APPLIED SCIENCES AND ENGINEERING

A skin-conformable wireless sensor to objectively quantify symptoms of pruritus

Keum San Chun[1†], Youn J. Kang[2,3†], Jong Yoon Lee[2,4†], Morgan Nguyen[5], Brad Lee[5], Rachel Lee[4], Han Heul Jo[4], Emily Allen[6], Hope Chen[3], Jungwoo Kim[4], Lian Yu[7], Xiaoyue Ni[2,8], KunHyuck Lee[2,3], Hyoyoung Jeong[2,3], JooHee Lee[4], Yoonseok Park[2,3], Ha Uk Chung[2,3,4], Alvin W. Li[6], Peter A. Lio[9], Albert F. Yang[6], Anna B. Fishbein[10], Amy S. Paller[3,6,11], John A. Rogers[2,3,12,13,14]*, Shuai Xu[2,3,10,11,12]*

Itch is a common clinical symptom and major driver of disease-related morbidity across a wide range of medical conditions. A substantial unmet need is for objective, accurate measurements of itch. In this article, we present a noninvasive technology to objectively quantify scratching behavior via a soft, flexible, and wireless sensor that captures the acousto-mechanic signatures of scratching from the dorsum of the hand. A machine learning algorithm validated on data collected from healthy subjects ($n = 10$) indicates excellent performance relative to smartwatch-based approaches. Clinical validation in a cohort of predominately pediatric patients ($n = 11$) with moderate to severe atopic dermatitis included 46 sleep-nights totaling 378.4 hours. The data indicate an accuracy of 99.0% (84.3% sensitivity, 99.3% specificity) against visual observation. This work suggests broad capabilities relevant to applications ranging from assessing the efficacy of drugs for conditions that cause itch to monitoring disease severity and treatment response.

INTRODUCTION

Itch and pain are the two cardinal examples of nociception in humans—itch leads to a scratch reflex, while pain leads to a withdrawal reflex (*1*). Unfortunately, itch is often overlooked and undertreated in the clinical setting (*2*) despite the wide range of medical conditions that cause itch—1% of all outpatient visits yearly involve the symptom of itch representing a major global disease burden (*3*). While a wide range of medical conditions leads to itch (e.g., renal failure, liver failure, and lymphoma), atopic dermatitis (AD) is likely the most common. AD is also the most widespread pediatric inflammatory skin disease, affecting 10 million U.S. children with a yearly prevalence of 13% (*4*). The hallmark of AD is itch, the primary driver of morbidity that leads to chronic sleep disturbance in ~60% of affected children (*5–8*). The consequences for millions of children include neurocognitive impairment and decreased growth (*9–12*). The quality of life of patients with moderate-to-severe AD consistently scores among the lowest of all chronic diseases (*13*).

In AD and other conditions that cause itch, a major unmet need is in objective measures of itch to assess the efficacy of new medications, quantify disease severity, and monitor treatment response. The subjective nature of itch creates difficulties in its quantification. Commonly used methods such as patient reported visual analog scales poorly correlate to visually observed scratching behavior (*14*). These self-reported measures are also prone to both sensitivity and anchoring bias. Furthermore, these subjective methods lack validity in young children and cognitively impaired adults.

Measuring behaviors associated with scratching provides one method to quantify itch severity and frequency. Direct visual inspection of scratching in video recordings represents the gold standard, but these schemes are both labor intensive and impractical in clinical practice. Wrist actigraphy, a method to record movements using accelerometers, estimates scratching by measuring wrist motion. This approach offers suboptimal accuracy because of an inability to distinguish, for example, hand waving from scratching. Additional aspects of the widely varying characteristics of scratching, including variabilities across individuals, represent additional unsolved challenges (*15*). Thus, current sensors and strategies do not offer sufficient performance or practicality for routine use in clinical trials or patient care.

Here, we report a strategy that exploits a soft, flexible, wireless, and low-profile sensor capable of capturing both vibratory and motion signatures of physiological processes with clinical-grade data quality. This device, which we refer to as the ADAM (ADvanced Acousto-Mechanic) (*16*) sensor, softly couples to the skin to enable capture of both low-frequency (e.g., motion) and high-frequency (acousto-mechanic) signals from the human body. An embedded rechargeable battery allows 7 days of continuous operation on a single wireless charge. When placed on the dorsum of the hand, the ADAM sensor is able to capture acousto-mechanic signals associated with scratching via a combination of motion and acousto-mechanic signals in a manner that is immune to ambient noise. Specifically, when placed between the second and third finger meta-carpal bones, the ADAM sensor can quantify scratching behavior initiated not only from motions of the wrist and/or arm but also from the fingers and fingertips across a wide temporal bandwidth that includes high-frequency vibratory motions associated with scratching itself. Using

[1]Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX 78712, USA. [2]Querrey Simpson Institute for Bioelectronics, Northwestern University, Evanston, IL 60208, USA. [3]Center for Bio-Integrated Electronics, Northwestern University, Evanston, IL 60208, USA. [4]Sibel Health, Niles, IL 60714, USA. [5]Northwestern University Feinberg School of Medicine, Chicago, IL 60611, USA. [6]Department of Dermatology, Northwestern University Feinberg School of Medicine, Chicago, IL 60611, USA. [7]Electrical and Computer Engineering, University of Illinois at Champaign-Urbana, Champaign, IL 61801, USA. [8]Department of Mechanical Engineering and Materials Science, Duke University, Durham, NC 27708, USA. [9]Chicago Eczema Center, Chicago, IL 60654, USA. [10]Department of Pediatrics (Allergy and Immunology), Ann & Robert H. Lurie Children's Hospital, Chicago, IL 60611, USA. [11]Department of Pediatrics (Dermatology), Ann & Robert H. Lurie Children's Hospital, Chicago, IL 60611, USA. [12]Department of Biomedical Engineering, Northwestern University, Evanston, IL 60208, USA. [13]Department of Materials Science and Engineering, Northwestern University, Evanston, IL 60208, USA. [14]Department of Neurological Surgery, Feinberg School of Medicine, Northwestern University, Chicago, IL 60611, USA.
*Corresponding author. Email: jrogers@northwestern.edu (J.A.R.); stevexu@northwestern.edu (S.X.)
†These authors contributed equally to this work.

the sensor data and an associated understanding of the physics of the processes of scratching, signal features that uniquely characterize scratching activities can be extracted to train and validate a machine learning (ML)–based algorithm for scratch detection.

