nature biotechnology

The Asthma Mobile Health Study, a large-scale clinical observational study using ResearchKit

Yu-Feng Yvonne Chan[1,2], Pei Wang[1], Linda Rogers[3], Nicole Tignor[1], Micol Zweig[1], Steven G Hershman[4], Nicholas Genes[1,2], Erick R Scott[1], Eric Krock[4], Marcus Badgeley[1], Ron Edgar[4], Samantha Violante[1], Rosalind Wright[3,5,6], Charles A Powell[3], Joel T Dudley[1,7] & Eric E Schadt[1]

The feasibility of using mobile health applications to conduct observational clinical studies requires rigorous validation. Here, we report initial findings from the Asthma Mobile Health Study, a research study, including recruitment, consent, and enrollment, conducted entirely remotely by smartphone. We achieved secure bidirectional data flow between investigators and 7,593 participants from across the United States, including many with severe asthma. Our platform enabled prospective collection of longitudinal, multidimensional data (e.g., surveys, devices, geolocation, and air quality) in a subset of users over the 6-month study period. Consistent trending and correlation of interrelated variables support the quality of data obtained via this method. We detected increased reporting of asthma symptoms in regions affected by heat, pollen, and wildfires. Potential challenges with this technology include selection bias, low retention rates, reporting bias, and data security. These issues require attention to realize the full potential of mobile platforms in research and patient care.

Three billion smartphones were in use worldwide in 2015, a figure expected to double by 2020 (ref. 1). Smartphones have replaced standard mail and landline phones for many people, creating a need to leverage mobile devices for research historically conducted by phone and mail. Mobile technology may also offer advantages over traditional data collection and management processes in research.

ResearchKit (Apple; Cupertino, CA, USA), an open source framework for mobile research can (i) obtain electronic informed consent, (ii) administer and collect questionnaires, (iii) actively and passively collect biometric data, (iv) provide reminders and notifications, and (v) reliably transmit and secure data in a central repository in compliance with regulatory requirements. Several research institutions and Sage Bionetworks (Seattle) collaborated with Apple to build the first mobile health applications using ResearchKit to demonstrate the feasibility of conducting research via this platform, and to provide an open source template to build third-party research apps[2–5]. To this end, we developed the Asthma Health Application (AHA) and conducted the Asthma Mobile Health Study (AMHS).

As many as half of the 25 million Americans with asthma lack optimal asthma control, contributing to $56 billion in annual disease costs[6]. A smartphone platform enabling large-scale, continuous collection of clinical, environmental, and passive biometric data may provide valuable insights for asthma research and clinical care. Our prospective observational mobile health study focused on assessing the following primary objectives: (i) feasibility of smartphone-based recruitment; (ii) characteristics of a study cohort recruited through the ResearchKit platform; (iii) user engagement and retention patterns; and (iv) user data sharing preferences. We tested the quality and utility of self-reported data collected by this method by assessing correlation with trusted external sources and concordance with expected patterns. Lastly, we evaluated the reported clinical impact of AHA use in a subset of participants.

RESULTS

Study enrollment, user experience, and data sharing

After its Apple App Store release on March 9th 2015, the AHA was downloaded 49,963 times over the first 6 months, 40,683 of which were from United States. Only US residents were eligible for the study. **Figure 1** describes the AHA enrollment process, cohort and key sub-cohort definitions (see **Supplementary Table 1a** for a comprehensive description of study sub-cohorts), user experience, and the geographic distribution of the study participants. A total of 7,593 users, out of 8,524 completed the enrollment process. Participants were asked to complete a series of intake surveys on demographics, comorbidities, and asthma history over four consecutive days after enrollment. Participants were also asked to complete daily asthma surveys to log symptoms, presumed triggers, and medication adherence for the duration of the study. In addition, the AHA administered weekly surveys to capture participants' healthcare utilization (HCU) and quality of life over the previous 7 d. EQ-5D health questionnaire

[1]Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, New York, USA. [2]Department of Emergency Medicine, Icahn School of Medicine at Mount Sinai, New York, New York, USA. [3]Department of Medicine, Pulmonary, Critical Care and Sleep Medicine, Icahn School of Medicine at Mount Sinai, New York, New York, USA. [4]LifeMap Solutions, Inc., New York, New York, USA. [5]Department of Pediatrics, Pulmonary and Critical Care, Icahn School of Medicine at Mount Sinai, New York, New York, USA. [6]Department of Environmental Medicine & Public Health, Icahn School of Medicine at Mount Sinai, New York, New York, USA. [7]Department of Population Health Science and Policy, Icahn School of Medicine at Mount Sinai, New York, New York, USA. Correspondence should be addressed to Y.-F.Y.C. (yu-fengyvonne.chan@mssm.edu) or E.E.S. (eric.schadt@mssm.edu).

Received 15 April 2016; accepted 13 February 2017; published online 13 March 2017; doi:10.1038/nbt.3826

and milestone surveys were also administered at less frequent intervals (see **Supplementary Table 1b** for survey descriptions and **Supplementary Note** for survey questions).

A total of 88% of participants chose to share de-identified data with researchers; 67% made their data available to all qualified researchers, and 21% shared data with study sponsors and partners. Study participants could send a copy of their own data to their e-mail by using the one-click 'Export Data' feature.

