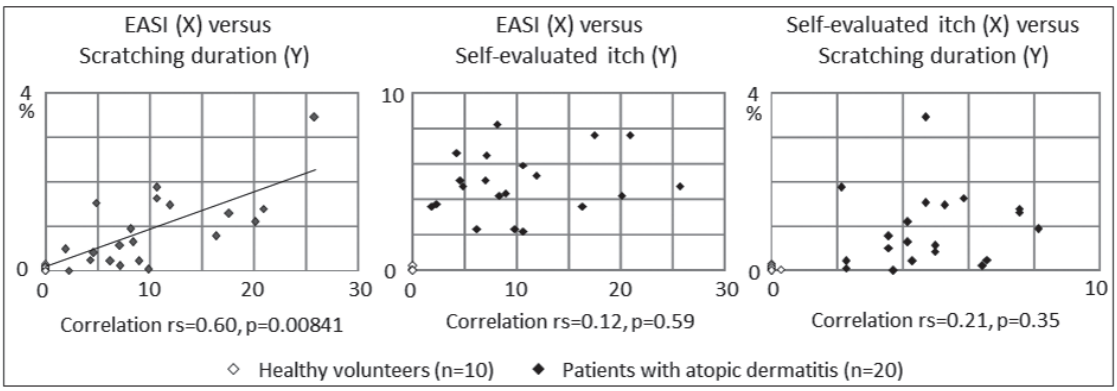

Fig. 5. Correlation among Eczema Area and Severity Index (EASI), self-evaluated severity of itch and watch-recorded scratching duration (% of total sleep duration) in Study 2. Correlation analysis (only patients with atopic dermatitis) with Spearman's rank correlation coefficient.

ActaDV Acta Dermato-Venereologica ActaDV Advances in dermatology and venereology

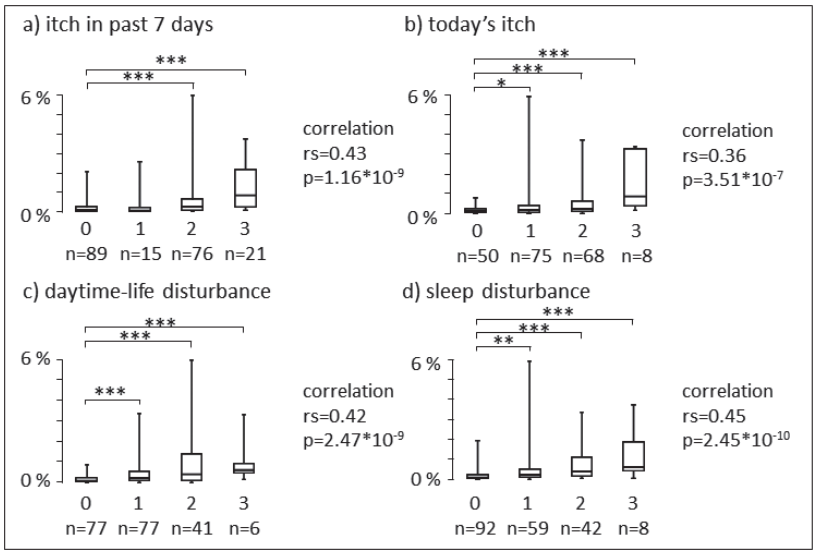

Fig. 6. Severity of itch/disturbance vs. total scratching duration (% out of total sleep duration) in Study 3. 0 = none (0/10), 1 = mild (1–2/10), 2 = moderate (3–6/10), 3 = severe/very severe (7–10/10). *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, Wilcoxon rank-sum test. Correlation analysis with Spearman's rank correlation coefficient.

hour was statistically positive, with Y = 1.10*X – 2 s (Fig. 3a), which is close to Y = X, suggesting high accuracy of this method in the measurement of the total duration of nocturnal scratching that lasts for 3 s or longer and is moderate or strong.

It is intriguing that the watch-recorded duration of scratching, but not the self-evaluated itch severity, correlated positively with EASI in a statistically significant manner in Study 2 (Fig. 5). This is compatible with other previous studies (5, 16) and indicates that the scratching duration during sleep is a better reflection of the severity of atopic dermatitis than self-evaluated itch severity.

In Study 3 with ResearchKit, the majority of enrolled participants were male in their 30s and 40s, reflecting the sex and age distribution of smartwatch users in general (17). Although the major ethnicity of subjects was Asian, perhaps due to the promotional activities for the app that were more intensive in Japan than in other countries, ResearchKit enabled recruitment of study participants worldwide. The watch-recorded scratching duration was significantly longer in participants who had itch or life/sleep-disturbance due to itch, compared with those who did not. Correlation analysis of the self-evaluated itch severity vs. the watch-recorded scratching duration showed a statistical significance, to be different from Study 2. This discrepancy might be partly explained by the large difference in the number of participants between the 2 studies, or the different rating scales applied for statistical analysis (a 0–10 numerical rating scale in Study 2 vs. a 4-level scale in Study 3). Notably, the correlation coefficient was < 0.5, denying the presence of a strongly positive correlation. The participants who self-reported atopic dermatitis, dry skin, chronic urticaria, and psoriasis had a longer watch-recorded scratching duration than those who did not, providing evidence that patients with those pruritic diseases scratch longer during sleep compared with healthy people. Combined with Study 2, it has been demonstrated that the scratching duration during sleep is a good reflection of pruritus with a better correlation with disease severity compared with self-evaluated itch severity.