The content begins with descriptions of systematic investigations of the mechanics of scratching in a set of controlled experiments, followed by validation of the combined use of sensor signals with data analytics approaches in two human subject studies: an algorithm development study (n = 10 healthy normal subjects) and a clinical validation study (n = 11 predominately pediatric patients with AD). The training study involves collection of scratching and nonscratching data in well-defined conditions to develop an ML algorithm to categorize and quantify scratching behavior. The clinical study validates the sensor and the algorithm in a cohort of patients with moderate to severe AD. The performance compares well to a gold standard defined by direct visual observation via manually labeled infrared (IR) camera recordings.
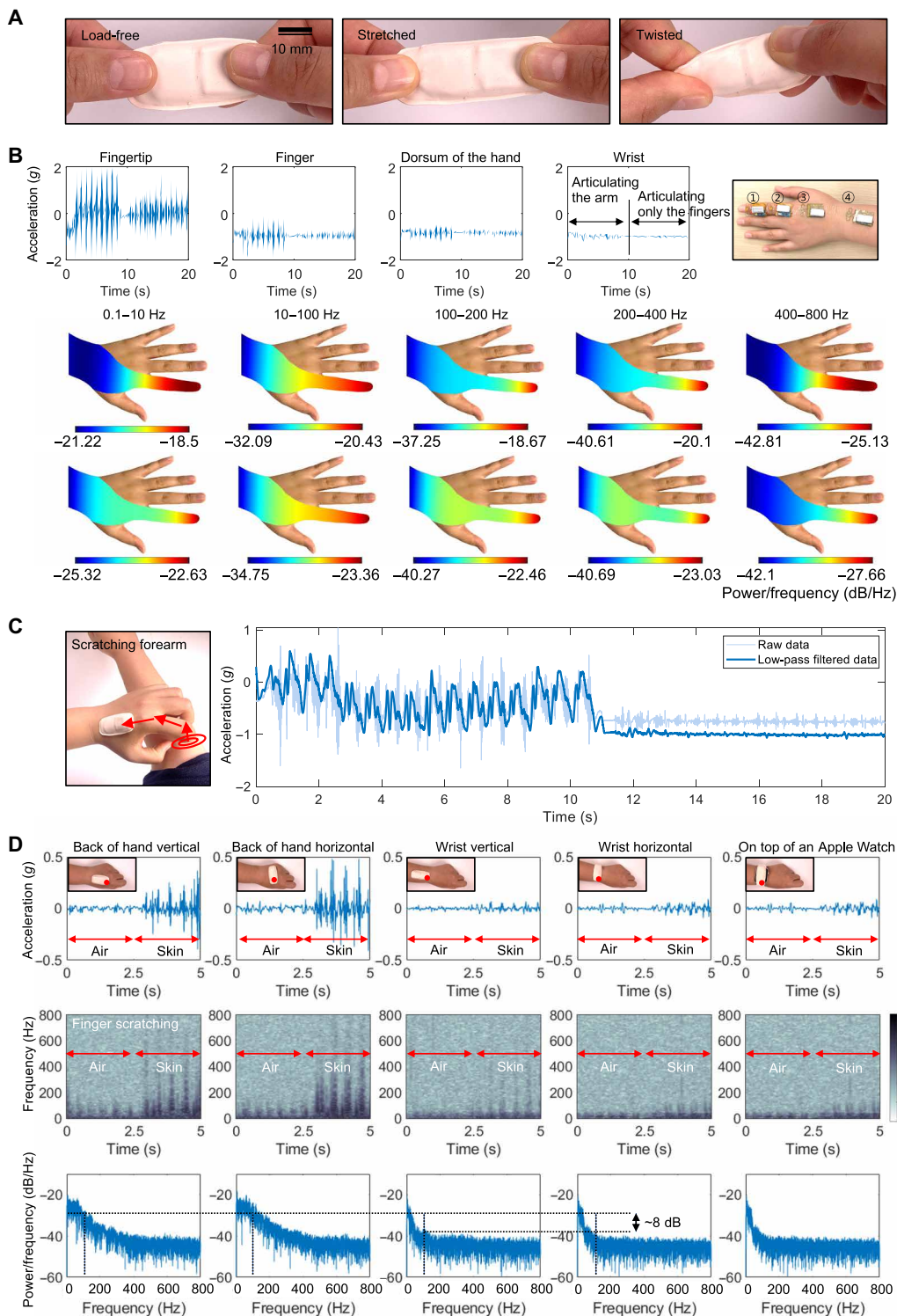
RESULTS

Sensor operation and performance

The ADAM sensor is a small, soft, stretchable wireless device that mounts using a thin adhesive onto the curvilinear surface of the dorsal hand, with an ability to be positioned across a range of possible anatomical locations and with various orientations (Fig. 1A). As described in detail elsewhere, the sensor features a Bluetooth Low Energy radio, electronics, and a rechargeable battery at one end of the device and, at the other, a millimeter-scale, three-axis accelerometer with a sampling rate of 1600 Hz at a resolution of 16 bits and a dynamic range of ±2g, where g is the gravitational acceleration, 9.8 m/s^2 (16). Studies reported here focus on accelerations measured in a direction perpendicular to the surface of the skin. A thin, low modulus silicone elastomer forms a skin-compatible encapsulating package, as a watertight enclosure for a flexible printed circuit board (fPCB) that supports and interconnects these various components. The fPCB adopts a patterned layout that includes a collection of freely deformable, serpentine interconnects specially designed by use of computational modeling to mechanically decouple the accelerometer from the other components of the device. This design optimizes the ability of the device to record subtle movements at the skin surface, accurately and without constraint. Figure 1A presents images that highlight these "soft" mechanical attributes and the overall "patch" form factor.

Scratching activities generate two types of signals. The first corresponds to gross movements of the hand, for which characteristic frequencies are in the range of a few hertz or less. The second, overlapping in time with the first, arises from subtle vibratory impulses generated by motions of the fingertips and fingernails against a contacting surface. Here, frequencies extend into the range of a few hundred hertz and exhibit amplitudes that decay rapidly with position along the fingers and into the hand, where they eventually pass through the wrist and to the arm. Detailed multi-accelerometer measurements reported elsewhere hint at the physics of this second type of signal, although in the context of haptic interfaces (17). Similar direct measurements of accelerations at different locations from the fingertip to the wrist reveal the spectral characteristics of signal attenuation for the scratching behaviors studied here (Fig. 1B). Measurements involve wireless accelerometers located at the fingertip, the midpoint along the finger, the dorsum of the hand, and the wrist, as in the right frame in Fig. 1B (fig. S1A). The tests capture signals due

to scratching the surface of the skin by articulating the arm for 10 s and then scratching with only the finger for 10 s. The peak-to-peak amplitudes of accelerations at the surface of the skin decay with increasing distance along the finger and the surface of the hand to the wrist. The color bars in spatial patterns in Fig. 1B represent the mean power of the signal per frequency in each frequency band. The left and right numbers indicate the maximum and minimum values for each of the respective plots. For scratching by articulating the arm (the first row of the spatial pattern), the power decays from 20 dB/Hz at the fingertips, the point of initiation, to −32 dB/Hz at the wrist in the frequency range from 10 to 100 Hz. The attenuation increases sharply with frequency. Specifically, the signal decreases from −19 to −37 dB/Hz and −20 to −41 dB/Hz for frequency ranges of 100 to 200 Hz and 200 to 400 Hz, respectively. For both types of scratching, the approximate power dissipation from the fingertip to the wrist is 2 dB in the frequency range from 0.1 to 10 Hz, 12 dB in the range from 10 to 100 Hz, 20 dB in the range from 100 to 200 Hz and 200 to 400 Hz, and 16 dB in the range from 400 to 800 Hz. As expected, the attenuation at low frequencies is much lower than that at high frequencies. The power of the signal decreases exponentially with increasing propagation distance and with increasing frequency. A decay constant of η, a frequency-dependent characteristic of viscoelastic material obtained by fitting the experimental data, characterizes the behavior. In the experiment in Fig. 1B, η has a value of 0.38 for the case of vibrations that propagate from the fingertip to the dorsum of the hand (fig. S1C). These results suggest that critically important high-frequency information associated with scratching can be captured most effectively by (i) locating the sensor in close proximity to the fingertips and (ii) operating the accelerometer in a high bandwidth mode.

Existing scratch sensors use relatively low bandwidth accelerometers (typically 50 Hz or less) and associated electronics in rigid enclosures that couple to the wrist using bands (18–20). Although these devices can capture the first type of signal, they cannot record the second with adequate fidelity, consistent with the results of the propagation studies described above. This limitation follows from (i) a nonideal mounting location, at the wrist, where high-frequency signals have small amplitudes; (ii) measurement bandwidths that are insufficient to capture high-frequency information; and (iii) a loose mechanical coupling to the skin. The ADAM sensor, by contrast, quantifies both types of signals with high precision because of its proximity to the fingertips/fingernails, its high bandwidth operation, and its intimate mechanical interface to the skin. The result is a superior capability for identifying scratching events with high specificity and sensitivity. Figure 1C presents representative data generated by scratching the arm with the pointer finger of the opposing hand, as illustrated in the left frame. The motions involve (i) translating the hand, wrist, and arm with the finger in a fixed position for 10 s and then (ii) articulating the fingers only for 10 s. The two types of signals discussed previously can be identified by passing the data through low-pass and high-pass filters with cutoffs at 2 Hz. The first type represents the movement of the hand and can serve as the basis for detecting scratch events but, when used in isolation, without the second, can produce false-positive results from hand waving, essential tremor, and other sources of motions unrelated to scratching (18). The critical high-frequency information originates at the fingertips and transmits through the soft tissue and metatarsal bones to the location of the sensor, as described above and illustrated by red arrows in Fig. 1C. These features, which are time-synchronized with

Fig. 1. Overview of the ADAM sensor and signal outputs. (**A**) Image of the ADAM sensor in various states of deformation to highlight its soft, flexible construction. (**B**) Results of measurements of spatiotemporal patterns of motions and vibratory signatures of scratching, with a focus on five different frequency ranges. The experiments involve four accelerometers, one each located at the fingertip, finger, dorsum of the hand, and the wrist, during scratching by articulating the arm for 10 s and then scratching with only the finger for the last 10 s. The color bars represent the mean power of the signal per frequency in each frequency band. The left and right values indicate the minimum and maximum of the respective plot. Upper spatial patterns are from scratching by articulating the arm, and lower spatial patterns are from scratching with only the finger. (**C**) Path of scratching and resulting data (raw and filtered) data from the sensor. (**D**) Comparison of signals captured during scratching with five different mounting locations as time series and spectrogram plots. The dorsal hand is an effective mounting location for capturing scratching with only the fingers. Photo credit: Keum San Chun, University of Texas at Austin.

sliding of the fingertips and, in some cases, with motions of the hand, include energy across a broad range of frequencies (*21*). Although transmission through the body leads to attenuation that increases with frequency, particularly above 100 Hz (*22*), the associated complex waveforms are well preserved at distances of several centimeters (*17*, *23*).