Different cohorts of AHA users and their baseline characteristics

Among the 7,593 enrolled participants, 6,470 responded to at least one survey in the study. We referred to these 6,470 users as the 'Baseline user' cohort. The 'Robust user' cohort (n = 2,317) included participants who: (i) were free from other lung disease and congestive heart failure, (ii) didn't smoke more than ten packs of cigarettes per year, and (iii) completed at least five daily or weekly surveys. Some of our analyses focused on the Robust users because of their more complete

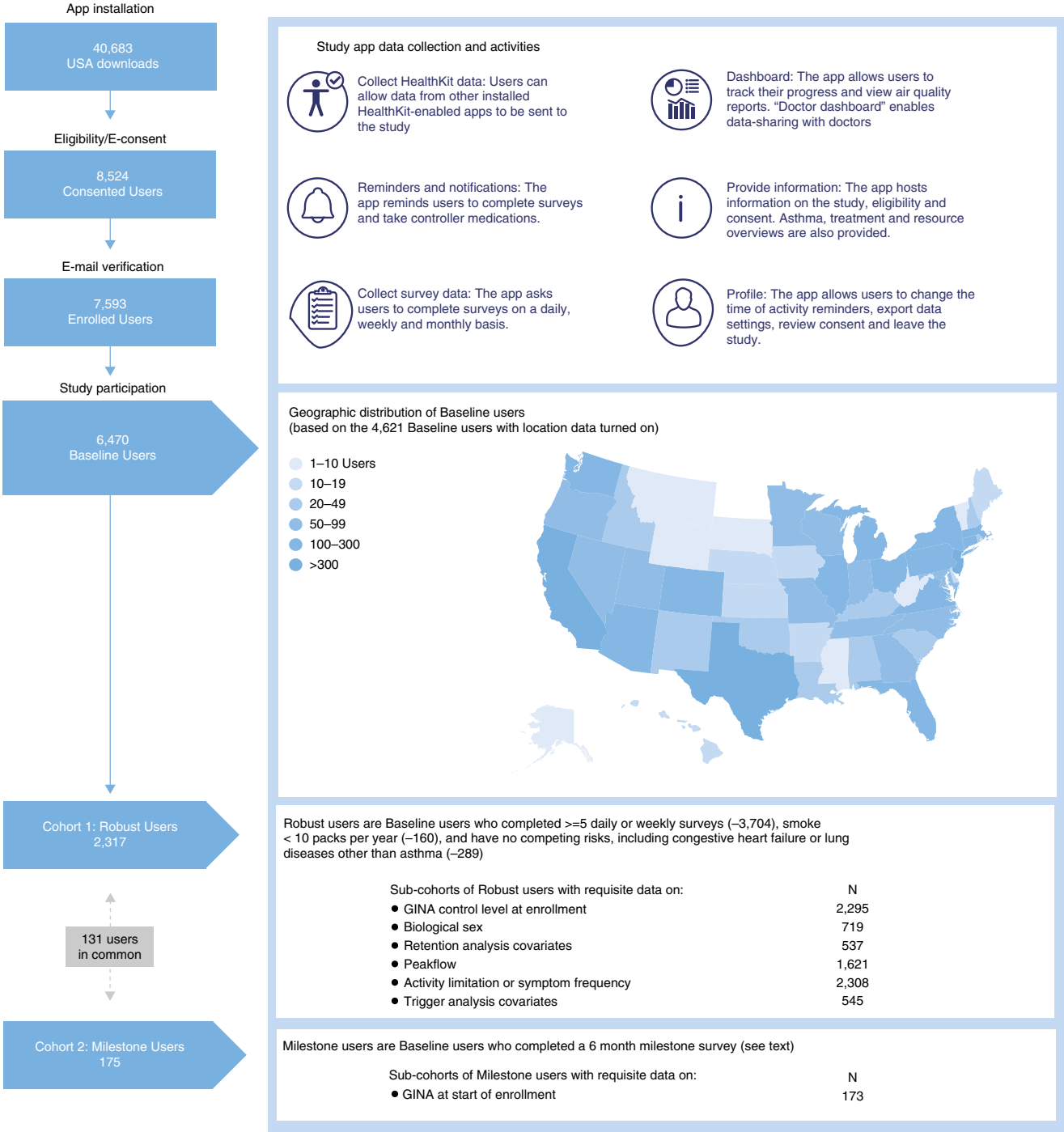

Figure 1 Recruitment process, user experience, and geographic distribution. The flow chart on the left indicates the recruitment process. The upper-right panel illustrates the activities experienced by users of AHA. The map illustrates the geographic distribution of Baseline users (i.e., Enrolled users who submitted at least one study survey). The box in the bottom outlines the selection of several of the key sub-cohorts used in our analyses, along with their sample sizes (see **Supplementary Fig. 1a** for more details).

baseline data, greater adherence to study survey completion, and fewer confounding variables.

The Milestone user cohort was a subset of 'Enrolled users' who completed the 6-month milestone survey. Of note, 131 out of the 175 Milestone users are included in the Robust user cohort. **Table 1** and **Supplementary Table 2** illustrate the demographic and baseline clinical characteristics of our Baseline, Robust, and Milestone users compared with population-based asthma statistics from the Centers for Disease Control and Prevention (CDC; Atlanta)[6], (http://www.cdc.gov/asthma/most_recent_data.htm). AHA users tended to be younger, wealthier, more educated, and were more often male than asthma patients in the CDC asthma population. Based on the location data from 4,621 Baseline users, **Supplementary Figure 1a** shows the high correlation (r = 0.85) of asthma prevalence, by state, between AHA users and the CDC cohort.

Baseline users had a higher rate of hospitalization (6%) and emergency department visits (11%) in the 6 months before enrollment than CDC rates (2% and 8%, respectively; **Table 1**)[3]. Of note, 13% of Baseline users reported a history of intubation and 37% reported the use of oral steroids (in the 6 months before enrollment) to treat an asthma exacerbation (**Supplementary Table 2**). During the study, high frequencies of symptoms were reported by a large number of users: (i) 47% of Baseline users reported symptoms in excess of twice weekly, (ii) 29% reported symptoms on most days or daily, and (iii) 37% reported the use of oral steroids (in the 6 months before enrollment) to treat an asthma exacerbation (**Supplementary Table 2**). Similar distributions were observed in Robust users and Milestone users. In addition, using Global Initiative for Asthma (GINA) criteria (http://www.ginasthma.org/) to assess asthma symptom control, uncontrolled asthma was reported in 43%, 44%, and 42% of Baseline users, Robust users, and Milestone users, respectively (Online Methods).