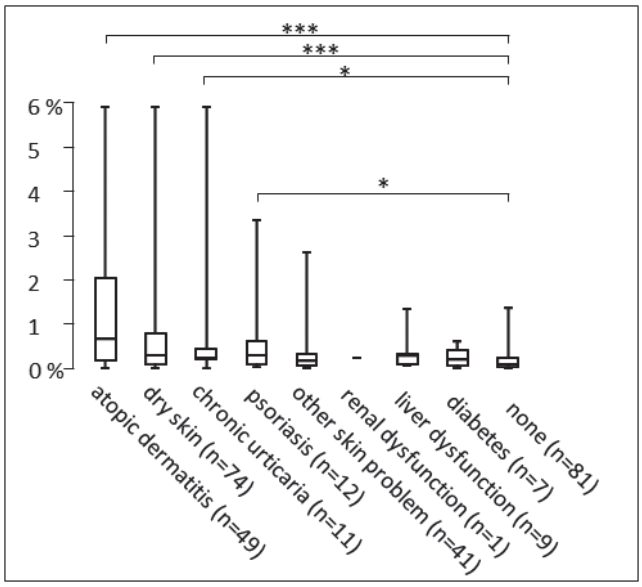

Fig. 7. Watch-recorded total scratching duration (% out of total sleep duration) in each disease category in Study 3. *$p < 0.05$, ***$p < 0.001$, Wilcoxon rank-sum test.

In conclusion, the data in the studies reported here, from a total of 236 study participants, provides evidence of the high reliability and clinical usefulness of the Itch Tracker app in the measurement of nocturnal scratching. Measurement of scratching is an indirect, but objective, way to evaluate the severity of itch. The ease of use of the Itch Tracker app makes it suitable for wide use in daily clinical settings. Further studies are required for the validation of the Itch Tracker app in other unstudied populations, such as children, elderly people, and patients with other pruritic conditions or neurological disorders.

ActaDV Acta Dermato-Venereologica ActaDV Advances in dermatology and venereology

Enhanced validation of the Itch Tracker app is needed to enable it to be developed to provide diagnostic information or medical guidance based on scratching data, which currently is not possible.

ACKNOWLEDGEMENT

Conflicts of interest: The present study was sponsored by Nestlé Skin Health SA. All authors except for Toshiya Ebata were employed by Nestlé Skin Health SA or its affiliate company Galderma SA. Toshiya Ebata had an advisory contract with Nestlé Skin Health SA at the time of the study.

REFERENCES

1. Felix R, Shuster S. A new method for the measurement of itch and the response to treatment. Br J Dermatol 1975; 93: 303–312.
2. Ebata T, Aizawa H, Kamide R. An infrared video camera system to observe nocturnal scratching in atopic dermatitis patients. J Dermatol (Tokyo) 1996; 23: 153–155.
3. Noro Y, Omoto Y, Umeda K, Tanaka F, Shiratsuka Y, Yamada T, et al. Novel acoustic evaluation system for scratching behavior in itching dermatitis: rapid and accurate analysis for nocturnal scratching of atopic dermatitis patients. J Dermatol 2014; 41: 233–238.
4. Ebata T. Measurement of itch: actigraphy. In: Misery L, Staender S, editors. Pruritus. Second edition. Springer-Verlag, London, UK; 2016: p. 97–102.
5. Ebata T, Iwasaki S, Kamide R, Niimura M. Use of a wrist activity monitor for the measurement of nocturnal scratching in patients with atopic dermatitis. Br J Dermatol 2001; 144: 305–309.
6. Benjamin K, Waterson K, Russel M, Schofield O, Diffey B, Rees JL. The development of an objective method for measuring scratch in children with atopic dermatitis suitable for clinical use. J Am Acad Dermatol 2004; 50: 33–40.
7. Hon KL, Lam MC, Leung TF, Kam WY, Lee KC, Li MC, et al. Nocturnal wrist movements are correlated with objective clinical scores and plasma chemokine levels in children with atopic dermatitis. Br J Dermatol 2006; 154: 629–635.
8. Hon KL, Lam MC, Leung TF, Chou CM, Wong E, Leung AK. Assessing itch in children with atopic dermatitis treated with tacrolimus: objective versus subjective assessment. Adv Ther 2007; 24: 23–28.
9. Fujita H, Nagashima M, Takeshita Y, Aihara M. Correlation between nocturnal scratch behavior assessed by actigraphy and subjective/objective parameters in patients with atopic dermatitis. Eur J Dermatol 2014; 24: 120–122.
10. Petersen J, Austin D, Sack R, Hayes TL. Actigraphy-based scratch detection using logistic regression. IEEE J Biomed Health Inform 2013; 17: 277–283.
11. Kaburagi T, Kurihara Y. Algorithm for estimation of scratching time. IEEE Sensors J 2017; 17: 2198–2204.
12. Moreau A, Anderer P, Ross M, Cerny A, Almazan TH, Peterson B. Detection of nocturnal scratching movements in patients with atopic dermatitis using accelerometers and recurrent neural networks. IEEE J Biomed Health Inform 2018; 22: 1011–1018.
13. 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.
14. Aoki T, Kushimoto H, Hishikawa Y, Savin JA. Nocturnal scratching and its relationship to the disturbed sleep of itchy subjects. Clin Exp Dermatol 1991; 16: 268–272.
15. Bender BG, Ballard R, Canono B, Murphy JR, Leung DY. Disease severity, scratching, and sleep quality in patients with atopic dermatitis. J Am Acad Dermatol 2008; 58: 415–420.
16. Bringhurst C, Waterston K, Schofield O, Benjamin K, Rees JL. Measurement of itch using actigraphy in pediatric and adult populations. J Am Acad Dermatol 2004; 51: 893–898.
17. The NPD Group, Inc., Port Washington, US. Wearable ownership by demographics. In: The NPD Group/Connected Intelligence Consumers and Wearables Report, January 6, 2015.

ActaDV Acta Dermato-Venereologica ActaDV Advances in dermatology and venereology