The effects appear in characteristics of signals captured by the ADAM sensor placed at the dorsum of the hand and the wrist with horizontal and vertical orientations, and on top of a smartwatch (Apple Watch Series 4) tightly strapped to the wrist. Measurements involve first articulating the fingers with a scratching motion in the air for 2.5 s and then scratching the surface of the skin with the same motions for 2.5 s. Time series and spectrogram representations of the data in Fig. 1D show the high-frequency content associated with scratching the skin versus the air, in all cases. The spectrograms in Fig. 1D use a Hamming window with a frame size of 0.2 s and an overlapping duration of 0.19 s. These identifying features extend in frequency up to ~600 Hz, with some energy even up to 800 Hz, but with the most notable amplitudes in the range of up to 200 Hz. The results show only modest dependence on mounting location provided that the sensor is on the radial half of the dorsum of the hand. Orientation has little effect, simply because the accelerometer resides at the front of the ADAM device, as described previously and highlighted with a red dot in Fig. 1D. On the wrist, the high-frequency components appear only as weak features in the data, consistent with the previous attenuation results in Fig. 1B (fig. S1B). As mounted on the back of the hand where the signal strength is high, the soft physical properties of the device and the medical-grade, hypoallergenic, single-use adhesive support a robust yet comfortable and nonirritating interface compatible with many daily activities involving various hand and wrist motions (fig. S2).

Characterizing scratching

Scratching varies with individuals in response to both internal and external stimuli. Scratching can occur via articulation of the elbow, wrist, or fingers alone, as examined in the tests summarized above and in Fig. 1. In the mode that engages the entire upper arm, the fingers press against the surface and remain stationary, where movement of the forearm articulating at the elbow joint drives the scratching action. The finger scratching mode, on the other hand, involves mainly local movements of the fingers, without substantial forearm or wrist motion. Mixtures of these two modes are also possible. Depending on the location of scratching and the degree of itch, individuals interchangeably use different modes of scratching. For any type of scratching action, the intensity is also important. The force at the point of contact determines the friction between the fingertips and the surface, and can therefore be considered, along with frequency and speed, as a metric of the intensity of a scratching event. These factors contribute to differences in signals captured by the ADAM sensor. Systematic studies for various types of scratching activities are important precursors to the development of a generalizable scratch detection algorithm.

Control experiments for this purpose involve two body locations: dorsum of the hand (DH) and forearm (FA) (Fig. 2A), with the ADAM sensor on the left dorsal hand (Fig. 1D), with two locations scratched with the left hand by articulating the arm only and then the fingers only, at a high intensity followed by a low intensity. The blue solid line in Fig. 2A shows the time-dependent *z*-axis acceleration from the vibrational motions at the surface of the skin in a direction

orthogonal to the plane of the scratching motion. The peak-to-peak amplitudes are larger for high-intensity scratching than for low intensity, as expected. For instance, the amplitude of the signal for high-intensity FA scratching (45 to 55 s) is 1.5*g*, which is more than two times greater than that for the low-intensity case (65 to 75 s; 0.7*g*). For an otherwise similar scenario but by articulating only the fingers, the amplitudes of the high (125 to 135 s) and low (145 to 155 s) are 0.4*g* and 0.15*g*, respectively. Results for articulating the arm only (0 to 80 s) show amplitudes larger than those for the fingers (80 to 160 s) for each scratch location (DH, FA) and intensity level. The frequency domain content associated with these signals appears as spectrograms (Fig. 2A), plotted using the same window and frame size as Fig. 1C. In all cases, even at low intensity with fingers only, the high-frequency content, especially up to ~200 Hz, features prominently in the data. The normalized spectral characteristics of high- and low-intensity scratching events are largely the same, as in fig. S3.

Scratching activities on five different body locations (head, arm, abdomen, knee, and leg) and nonscratching activities (simulating scratching with fingers in the air, hand waving, texting using a cell phone, typing on a keyboard, and clicking a mouse) highlight the key differences in the signals and the importance of the high-frequency content in a range of practical scenarios (Fig. 2, B to D). Specifically, the data for all scratching activities have substantial energy above 200 Hz, independent of location (Fig. 2C). As expected based on previous discussions, the spectrograms of most nonscratching activities exhibit minimal energy in this frequency range (Fig. 2E). Text messaging, typing on a keyboard, and clicking a computer mouse represent exceptions, where the data show contributions between 100 and 200 Hz because of impulse components of these motions. Nevertheless, the spectral content across all frequencies, the temporal characteristics of the signals, and the low-frequency information can help to distinguish these activities from scratching. In other words, the high-frequency components of the recorded signals are important indicators of scratch, but they must be used together with other features of the data to achieve high selectivity against a certain, relatively small, set of confounding activities. Furthermore, measurement of nocturnal scratching mitigates the need to differentiate from awake activities such as typing or text messaging. As an illustration, movies S1 and S2 highlight a form of manual signal analysis that involves translation of the raw data into an audio file that can be interpreted in a manner analogous to that of sounds from a stethoscope. Various activities such as those described above can be immediately distinguished from scratching in this manner. These collective considerations motivate an ML approach to data analysis, in which various features, including but not limited to high-frequency content, contribute to a classification scheme for unmatched combined levels of sensitivity and specificity.