Download, enrollment, and retention patterns

Presumably, initial media publicity led to a high rate of AHA downloads, starting with 43,949 in the first month, yet decreasing to 300–400 per month by 5–6 months after the app launch. During months 3 to 6 of the study, the average monthly enrollment rate was 30% of downloads, with approximately 21 participants becoming Robust users per month in these later 3 months (**Supplementary Table 3a**).

Patients with worse asthma control enrolled at a relatively higher rate over time. For example, in the Robust user cohort, the proportion of GINA-uncontrolled participants in the first half of the study was

Table 1 Characteristics of patients at enrollment

		AHA			CDC	
Characteristic		Baseline	Robust	Milestone		% Dist.
Age[a]	18–34	756 (0.54)	369 (0.52)	56 (0.33)		36
	35–64	573 (0.41)	317 (0.44)	94 (0.56)		48
	65+	66 (0.05)	28 (0.04)	19 (0.11)		16
Gender[a]	F	539 (0.39)	285 (0.4)	80 (0.46)		59
	M	859 (0.61)	434 (0.6)	95 (0.54)		41
Race[b]	Black	203 (0.05)	108 (0.05)	8 (0.05)		14
	White	2,689 (0.68)	1,540 (0.7)	125 (0.77)		64
	Other	316 (0.08)	144 (0.07)	7 (0.04)		5
	Multirace	185 (0.05)	98 (0.04)	5 (0.03)		2
	Hispanic	587 (0.15)	315 (0.14)	18 (0.11)		13
Education[b]	HS nongrad	90 (0.02)	29 (0.01)	1 (0.01)		16
	HS grad	358 (0.09)	179 (0.08)	9 (0.05)		28
	Some college	1,521 (0.37)	844 (0.37)	60 (0.34)		33
	College grad	2,138 (0.52)	1,200 (0.53)	104 (0.6)		23
Income[b]	<$14,999	232 (0.06)	114 (0.05)	7 (0.04)	<$15,000	16
	$15,000–21,999	207 (0.05)	105 (0.05)	7 (0.04)	$15,000–24,999	17
	$22,000–43,999	550 (0.14)	314 (0.15)	21 (0.13)	$25,000–49,999	20
	$44,000–60,000	516 (0.13)	282 (0.13)	24 (0.15)	$50,000–79,999	11
	>$60,000	2,121 (0.55)	1,217 (0.57)	97 (0.6)	$75,000	21
	I don't know	235 (0.06)	94 (0.04)	6 (0.04)		NA
Visited ER[c]	Yes	707 (0.11)	222 (0.07)	13 (0.07)		8
	No	5,484 (0.89)	2,087 (0.93)	162 (0.03)		92
Hospitalized[c]	Yes	398 (0.06)	104 (0.03)	6 (0.05)		2
	No	5,789 (0.94)	2,205 (0.97)	169 (0.95)		98
Age of diagnosis[c]	0–18	4,847 (0.8)	1,769 (0.78)	117 (0.68)		NA
	<18	1,208 (0.2)	492 (0.22)	56 (0.32)		NA
Asthma Control medication[d]	Yes	3,772 (0.64)	1,613 (0.7)	142 (0.82)		NA
	No	1,896 (0.32)	630 (0.27)	30 (0.17)		NA
	Not sure	192 (0.03)	49 (0.02)	1 (0.01)		NA
Daily inhaled medicine[d]	ICS/LABA[f]	2,233 (0.65)	948 (0.65)	87 (0.68)		NA
	ICS	1,202 (0.35)	506 (0.35)	41 (0.32)		NA
GINA[e]	Uncontrolled	2,534 (0.43)	1,004 (0.44)	73 (0.42)		50[g]
	Partly controlled	2,246 (0.38)	947 (0.41)	82 (0.47)		NA
	Well controlled	1,067 (0.18)	344 (0.15)	18 (0.1)		NA

Comparison of demographic distributions for Baseline, Robust and 6-month Milestone users with CDC national asthma statistics (https://www.cdc.gov/asthma/most_recent_data.htm).
[a]Based on data from 1,430 Baseline, 738 Robust, and 175 Milestone users. [b]Based on data from 4,274 Baseline, 2,317 Robust, and 175 Milestone users. [c]Based on data from 6,240 Baseline, 2,311 Robust, and 175 Milestone users. [d]Based on data from 5,898 Baseline, 2,300 Robust, and 175 Milestone users. [e]Based on data from 5,897 Baseline, 2,295 Robust and 175 Milestone users. All percentages rounded to the nearest 100th. [f]ICS/LABA inhaled corticosteroid and long acting beta agonist combination therapy. [g]Instead of using the GINA criteria, the CDC used a slightly different criteria to define uncontrolled asthma patients as those who reported any of the following: (1) asthma symptoms more than two days a week in the past 30 days, (2) nighttime awakenings for more than one time a week in the past 30 days, or (3) short-acting β2-agonists use more than two days a week in the past three months (https://www.cdc.gov/asthma/asthma_stats/uncontrolled_asthma.htm). NA, not available.