Hand-mounted versus wrist-mounted scratch sensing

As mentioned previously, scratching behavior can occur via articulation of the fingers, wrist, elbow, or shoulder. A key limitation of wrist-mounted sensors is the inability to measure motion associated with finger-only scratching and in difficulties in distinguishing between hand waving and scratching. Control studies reveal the performance of the ADAM system compared to a wrist-bound system (Apple Watch Series 4) with an embedded mobile application published previously to quantify scratch (*18*). The comparison focuses on two modes of scratching: one by articulating the arm and the other by only the fingers. To examine scratching by articulating the
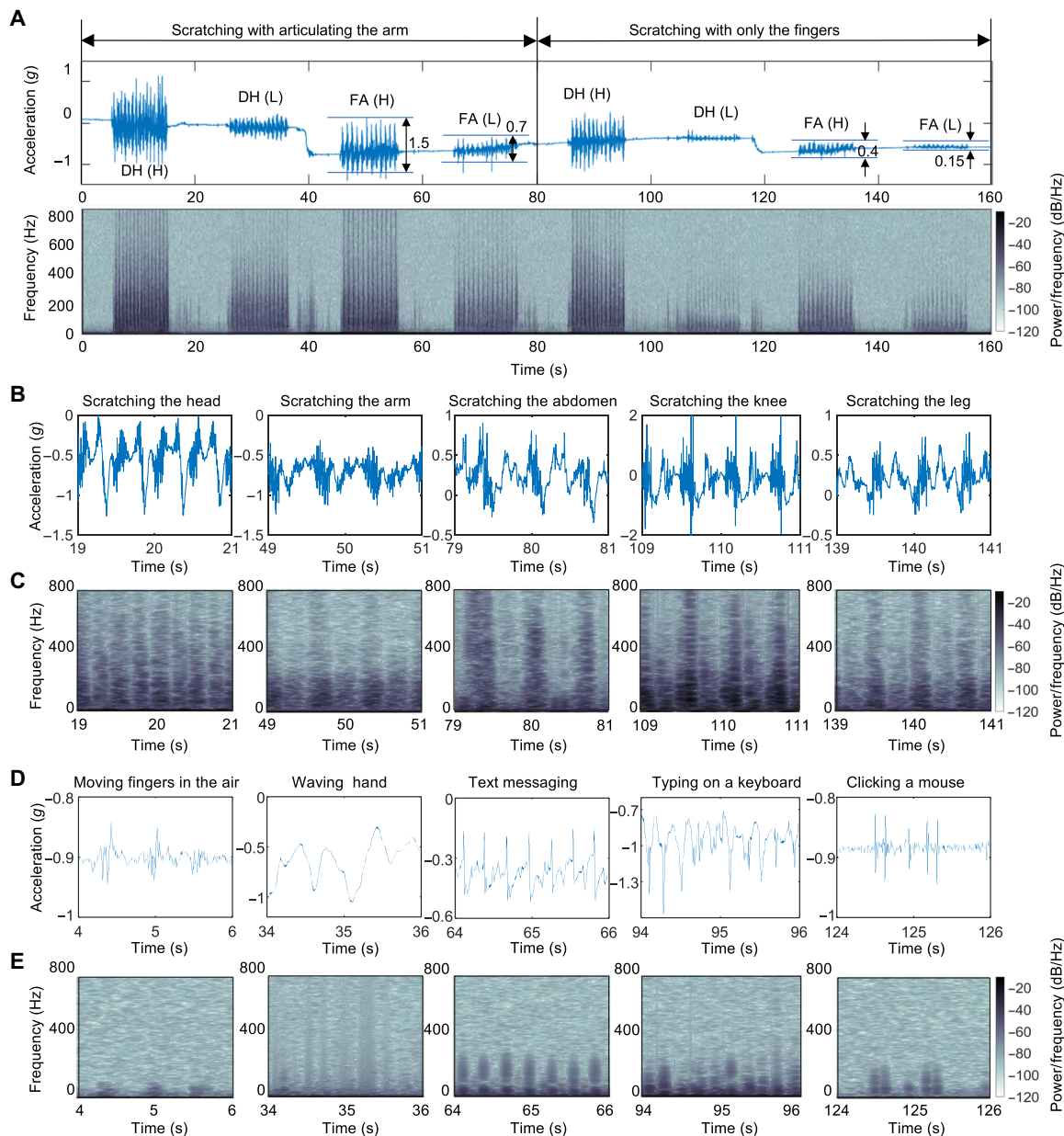
Fig. 2. Representative data collected by the ADAM sensor. (**A**) Sample time series data and spectrogram corresponding to scratching activities. Two modes of scratching are conducted on two body parts, dorsum of the hand (DH) and forearm (FA), in two intensities of high (H) and low (L). (**B**) Time series data of scratching activity in which five different body parts include head (hairy skin), arm (normal skin), abdomen (soft skin), knee (bony prominence skin), and the leg (hard skin). (**C**) Spectrogram of each time series data in (B). The signals due to the scratching activities have the energy in the 0- to 800-Hz frequency range. (**D**) Time series data in a wide range of nonscratching activities including simulated moving fingers in the air, waving hand, text messaging, typing on the keyboard, and clicking the mouse. (**E**) Spectrogram of each time series data in (D). The signals due to nonscratching activities have energy mainly in the range less than 200 Hz.

arm, the set of activities includes knee scratching, hand waving, abdomen scratching, and head scratching in a sequence for 10 s separated by 5 s of pause. The blue line represents the time series of the raw data, and the red line indicates the final binary classification results where 0 and 1 correspond to nonscratch and scratch, respectively. The black dashed line highlights the classification results from Itch Tracker (*18*), with 0 representing nonscratch activity and 0.5 representing scratch activity (Fig. 3A). While Itch Tracker captures knee scratching and abdomen scratching, it misclassifies hand waving

and head scratching. The former likely follows from a heavy reliance on hand motions for scratch detection. The latter probably occurs because head scratching involves only short range of motions of the wrist. The ADAM system performs well across all of these and other scenarios.

Additional experiments explore scratching modes that involve only finger articulation (Fig. 3B). As with the previous comparison using the arm for scratching, the four activities include knee scratching, simulating scratching in the air, abdomen scratching, and head
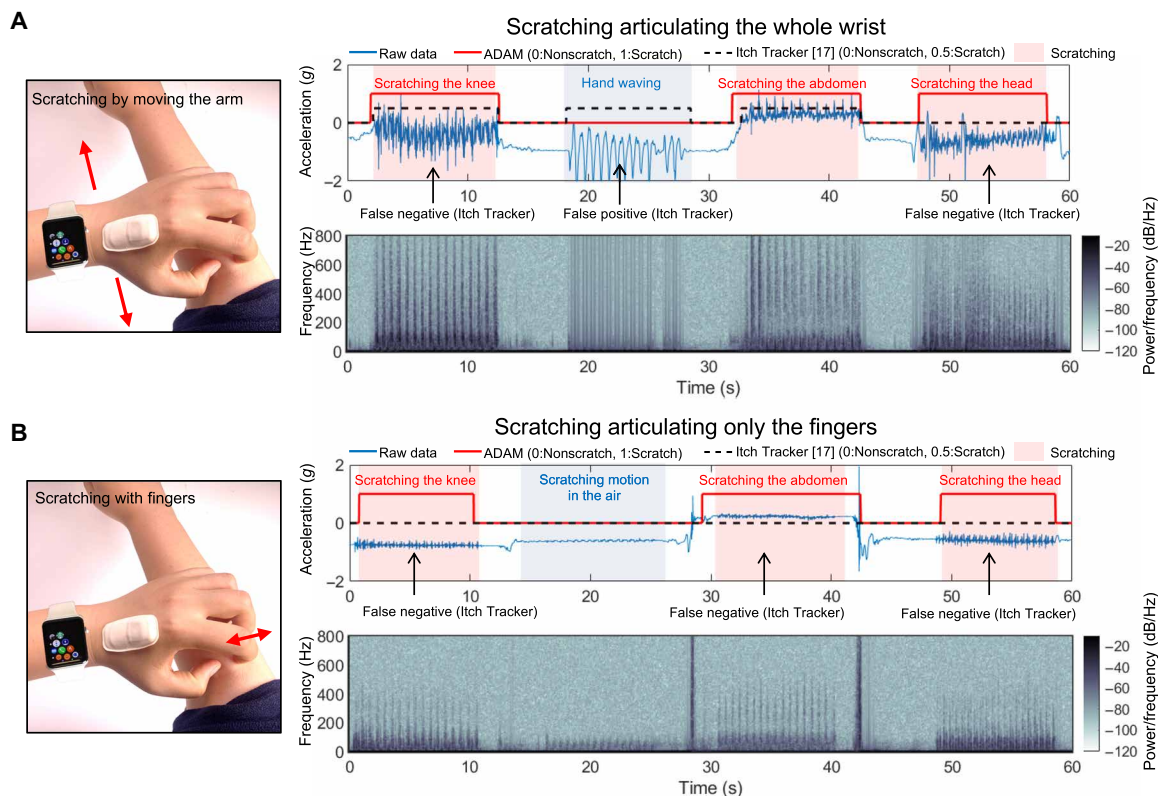
Fig. 3. Comparison of results obtained with an ADAM device on the hand and an Apple Watch with the Itch Tracker mobile application. Time series data and the spectrograms for various scratch (skin over the knee where there is a bony prominence, skin of the soft abdomen, and skin of complex surface head) and nonscratch activities (waving of hand and scratching motion in the air), with a 5-s pause between each activity. (**A**) Scratching by articulating the whole wrist. The Itch Tracker misclassifies hand waving as scratching and fails to detect scratching on the head. (**B**) Scratching by articulating only the fingers. The ADAM sensor reliably discriminates scratching the skin using only articulation of the fingers. The Itch Tracker cannot detect scratches using only fingers, given its mounting location on the wrist and its inability to record high-frequency information. Photo credit: Keum San Chun, University of Texas at Austin.

scratching. In these activities, the Itch Tracker does not detect any instances of scratching because of the lack of motion at the wrist. On the other hand, the ADAM system detects all instances, and it even accurately discriminates scratching from pseudo-scratching (Fig. 3B). Simulated scratching motion in the air provides an additional test of the ADAM system. The major difference between scratching and pseudo-scratching is the presence of frequency signatures above 100 Hz. While the ADAM system can capture frequencies as high as 800 Hz, most wrist-bound accelerometers such as the Apple Watch record frequencies only as high as 50 Hz (i.e., sampling rate of 100 Hz).

Validation

A random forest (RF) classifier allows automated detection of scratching activities. The training dataset contains scratching data from 15 different body locations and 6 types of nonscratching activities collected from 10 healthy normal subjects (Fig. 4, A and B). Results from leave-one-subject-out cross validation (LOSO-CV) applied to the training dataset form the basis for optimization of the RF classifier. In this process, a dataset from a single participant serves as the testing dataset, and the remaining data allow training. Iterating this process across all participants allows each participant's data to be used for testing once. The classification performance corresponds to averages from all iterations of the validation. LOSO-CV,

unlike *k*-fold CV approaches, prevents the classifier from learning a confounding relationship within a subject's dataset, with unrealistically high classification accuracy as reported by Saeb *et al.* (*24*). The RF model can be optimized by comparing the LOSO-CV outcomes. The parameters for this optimization include the number of trees in the RF classifier and the minimum required scratching (MRS) duration. The RF classifier is a meta estimator that fits a number of decision tree classifiers on various subsamples of a dataset. The choice of number of trees in the forest fairly balances the complexity of the model and its performance. Increasing the number increases the training time and the potential for overfitting. Thus, the optimization focuses on examining a different number of estimators, from 10 to 200 trees, and comparing the LOSO-CV performance. The classification performance reaches a maximum value at approximately 30 decision trees.

The MRS duration represents another optimization parameter. In the signal processing pipeline (Fig. 4C), the RF classifier returns predictions at a 1-s frame level. The outputs are subsequently clustered using density-based spatial clustering of applications with noise (DBSCAN). The purpose of DBSCAN is to join two scratching frames in close proximity as a single scratching event. MRS defines the minimum number of frames to form a cluster of scratching events. MRS values from 1.5 to 8.5 s yield different classification performance using LOSO-CV. The performance increases with increasing MRS
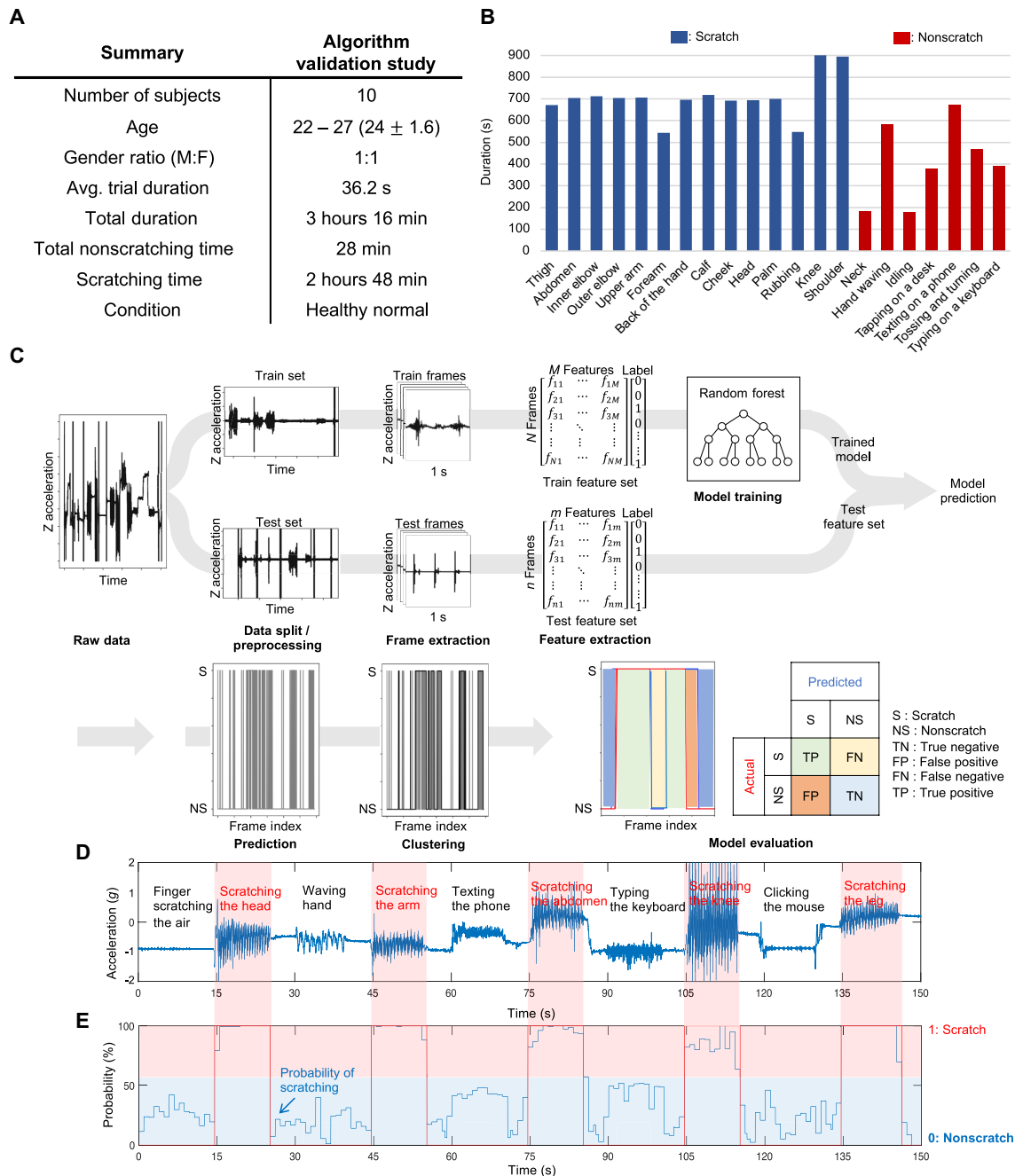
Fig. 4. Overview of datasets, signal processing pipelines, and validation methods. (**A**) The algorithm validation data were collected from 10 healthy normal volunteers. (**B**) Activities performed by each participant and their corresponding activity durations. (**C**) Preprocessing subsequently applied to the sets and segmented into frames by applying a sliding window of 1 s. From each frame, a set of features was extracted. With the feature set extracted from the training frames, a random forest (RF) classifier was trained. The trained classifier was applied to the test feature set for prediction and validation. (**D**) Time series data of our validation cohort with a single subject performing various scratching and nonscratching activities include 5-s pause periods between each activity. (**E**) The output probability for scratching is characterized as "1:scratch" if the probability exceeds 50%. The blue line is the probability of scratching for each frame, and the red line indicates the final classified result, where 0 and 1 correspond to nonscratch and scratch, respectively.

up to 4.5 s and then remains constant for values up to 8.5 s (fig. S5). Improved performance can be achieved with further increases in MRS but with diminishing returns. For the analysis below, MRS is defined as at least 4.5 s.

The optimized RF classifier is applied to data collected from a single subject performing various scratching (the head, the arm, the abdomen,

the knee, and the leg) and nonscratching activities (simulating scratching in the air, hand waving, texting on the phone, typing on a keyboard, and clicking a mouse). The corresponding time series z-axis acceleration data are in Fig. 4D, and the associated classification results are in Fig. 4E.

With the optimized RF classifier (30 trees) and MRS duration (4.5 s), LOSO-CV can be applied to the healthy, normal training

dataset to evaluate the performance. The results indicate an average precision of 94.4% and an average sensitivity of 87.8%. The overall accuracy is 89.1%. Table 1 summarizes the overall performance.

The RF classifier trained on the algorithm validation dataset can be used directly on the patient dataset from the clinical validation study. This evaluation reveals an average precision of 82.5% and an average sensitivity of 84.3% (Table 1). This is a notable improvement from most of the previous work on automated scratch detection (table S2). Movie S1 illustrates a pediatric patient with AD wearing our sensor while scratching. The IR camera demonstrates scratching behavior that is subsequently coded. Note that scratching intensity varied from the beginning to the end of the scratching event, as reflected in the IR video as a rigorous scratching motion and in the ADAM sensor data as a large amplitude signal.