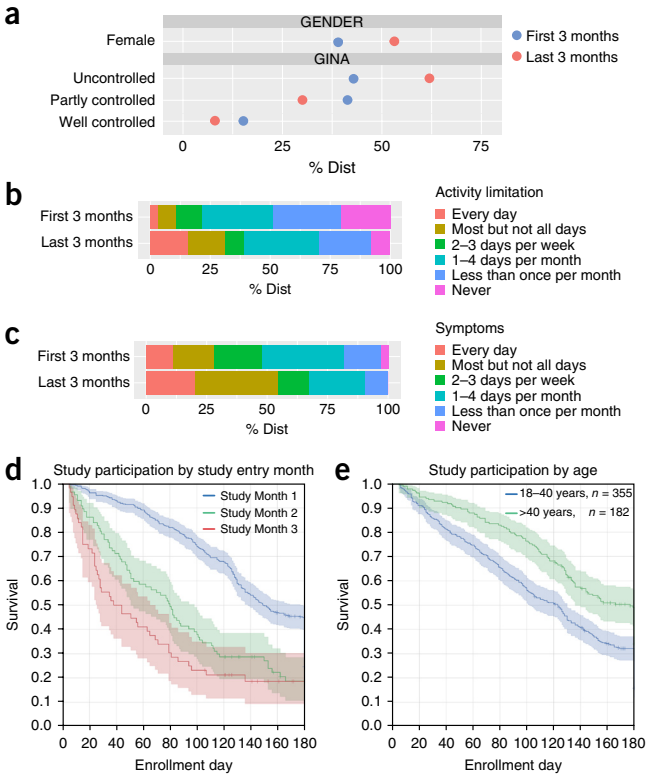

Figure 2 Enrollment and retention over time for Robust users. (**a**–**c**) The distributions of gender and GINA control (**a**), the frequency of activity limitation (**b**), and the frequency of daily asthma symptoms (**c**) are all significantly different between the first and the second halves of the study. Specifically, in the latter half of the study, the Robust users have a higher percentage of females (chi-squared = 4.74, d.f. = 1, $P = 0.03$, and $n = 719$), a higher percentage of users with uncontrolled asthma (chi-squared = 8.97, $P = 0.01$, and $n = 2,295$), and an increased frequency of symptoms (chi-squared = 22.3, d.f. = 5, $P = 0.001$, $n = 2,308$) and activity limitation (chi-squared = 36.9, d.f. = 5, $P = 0.0004$, $n = 2,308$). (**d**,**e**) Daily survey participation survival curves stratified by study entry month and reported age (Robust users, $n = 537$ participants, >90 d of post-enrollment follow-up). (**d**) Kaplan-Meier survival curve of daily survey participation stratified by study entry month and excluding participants entering after May ($n = 15$ participants). Study entry month of the participant was statistically significantly associated with daily survey participation longevity using a Cox proportional hazards model, $P = 2.99^{-23}$, hazard ratio 1.847 (95% CI, 1.64–2.08) for each passing month. (**e**) Kaplan-Meier survival curve of daily survey participation stratified by age (18–40 years and >40 years of age). Age was statistically significantly associated with daily survey participation longevity using a Cox proportional hazards model, $P = 1.59 \times 10^{-7}$, hazard ratio 0.976 (95% CI, 0.806–0.96) for each additional year of age. Colored bands show 95% confidence intervals for each strata.

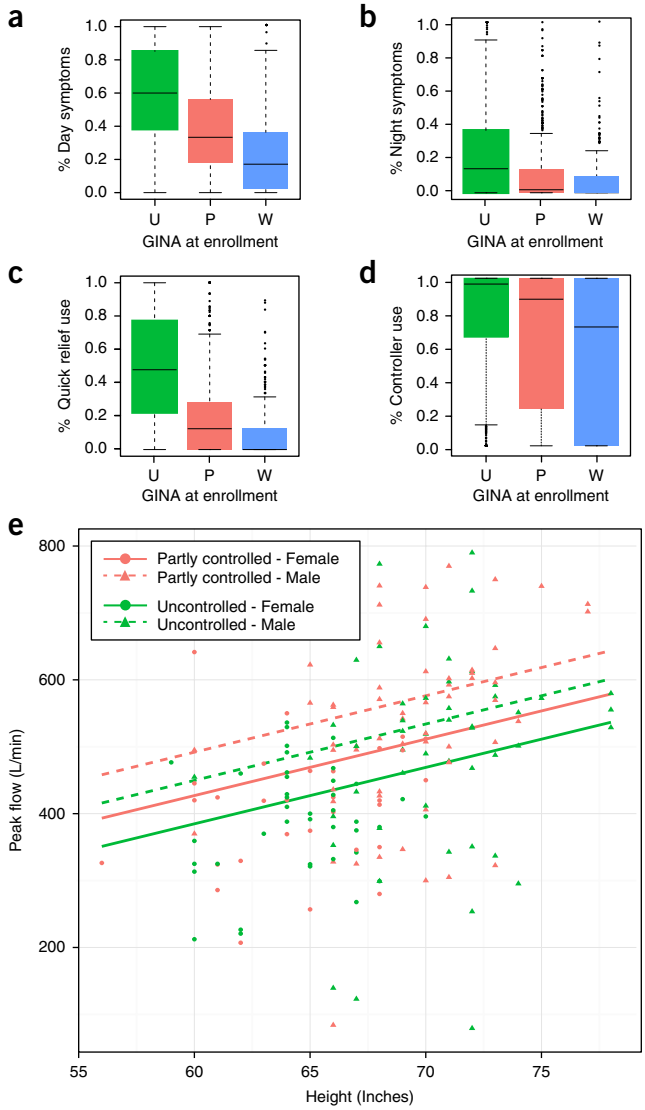

Figure 3 Concordance between GINA control at enrollment and prospectively collected daily symptoms reports during the study. (**a**–**d**) Distributions of frequencies of daytime symptoms (Kruskal–Wallis test; H(2) = 471.94, $P < 2.2^{-16}$, $n = 2,295$) (**a**), nighttime symptoms (Kruskal–Wallis test; H(2) = 232.23, $P < 2.2^{-16}$, $n = 2,295$) (**b**), inhaler puffs usage (Kruskal–Wallis test; H(2) = 677.12, $P < 2.2^{-16}$, $n = 2,295$) (**c**), and controller medicine usage (Kruskal–Wallis test; H(2) = 63.73, $P = 1.4^{-14}$, $n = 2,285$), and (**d**) among Robust users stratified according to their GINA control level at enrollment. (U, uncontrolled; P, partly controlled; W, well controlled). (**e**) Based on data from 183 Robust users, the lines illustrate a multiple linear regression model for peak flow trained on users' daily peak flow responses, GINA control assessed at enrollment, and HealthKit physique data, which demonstrates that male sex ($\beta = 64.847$, $t(179) = 2.836$, $P = 0.005$), controlled asthma ($\beta = 42.224$, $t(179) = 2.364$, $P = 0.02$), and height ($\beta = 8.435$, $t(179) = 2.695$, $P = 0.002$) are associated with greater peak flows.