DISCUSSION

Itch is one of the most common clinical symptoms of disease affecting both children and adults across dermatological and non-dermatological conditions. Despite the prevalence of itch and its profound impact on quality of life, accurate and convenient methods to fully characterize itch and itch-associated scratching behavior do not exist for use either in clinical practice or in the home. Technologies that are able to precisely and reliably assess itch will serve as critically important tools for drug development, quantification of itch severity, and monitoring of treatment response.

Video surveillance with manual labeling is considered the gold standard as it allows accurate identification and quantification of scratching behavior. This technique is, however, both labor intensive and difficult to automate. These drawbacks, together with a range of associated privacy concerns, render this approach impractical for routine use (25). Wrist-bound accelerometers represent popular tools for collecting objective data in a wide range of clinical studies with adequate sensitivity but poor specificity (26). Some studies cast doubt on the clinical applicability of actigraphy data, due partly to poor correlations with AD disease severity and quality of life. Emerging approaches that exploit smartwatches in combination with ML algorithms outperform simple actigraphy-based methods (15).

However, these systems cannot detect finger-only scratching, and hand waving and related gestures are misclassified as scratching (Fig. 3). As an example of an unusual approach, Noro *et al.* developed a wristwatch-shaped sound detection system to measure scratching sounds through bone conduction (27). Although the accuracy can approach that of video analysis, microphone-based systems involve substantial privacy concerns, and their performance degrades with ambient noise.

The ADAM sensor and scratch algorithm reported here offer key advantages. First, the soft flexible nature of the devices allows them to conform to the hands of a wide range of patients, as validated on children as young as four to full-grown adults, where they remain softly but robustly adhered during a range of natural activities. This aspect of the ADAM sensor is particularly suitable for investigating long-term changes of pathological symptoms such as nocturnal pruritus. For instance, with ADAM sensor, we observed notable nightly variation in scratching in a subset of patients with AD who wore the sensor for an extended period of time (10 days or more). This implies that averaging scratch duration over a more sustained period of wear time is likely necessary rather than a single night (fig. S6). Second, the system captures both low- and high-frequency acousto-mechanic signatures associated with scratching. The ability to effectively capture "acoustic" signals generated by scratching, without the need for a microphone, ensures both patient privacy and performance in noisy environments with no deterioration in signal quality. Last, the system is waterproof, wireless, and rechargeable with 7 days of continuous operating efficiency. These features enhance the ease of use and reduce user burden in both clinical practice and research trials.

Several limitations must, however, be noted. First, the clinical validation studies reported here involve only $n = 11$ subjects, due partly to the high degree of manual labor required to evaluate IR camera footage. However, this represents one of the largest cohorts studied to date for scratching (table S2) given the time intensity of data labeling—and the first, to our knowledge, in a pediatric cohort where AD is most prevalent. Future efforts will include additional patient testing across a wide range of clinical conditions beyond AD, with patients even younger than those studied here. Nevertheless, the findings in this paper show that training on healthy normal subjects performing voluntary scratching activities can apply successfully to clinical data in a predominately pediatric population affected by AD. The outcomes, which follow from the unique distinguishing features in the data from the ADAM device, suggest remarkable levels of generalizability and opportunities to develop algorithms tailored to patient populations. Second, all measurements of scratching described here focus only on the dominant hand of the subjects. An additional sensor on the contralateral hand will provide additional information. Previous studies indicate no difference in scratching tendency for dominant or nondominant hands, suggesting that a single sensor does not lead to major information loss (26). Third, the clinical validation studies focus on AD patients only. Future work will examine applications to other medical conditions where itch is a predominant symptom. As a final comment, AD affects sleep to a substantial degree, particularly in children (5). The ADAM sensor itself can recapitulate sleep stages against polysomnography, the gold standard method to assess sleep (16). An immediate opportunity for additional work is in using two time-synchronized ADAM sensors to assess both scratch and sleep concomitantly.

This paper reports a unique hand-mounted, soft, flexible sensor and a data analysis pipeline for detecting scratching activities using

Table 1. Classification results. The algorithm validation study used LOSO-CV and an RF classifier. The clinical study included manually labeled datasets from all 46 nights with the ADAM sensor mounted on the dominant hand of each subject. The algorithm developed in the algorithm validation study was then deployed on the raw data from the clinical study. The overall accuracy in the clinical study was 99.0% with a sensitivity of 84.3% and a specificity of 99.3%. The accuracy of the algorithm validation was lower (89.1%) due to a higher number of confounding activities to train the algorithm (e.g., typing and texting) that are not seen in nocturnal settings.

	Algorithm validation study ($n = 10$)	Clinical study ($n = 11$)
Sensitivity	87.8%	84.3%
Specificity	88.1%	99.3%
Accuracy	89.1%	99.0%
Precision	94.4%	82.5%
F1 score	89.8%	82.9%
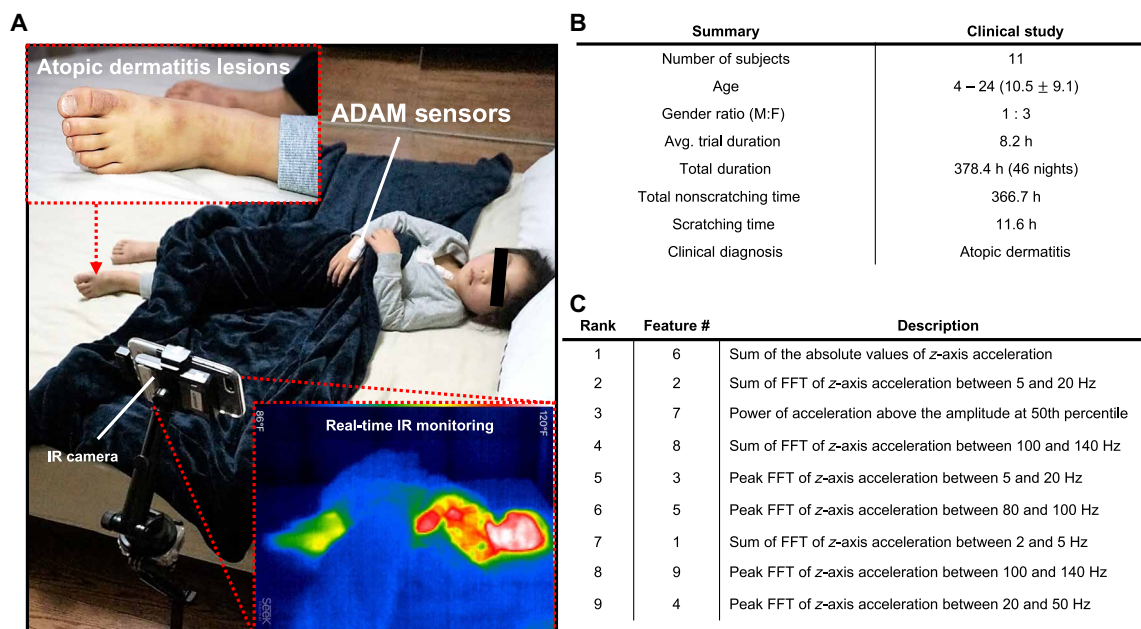
Fig. 5. Validation data summary and extracted features. (**A**) The clinical validation was conducted in a natural home environment with predominately pediatric AD patients (median age, 10.5 years). An IR camera was used to record scratching behavior of human subjects and was manually graded by two clinical research staff members. (**B**) A total of 11 AD patients were recruited, generating a total of 46 nights of data in a pooled analysis. (**C**) A total of nine features were used for LOSO-CV. They are listed in order of their feature importance obtained from the RF classifier. Photo credit: Jan-Kai Chang, Wearifi Inc.

high-frequency three-axis accelerometer data that capture features associated with both motions and acousto-mechanical signals. From a training study conducted with 10 participants, detection of scratching activity has an overall accuracy of 89.1% with a sensitivity of 87.8% and a specificity of 88.1%. Clinical validation studies with 11 participants reveal an overall accuracy of 99.0% with a sensitivity of 84.3% and a specificity of 99.3%, even in naturalistic home environments. The small, flexible characteristics of the ADAM sensor and its reusability are key features that allow deployment in both research and clinical care settings. The promising results presented herein suggest utility as a powerful methodological tool with the potential to accurately capture scratching activities not only in patients with AD but also in patients suffering from any generalized pruritic disorder.