43%, which was significantly lower than that in the second half of the study (62%) (chi-squared test: $P = 0.01$) (**Fig. 2a**). The distributions of different activity limitation levels and symptom frequencies indicated the same trend (**Fig. 2b**,**c**). The percentage of female users was also significantly higher in the second (53%) versus the first half (38%) of the study (chi-squared test: $P = 0.03$; **Fig. 2a**). The gender distribution of the cohort recruited later more closely approximated the CDC asthma population statistics. We also confirmed the significant associations between enrollment time and gender, GINA scores, activity limitation, as well as symptom frequencies from baseline surveys through logistic and regular regression analyses (**Supplementary Table 3b**).

Our study participants completed 79,297 daily and 10,969 weekly surveys over the 6-month study period. Of the 6,470 Baseline users, 6,023 and 2,520 participants responded to at least one daily and weekly survey, respectively. Total survey numbers collected in each month decreased over time (**Supplementary Table 3a**), consistent with an exponential decay function as observed in the other mobile health research studies[7]. To evaluate the impact of various factors on user

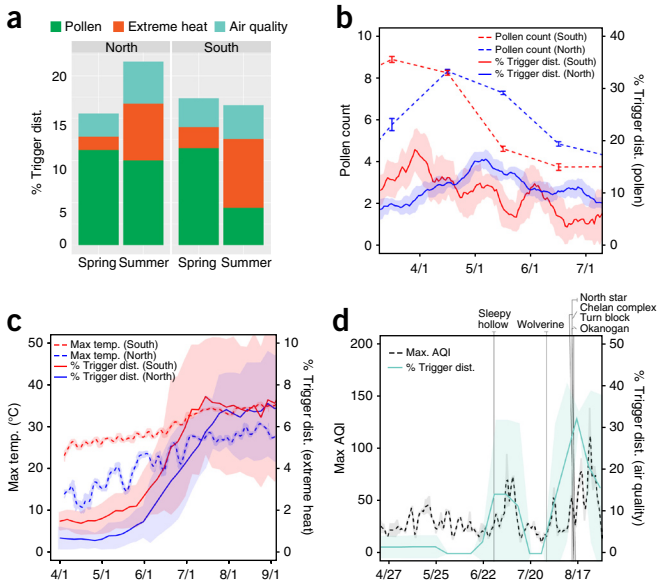

Figure 4 Geographic and seasonal trends in asthma triggers for Robust users. (**a**) The percentages of users reporting pollen, extreme heat or air quality as an asthma trigger (*y* axis) for southern (red) and northern (blue) regions of the contiguous US in the spring (March–May) and summer (Jun–Aug) respectively (based on $n = 545$ Robust users). (**b**) The percentage of users reporting pollen as an asthma trigger (solid) and the monthly pollen level (dashed) for southern (red) and northern (blue) regions of the US (based on $n = 64$ Robust users). (**c**) The percentage of users reporting extreme heat as their asthma triggers in southern and northern US regions for the spring and summer months (based on $n = 545$ Robust users). (**d**) The percentage of users reporting air quality as an asthma trigger for Washington state wildfires (solid, left *y* axis) and daily PM2.5 concentration (dashed, right-axis) in the same area (based on $n = 37$ Robust users). In (**b**–**d**), the shaded regions represent the ± 1 s.d. interval bands.

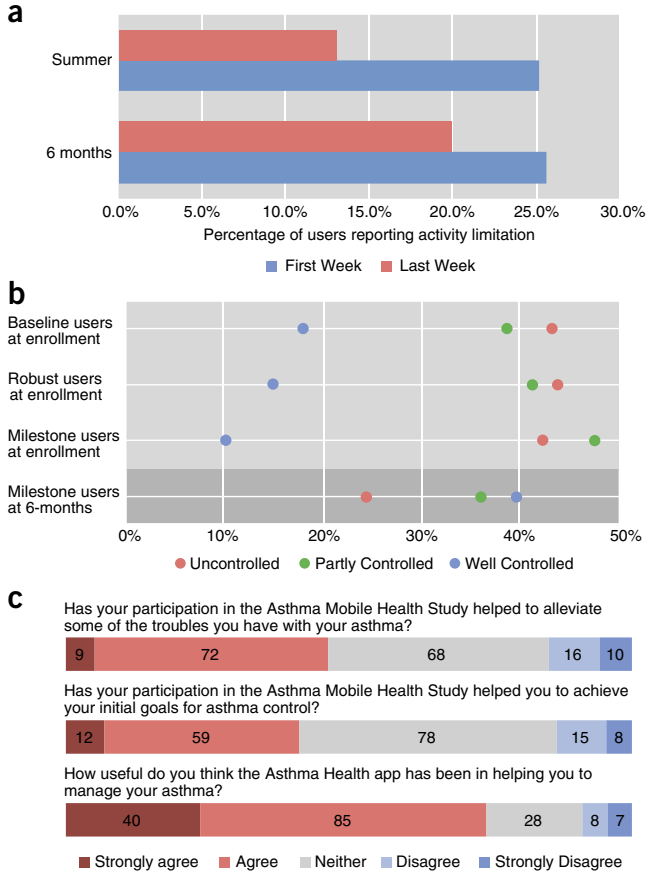

Figure 5 Positive impact of the app on user group. (**a**) The percentage of users reporting activity-limitation in their first week versus their last week in the summer (top, based on $n = 331$ Robust users) and in the entire 6-month study period (bottom, based on $n = 1{,}926$ Robust users). (**b**) The percent distribution of GINA control for all cohorts at enrollment (top three) and after 6-months of study participation (bottom). (**c**) Feedback and Milestone survey results based on data from Milestone users.

retention patterns, we focused on a subset of 537 users from the Robust user cohort, who were enrolled in the study for more than 90 d and who provided data for all the co-variates considered in the analysis. Both univariate and multivariate survival analyses of these 537 users, found earlier entrance into the study, (hazard ratio = 2.01 (95% CI, 1.73–2.33)) for each month following AHA launch, and increasing age, (hazard ratio = 0.978 (95% CI, 0.969–0.987)) for each additional year, significantly associated with greater likelihood of daily survey participation (**Fig. 2d**,**e**, **Supplementary Fig. 2a** and **Supplementary Table 4a**,**b**).

We also investigated the 'individual response rate', defined as the number of days with at least one daily survey question completed divided by the number of days enrolled through September 9, 2015, for each user. For the 537 users considered in the above retention analysis, the average individual response rate was 31%, with 104 of these users having an individual response rate >50% (**Supplementary Fig. 2b**). Increasing age and earlier study entry month were also associated with higher individual response rate (**Supplementary Table 4c**).

Relationship between baseline asthma control and prospectively collected data

Participants completed intake questionnaires assessing asthma control upon study enrollment and then prospectively reported daily and nightly asthma symptoms, quick-relief inhaler usage, controller medicine usage, and peak flow measures over the course of their participation in the study. Patients' daily survey responses for the aforementioned four parameters were all found to be significantly associated with the GINA control levels calculated based on intake questionnaires from Robust users reporting daily (Kruskal–Wallis test; $H(2) = 471.94$, $P < 2.2^{-16}$, $n = 2{,}295$) and nightly symptoms (Kruskal–Wallis test; $H(2) = 232.23$, $P < 2.2^{-16}$, $n = 2{,}295$), quick-relief inhaler usage (Kruskal–Wallis test; $H(2) = 677.12$, $P < 2.2^{-16}$, $n = 2{,}295$), and controller medicine usage (Kruskal–Wallis test; $H(2) = 63.73$, $P = 1.4^{-14}$, $n = 2{,}285$) (**Fig. 3a**–**d** and **Supplementary Fig. 3**).

Of those in the Robust user cohort, 1,621 voluntarily submitted at least one peak flow measurement during the study period. As expected, patients with well-controlled asthma, and who were male and tall, had higher average peak flows throughout the study period (**Fig. 3e**). Those with uncontrolled asthma at baseline reported peak flows 42 liters/min lower than their well-controlled counterpart after adjusting for height and gender ($n = 183$).

We examined concordance of reported asthma symptoms, rescue inhaler use, and peak flow measurements using time series of daily survey responses from Robust users. Consistent with clinical expectation, we detected a positive correlation between daily or nightly symptoms and rescue inhaler use ($n = 979$ and $n = 761$, respectively), whereas these same variables were negatively correlated with peak flow values ($n = 235$ and $n = 173$, respectively) (**Supplementary Fig. 4**).

Geographic and temporal trends in asthma triggers

Animals, pollen, and upper respiratory tract infections were the top three asthma triggers reported by participants at time of enrollment (**Supplementary Fig. 5**). **Figure 4a** illustrates the distribution of three self-reported asthma triggers from the 545 users of the Robust user cohort for whom we had consent to obtain geolocation data for asthma triggers during the spring and summer months (**Supplementary Fig. 1b**,**c**). For example, before 1 April 2015, a greater percentage of users reported pollen as a trigger in Southern regions of the United States (where pollen counts were higher) than in Northern regions (**Fig. 4b**). Additionally, the percentage of users who reported extreme heat as a symptom trigger correlated with maximum daily temperature trends (i.e., climate data reports of the highest temperature recorded over a specified period of time)[8]. As expected, a higher percentage of Southern participants reported extreme heat as a trigger (**Fig. 4c**).

Our data also indicate that some asthma patients were sensitive to air quality changes caused by environmental events, such as wildfires. There was a marked increase of participants reporting air quality triggers in regions affected by the summer 2015 Washington state wildfires during the corresponding time periods (**Fig. 4d**).

Self-reported clinical impact of AHA use

We assessed reported activity levels in the subset of our study participants comprising the 1,926 users from the Robust user sub-cohort who had been enrolled for more than 90 d and for whom we had relevant data from the first and last weeks of participation. These participants reported a significant decrease of activity limitation from 25–20% (Wilcoxon signed-rank test, $P < 0.0001$; **Fig. 5a** and **Supplementary Table 5a**). The same trend persisted during the three summer months, from 25–13% (Wilcoxon signed-rank test, $P < 0.0001$, $n = 331$, **Fig. 5a** and **Supplementary Table 5b**), suggesting that milder weather did not drive changes in activity levels.

Likewise, 173 Milestone users for whom the 6-month milestone survey results and baseline GINA information were available, reported that their asthma control substantially improved over the study period. Specifically, the percentage of Uncontrolled users changed from 42–24% (paired Wilcoxon signed-rank test, $P < 0.0001$; **Fig. 5b** and **Supplementary Table 5c**). A modest majority of participants also reported that the AHA was useful in helping them manage their asthma (**Fig. 5c** and **Supplementary Table 5d**,**e**).

DISCUSSION

The AMHS is one of the few studies to examine the value and validity of the novel mobile health research platform, ResearchKit[3]. We conducted a prospective, observational study focused on the feasibility of conducting research remotely via this platform and observed certain strengths as well as limitations to this methodology. In terms of strengths, the AMHS demonstrated that a broad-scale asthma study can be conducted in its entirety via a smartphone application, including remote recruitment, consent, enrollment, and secure bi-directional data flow between investigators and participants. We prospectively collected detailed, multi-dimensional, longitudinal data on an asthma cohort more efficiently than traditional epidemiological studies by automating, standardizing, and accelerating various costly and time-consuming processes. Our study's rapid recruitment and participants' willingness to share de-identified data broadly highlight users' acceptance of this methodology for low-risk health studies.