MATERIALS AND METHODS
Sensor operation and performance
The high-bandwidth three-axis accelerometer in the ADAM sensor captures skin-conducted vibrations and overall motions associated with scratching behavior. The ADAM sensor captured acceleration along the x and y axes at a rate of 200 Hz, and the z-axis acceleration was sampled at a rate of 1600 Hz. The x and y axes were oriented parallel to the skin surface, whereas the z axis was perpendicular to the skin surface. The high-frequency content attenuates as the disturbance propagates from its point of origination at the fingertips, along the fingers and to the hand and, ultimately, the wrist. The location of the sensor determines, in this manner, the nature and quality of the signals. Experiments to examine the dependence on mounting location involved a set of scratching activities (10 s of scratching by articulating the arm and 10 s of finger scratching were performed, each separated by 5 s of idle state) repeated with different placements and orientations of the sensor. On the basis of signal

quality and considerations in wearability, the ideal location is between the metacarpal first and second digit.

Characterizing scratching
Scratching activity refers to a collection of movements that involve pressing the fingertips against a surface and moving the fingertips, causing friction. The friction between the fingertips and the surface generates vibrations that propagate along the surface. Scratching signals can differ from one another as the vibrations between the fingertips and the surface depend on many factors including, but not limited to, scratching speed, scratching intensity, and mode of scratching (e.g., only using fingers or moving the entire arm). Experiments to explore the time and frequency characteristics associated with these factors included scratching the head, arm, knee, and leg and various nonscratching events such as hand waving, texting messaging, typing on a keyboard, clicking a computer mouse, and moving fingers in the air. In fig. S8, we show the predicted labeling for these confounders. The system performs best (>0.7) in distinguishing finger scratching in the air, hand waving, and clicking a mouse.

Study design
Two datasets were collected independently for model development (training) and validation. The training data, collected in a controlled environment, served as the basis for developing an ML algorithm for detecting scratching behavior. The validation dataset was collected from AD patients in their naturalistic home environments (Fig. 5A). These data were used to evaluate the generalizability of the ML algorithm. A total of 10 participants (5 males and 5 females) participated in the training data collection, with ages between 22 and 27 years old (24 ± 1.6). The validation data collection involved 11 patients (2 males and 9 females) with moderate to severe AD, with ages between 4 and 24 years old (10 ± 9.1).

The training data were collected from healthy normal individuals. The participants performed a set of scratching and nonscratching activities, each lasting for approximately 35 s. The scratching activities were performed over clothes on the thigh, abdomen, upper arm, calf, knee, and shoulder. Each participant then scratched the skin on the inner elbow, outer elbow, forearm, dorsum of the hand, cheek, head, palm, and neck. Rubbing was considered as a variation of scratching for purposes of these studies. These scratching activities were each performed twice, once while sitting on a chair and another while lying on a bed. For nonscratching activities, the participants waved their hands, sent text messages to friends, tossed and turned on a bed, sat idly, typed on a keyboard, and tapped on a desk. These nonscratching activities were included as potential confounders for actual scratching activities. Data captured from each activity were saved as separate files with descriptions of each activity, annotated as scratching or nonscratching.

The validation data were collected from AD patients in their home environment as part of a clinical study approved by the Institutional Review Board of Northwestern University and the Ann & Robert H. Lurie Children's Hospital (IRB 2018-2111). The inclusion criteria included any patient with mild to severe AD older than 2 years of age. Patients with active skin or systemic infection were excluded. Eligible participants were screened in the Northwestern Medicine Department of Dermatology clinics and other pediatric dermatology or allergy providers at Ann & Robert H. Lurie Children's Hospital. A total of 11 subjects were recruited for the validation data collection (Fig. 5B). For subjects who agreed to participate, a written consent form was first obtained. Clinical research personnel then completed an eczema area and severity index (EASI) assessment and investigator's global assessment (IGA) of each patient. After completion of EASI and IGA, each patient was trained on procedures for operating the ADAM sensor and collecting nocturnal data. During the training, the patients were provided with specific instructions about placing the ADAM sensor on the dorsum of the hand and securing it with a Coban self-adherent wrap. The patients then received a package containing the ADAM sensor, an iPhone, tripod, IR camera, charging equipment, instruction manual, and checklist. During the collection of validation data, each patient's nocturnal activities were recorded with an IR video camera. The ADAM sensor data were compared with the IR video data for data annotation. This annotation process was performed using a data visualization and annotation software (fig. S7) (28).

Validation

CV has been frequently used in many ML applications as it provides a simple and effective way of evaluating classification models (29). In CV, the number of folds is a parameter that specifies how data are split into training and testing sets. We used LOSO-CV to assess the performance of the model as it is known to perform well on real world data (30). After the model was optimized on the training dataset, we evaluated the performance of the model in cross-study validation where the trained model is applied to the clinical validation dataset collected in the home environment. The cross-study validation examines the generalizability of the algorithm as the training dataset, and the validation datasets were independently collected.

Data analysis

In this study, we focused on the z-axis acceleration data because the critical information about scratching pertains mainly to out-of-plane skin vibrations. The raw z-axis acceleration data were processed

through a signal processing pipeline as illustrated in Fig. 4C. In the first step, the raw data were passed to the preprocessing block to remove the baseline wander of the data caused by slow, large-scale motions. Thus, low-frequency baseline wandering is removed from the signal processing pipeline to allow greater emphasis of the more specific high-frequency signal components. Subsequently, the preprocessed z-axis acceleration data were segmented into frames by sliding a 1-s window with an overlap. For training, 90% overlap was used, and for testing, 50% overlap was used. Each frame was assigned a label of 0, indicating nonscratching activity if less than 50% of the data points in a given frame were labeled as scratching activity. Otherwise, the frame was assigned a label of 1, indicating scratching activity. From each frame, a set of features shown in Fig. 5C were extracted, and these features were used to train and validate an RF classifier.

Comparison with wrist-mounted scratch detection

An Apple Watch Series 4 was placed on the left wrist, and an ADAM sensor was placed on the dorsum of the left hand. This configuration was intended to synchronize the data collected from both devices for purposes of direct comparisons. Tight coupling of the Apple Watch Series 4 to the dorsum of the left hand was verified by ensuring that the smartwatch could not be rotated freely around the wrist. Similarly, the tight attachment of the ADAM sensor was verified by flexing and extending the left hand, ensuring that the ADAM sensor stayed in contact with the skin. After placing the sensors, a set of scratching and nonscratching activities was performed in sequence (Fig. 3). One of the scratching activities involved the movement of the entire arm/hand, while the other scratching method only used fingers. In addition, a simple hand waving activity was performed to determine an ability to differentiate simple hand motion from scratching activity.

Interrater reliability

For the clinical validation study, IR video served as the ground truth. Annotation was performed on the basis of visual inspection and human judgment. To evaluate the interrater reliability of annotation, annotation was performed again by a different researcher (E.A. or H.C.) on data from three patients within the validation dataset. The labels from two researchers were compared using Fleiss' kappa (31). The average of kappa was 0.88. According to Landis and Koch (32), annotations with such a kappa value are considered to be in good agreement.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/7/18/eabf9405/DC1

REFERENCES AND NOTES

1. G. Yosipovitch, M. W. Greaves, M. Schmelz, Itch. *Lancet* **361**, 690–694 (2003).
2. A. Bathe, E. Weisshaar, U. Matterne, Chronic pruritus-more than a symptom: A qualitative investigation into patients' subjective illness perceptions. *J. Adv. Nurs.* **69**, 316–326 (2013).
3. M. Shive, E. Linos, T. Berger, M. Wehner, M.-M. Chren, Itch as a patient-reported symptom in ambulatory care visits in the United States. *J. Am. Acad. Dermatol.* **69**, 550–556 (2013).
4. T. E. Shaw, G. P. Currie, C. W. Koudelka, E. L. Simpson, Eczema prevalence in the United States: Data from the 2003 National Survey of Children's Health. *J. Invest. Dermatol.* **131**, 67–73 (2011).
5. S. L. Chamlin, C. L. Mattson, I. J. Frieden, M. L. Williams, A. J. Mancini, D. Cella, M.-M. Chren, The price of Pruritus. *Arch. Pediatr. Adolesc. Med.* **159**, 745–750 (2005).
6. K. Hon, T. Leung, K. Wong, C. Chow, A. Chuh, P. Ng, Does age or gender influence quality of life in children with atopic dermatitis? *Clin. Exp. Dermatol.* **33**, 705–709 (2008).