As participant recruitment is a significant challenge in research today[9,10], digital health and social media could play a role in addressing that challenge. Within 1 d of the launch, the five ResearchKit studies collectively enrolled over 15,000 participants across the country, beyond university catchment areas, demonstrating the power and potential of this technology[11]. Characteristics of our study cohort compared with CDC Asthma Surveillance Data identified similarities and differences. Owners of iPhones have higher education levels and income than other smartphone users, who as a group have higher income and education levels than the general population[12] (**Supplementary Fig. 6**). Of note, only 5% of AHA users with asthma were Black, compared with 13% of the US population, an under-representation commonly encountered in clinical research in general. In the United States, 92% of Hispanics, 91% of Whites, and 94% of Blacks report using a mobile phone, with 64% of Hispanics, 66% of Whites and 64% of Blacks using a smartphone[12]. Use of the Android platform is more common in some racial and/or ethnic groups, thus the availability of an Android version of AHA (facilitated by the open source ResearchKit framework) could capture a more representative sampling of the general population. Other nuances, such as a propensity to text rather than use apps in low-income Hispanic communities should be considered in attaining diverse participation in mobile health research[13,14].

An understanding of use patterns and the ability to reach and impact diverse populations regardless of platform used will become increasingly important as mobile health technology expands in healthcare and research. Given the overwhelming trend toward use of digital communication, responses to landline telephone and mailed surveys may soon represent a more restricted and non-representative population[12]. For example, the vast majority of New York City residents (96%) own a cell phone, 79% of which are smartphones, including smartphone ownership among 67% of low income New Yorkers[12]. Therefore, the ability to recruit participants via digital technology without direct voice contact or print mailing will likely be needed to conduct population-based research. The possibility exists that the gold-standard population metrics obtained via traditional research methodology used for comparisons in our study may already contain biases. Offering and leveraging the strength of digital health and traditional research methods could optimize clinical study enrollment, participant retention, and data capture more than either method alone, getting us closer to a representative sampling of the general population. Moreover, offering a process that automates certain time-consuming and labor-intensive components, while incorporating a 'human touch' at selective key points may improve study retention and decrease costs[15–17].

Higher rates of HCU in AHA users than in CDC data may indicate a potential for selection bias in mobile health recruitment via uptake of this type of study by sicker patients. Conversely, use of a mobile platform may increase participant diversity in research by enabling the enrollment of traditionally under-represented and difficult-to-recruit populations (e.g., those with disabling, severe disease or limited healthcare access). The ability to complete study requirements at home on flexible time schedules may remove barriers (i.e., location, mobility, psychosocial factors, work hours) to research participation. The impact of selection bias and other potential threats to validity of mobile health research are currently poorly characterized and deserve ongoing investigation.

An important goal of our study was to evaluate the quality and validity of study data obtained via this mobile research platform. Patient-reported outcomes (PROs) are an important component of research in asthma and are often collected via paper diaries, where important issues in authenticity of data have been raised, including 'back filling', 'forward filling' and falsified data[18]. Because the accuracy of PRO data may be partially dependent on the relationship between

the reporter and the recipient[19], the validity of PRO data obtained via mobile health platforms (without direct participant–investigator contact) warrants rigorous evaluation.

More widespread use of electronic versus paper diaries outside of industry research studies is currently prohibited by cost but may be fostered by an open-source platform such as ResearchKit. In clinical trials, the PROs are correlated with objective measures of lung function, but in epidemiologic asthma research, it is commonplace to use symptom-based surveys without corresponding lung function measurements[20]. In this respect, our surveys and the types of epidemiologic data gathered do not differ from common practice, apart from the use of technology to scale and accelerate the process. Whereas concerns about falsifying information, such as identity authentication, have been raised in regard to mobile health studies. The consent and registration process of the AMHA is fairly vigorous and may mitigate such risk (e.g., e-mail verification of identity and entering a passcode to access the app). Validity of our data is supported by concordance between our cohort's self-reported asthma status at baseline and prospectively collected data. For example, participants' daily survey responses for day and night symptoms as well as inhaler and controller medication usage were all found to be significantly associated with the GINA control levels calculated based on intake questionnaires for these parameters. Similarly, the peak flow measurements submitted by participants were of expected range based on known trends for patients' sex, height, and asthma control status. Furthermore, we detected that patients' asthma symptoms correlated well with the frequency of rescue inhaler usage and peak flow values as would be expected based on the clinical behavior of asthma. Likewise, the self-reported asthma triggers (e.g., pollen, extreme temperature, air quality, pollutant exposures), mapped based on geography and time, correlated well with objective measures (e.g., external, validated environmental sources).

Of particular note, the summer 2015 Washington state wildfire analysis highlights that smartphone-based technology could provide innovative, scalable solutions for clinical research aspirations that were logistically not feasible or cost-prohibitive in the past. Specifically, we correlated and detected a marked increase in our study participants' daily asthma symptoms (air quality triggers) with real-time fine particulate matter (Environmental Protection Agency (EPA) air quality logs PM2.5) levels in regions affected by wildfires during those corresponding time periods. Conventional assessments of the effect of natural phenomena on disease are usually very limited due to the aforementioned difficulties. Since our AHA data set already contain location-specific environmental data, such exploratory analyses require minimal additional effort or cost to accomplish. In summary, the consistent trending of variables that we expect to be interrelated based on our knowledge of the disease, the tracking of symptoms with known environmental triggers, and the expected correlation among symptoms, medication use, and lung function in a subset of patients for whom lung function was measured strongly suggest the validity of this new research-data collecting method.

Characterizing survey participation rates and engagement with the AHA over time was another major aim of this study. We observed a large initial number of downloads likely driven by media publicity, which decreased to a steady rate over time. We attribute the moderate enrollment rate (14–38% of downloads) to the relative ease of app download with some dropoff related to the rigor of the consent process. Additionally, the significant rate of attrition of participants observed from the initial cohort to some of the sub-analyses conducted raises issues of generalizability. Also of note, although we did not actively recruit participants (thus, no incremental budget was required), we continued to enroll new participants daily a year after the launch of the study. (Please see **Supplementary Note** for study marketing details.)