7. D. Camfferman, J. D. Kennedy, M. Gold, A. J. Martin, K. Lushington, Eczema and sleep and its relationship to daytime functioning in children. *Sleep Med. Rev.* **14**, 359–369 (2010).

8. R. Sack, J. Hanifin, Scratching below the surface of sleep and itch. *Sleep Med. Rev.* **14**, 349–350 (2010).

9. A. B. Fishbein, O. Vitaterna, I. M. Haugh, A. A. Bavishi, P. C. Zee, F. W. Turek, S. H. Sheldon, J. I. Silverberg, A. S. Paller, Nocturnal eczema: Review of sleep and circadian rhythms in children with atopic dermatitis and future research directions. *J. Allergy Clin. Immunol.* **136**, 1170–1177 (2015).

10. A. B. Fishbein, K. Mueller, L. Kruse, P. Boor, S. Sheldon, P. Zee, A. S. Paller, Sleep disturbance in children with moderate/severe atopic dermatitis: A case-control study. *J. Am. Acad. Dermatol.* **78**, 336–341 (2018).

11. J. I. Silverberg, N. K. Garg, A. S. Paller, A. B. Fishbein, P. C. Zee, Sleep disturbances in adults with eczema are associated with impaired overall health: A US population-based study. *J. Invest. Dermatol.* **135**, 56–66 (2015).

12. M. K. LeBourgeois, D. C. Dean, S. C. Deoni, M. Kohler, S. Kurth, A simple sleep EEG marker in childhood predicts brain myelin 3.5 years later. *NeuroImage* **199**, 342–350 (2019).

13. G. Kiebert, S. V. Sorensen, D. Revicki, S. C. Fagan, J. J. Doyle, J. Cohen, D. Fivenson, Atopic dermatitis is associated with a decrement in health-related quality of life. *Int. J. Dermatol.* **41**, 151–158 (2002).

14. C. S. Murray, J. L. Rees, Are subjective accounts of itch to be relied on? The lack of relation between visual analogue itch scores and actigraphic measures of scratch. *Acta Derm. Vener.* **91**, 18–23 (2011).

15. A. Moreau, P. Anderer, M. Ross, A. Cerny, T. H. Almazan, B. Peterson, Detection of nocturnal scratching movements in patients with atopic dermatitis using accelerometers and recurrent neural networks. *IEEE J. Biomed. Health Inform.* **22**, 1011–1018 (2017).

16. K. H. Lee, X. Ni, J. Y. Lee, H. Arafa, D. J. Pe, S. Xu, R. Avila, M. Irie, J. H. Lee, R. L. Easterlin, D. H. Kim, H. U. Chung, O. O. Olabisi, S. Getaneh, E. Chung, M. Hill, J. Bell, H. Jang, C. Liu, J. B. Park, J. Kim, S. B. Kim, S. Mehta, M. Pharr, A. Tzavelis, J. T. Reeder, I. Huang, Y. Deng, Z. Xie, C. R. Davies, Y. Huang, J. A. Rogers, Mechano-acoustic sensing of physiological processes and body motions via a soft wireless device placed at the suprasternal notch. *Nat. Biomed. Eng.* **4**, 148–158 (2020).

17. Y. Shao, V. Hayward, Y. Visell, Spatial patterns of cutaneous vibration during whole-hand haptic interactions. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 4188–4193 (2016).

18. A. Ikoma, T. Ebata, L. Chantalat, K. Takemura, F. Mizzi, M. Poncet, D. Leclercq, Measurement of nocturnal scratching in patients with pruritus using a smartwatch: Initial clinical studies with the itch tracker app. *Acta Derm. Vener.* **99**, 268–273 (2019).

19. J. Lee, D. Cho, J. Kim, E. Im, J. Bak, K. H. Lee, K. H. Lee, J. Kim, Itchtector: A wearable-based mobile system for managing itching conditions, *Proc. CHI Conf. Hum. Factor. Comput. Syst.* 893–905 (2017).

20. K. Benjamin, K. Waterston, M. Russell, O. Schofield, B. Diffey, J. L. Rees, The development of an objective method for measuring scratch in children with atopic dermatitis suitable for clinical use. *J. Am. Acad. Dermatol.* **50**, 33–40 (2004).

21. M. Wiertlewski, J. Lozada, V. Hayward, The spatial spectrum of tangential skin displacement can encode tactual texture. *IEEE Trans. Robot.* **27**, 461–472 (2011).

22. M. Wiertlewski, V. Hayward, Mechanical behavior of the fingertip in the range of frequencies and displacements relevant to touch. *J. Biomech.* **45**, 1869–1874 (2012).

23. B. Delhaye, V. Hayward, P. Lefèvre, J.-L. Thonnard, Texture-induced vibrations in the forearm during tactile exploration. *Front. Behav. Neurosci.* **6**, 37 (2012).

24. S. Saeb, L. Lonini, A. Jayaraman, D. C. Mohr, K. P. Kording, The need to approximate the use-case in clinical machine learning. *Gigascience* **6**, 1–9 (2017).

25. T. Ebata, H. Aizawa, R. Kamide, An infrared video camera system to observe nocturnal scratching in atopic dermatitis patients. *J. Dermatol.* **23**, 153–155 (1996).

26. T. Ebata, S. Iwasaki, R. Kamide, M. Niimura, Use of a wrist activity monitor for the measurement of nocturnal scratching in patients with atopic dermatitis. *Br. J. Dermatol.* **144**, 305–309 (2001).

27. Y. Noro, Y. Omoto, K. Umeda, F. Tanaka, Y. Shiratsuka, T. Yamada, K. Isoda, K. Matsubara, K. Yamanaka, E. C. Gabazza, M. Nishikawa, H. Mizutani, Novel acoustic evaluation system for scratching behavior in itching dermatitis: Rapid and accurate analysis for nocturnal scratching of atopic dermatitis patients. *J. Dermatol.* **41**, 233–238 (2014).

28. O. Crasborn, H. Sloetjes, E. Auer, P. Wittenburg, Combining video and numeric data in the analysis of sign languages with the ELAN annotation software, in *2nd Workshop on the Representation and Processing of Sign Languages: Lexicographic Matters and Didactic Scenarios* (2006), pp. 82–87.

29. S. Arlot, A. Celisse, A survey of cross-validation procedures for model selection. *Stat. Surv.* **4**, 40–79 (2010).

30. R. Kohavi, A study of cross-validation and bootstrap for accuracy estimation and model selection. *IJCAI* **14**, 1137–1145 (1995).

31. J. L. Fleiss, J. Cohen, B. S. Everitt, Large sample standard errors of kappa and weighted kappa. *Psychol. Bull.* **72**, 323–327 (1969).

32. J. R. Landis, G. G. Koch, The measurement of observer agreement for categorical data. *Biometrics* **33**, 159–174 (1977).

Science Advances

A skin-conformable wireless sensor to objectively quantify symptoms of pruritus

Keum San Chun, Youn J. Kang, Jong Yoon Lee, Morgan Nguyen, Brad Lee, Rachel Lee, Han Heul Jo, Emily Allen, Hope Chen, Jungwoo Kim, Lian Yu, Xiaoyue Ni, KunHyuck Lee, Hyoyoung Jeong, JooHee Lee, Yoonseok Park, Ha Uk Chung, Alvin W. Li, Peter A. Lio, Albert F. Yang, Anna B. Fishbein, Amy S. Paller, John A. Rogers, and Shuai Xu

View the article online
https://www.science.org/doi/10.1126/sciadv.abf9405
Permissions
https://www.science.org/help/reprints-and-permissions

Use of this article is subject to the Terms of service