The drop-off in user retention over time that we observed appears to be shared by multiple 'digital' use cases (e.g., mobile apps including entertainment 'gaming' apps, tutorial videos, open online courses) and speaks to the hardwired biopsychosocial tendencies of users. Because the ultimate goals of digital health generally rely on prolonged participation, the creators of these tools must understand users' psychosocial-behavioral needs and predilections to keep them continually engaged. Attention and resources must be devoted to incorporating social and behavioral principles in digital health design beyond technical ones. For the AMHS, our study participants were not offered financial incentives—a standard practice for clinical studies when participation for prolonged periods of time is required[21,22]. In fact, we did not offer users any other 'tangible' rewards beyond features like environmental data, educational modules, and the ability to track entered data. Monetary incentivization, possibly in the form of micro-payments and/or 'advanced gamification'[23,24] may improve retention, especially if a study requires long term follow-up.

Despite attrition of longitudinal participation over time, 85.2% of the 7,593 enrolled users responded to at least one survey in the study. Although the comparisons are not analogous, we referred to CDC's Behavioral Risk Factor Surveillance System (BRFSS-CDC) statistics for context and as a reference point in interpreting our findings. The BRFSS cooperation rates (defined as the number of complete and partial complete interviews divided by the number of contacted and eligible respondents) were 62.5% for landline-based surveys and 71.6% for cell-phone-based surveys[25]. These findings suggest that studies relying on mobile apps could achieve similar or better cooperation rates than traditional (landline or mobile) phone-based population research methods.

We encountered some additional challenges in performing our study. We were unable to incorporate certain standard validated asthma surveys into our study due to licensing constraints. Besides the Euroquol-5D, our AHA surveys were developed by asthma specialists who incorporated general content used by validated survey instruments. Because of an initial technical issue with the integration of HealthKit and ResearchKit data, we obtained only gender and/or age information for 1,398 participants in the study, limiting several analyses. Multiple versions of the AHA were released during the study period to address these software-related concerns and to implement new features (**Supplementary Table 6**). Similar to the other mobile health studies and large-scale research studies in general, we encountered substantial missing values in our AHA data, such that many parallel analyses were presented based on different sample sizes (**Fig. 1** and **Supplementary Table 1a**).

Furthermore, the analyses on the self-reported clinical impact of AHA use were based on subsets of our study participants and may not be representative of our cohort at large. We did note consistent positive feedback from our users (e.g., decreased activity limitation, helpful in disease management), and these sentiments were echoed by some AHA users on our participant panel (http://apps.icahn.mssm.edu/asthma/participant-stories/). However, validating the clinical impact of AHA usage warrants rigorous future clinical trials.

Although mobile health apps and devices may promote health literacy and medication adherence, and mitigate exacerbations of chronic diseases, research on this technology thus far has been mostly small scale and yielded conflicting results. Studies to date have not found advantages in terms of HCU or cost, and adoption of the technology in healthcare settings remains low[26–29].

However, the field of devices and mobile health research has witnessed some promising recent advances. For example, D'Arcy *et al.*[30] developed a Bluetooth inhaler device that used acoustic recordings of inhaler usage to monitor temporal and technique adherence and assess the correlation between clinical outcomes and adherence. Furthermore, studies using the US Food and Drug Administration (FDA)-approved Propellor Health (Madison, WI, USA) digital platform and inhalers demonstrated the feasibility of tracking inhaler usage and triggers[31–34]. Moreover, individuals randomized to the Propeller Health arm had more significant reductions in inhaler usage than those under routine care[35]. Additionally, an observational mobile health study of Parkinson disease that leverages the ResearchKit platform, named mPower, interrogated aspects of the movement disorder and assessed high-resolution activity data collected through surveys and frequent sensor-based recordings from participants with and without Parkinson disease[36]. The large-scale and repeated measurements of thousands of individuals may help establish baseline variability of real-world activity measurements collected via smartphones and lead to quantification of the ebbs and flows of Parkinson symptoms[36]. The mPower research team developed a framework that accounts for research participant choices regarding clinical data sharing and also qualifies researchers requesting to access the de-identified data for secondary analyses[37]. Lastly, the investigators from the five ResearchKit launch partners published a report on their collective experience and the platform's potential and limitations, including selection bias, identity uncertainty, design limitations, retention, and privacy[38].

Based on the studies above and the initial results from the AMHS, we believe research hypotheses with the following characteristics are a good match for the current ResearchKit methodology: 1) a requirement for rapid enrollment across diverse geographical locations; 2) a design that presents minimal risk to participants, allowing the use of electronic consent; 3) a hypothesis that can be answered in a short time period (1–3 weeks); 4) a requirement for frequent data collection events; 5) data collection that is passive (e.g., GPS, physical activity); 6) no assumption that results will be generalizable to participants recruited via traditional methods; and 7) a sample size and statistical analysis plan that account for the known attrition and/or missing data historically seen in internet and mobile app studies.

Mobile health research represents a promising new avenue for clinical research. It has the potential to open up new possibilities for data collection, provide novel insights into disease, and reach participants that traditional studies may fail to adequately represent. Mobile health research must be amenable to the demands of a fast-paced, variable research environment, as well as be methodologically rigorous. The challenges associated with this technology, including selection bias, potential reporting bias, data security, and low user-retention rate will need to be addressed in order to better understand the technology's true value and role in research and patient care. In the future, mobile platforms may serve as a primary driver for conducting large-scale studies, perhaps complemented by traditional means to leverage the strengths of both methods. Looking forward, the potential of ubiquitous smartphone technology to address the needs of clinical research to better understand health and disease appears to be more promising than ever.

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

The study is funded by the Icahn School of Medicine at Mount Sinai, UL1TR001433-01, and with technology support from LifeMap Solutions.

AUTHOR CONTRIBUTIONS

Y.-F.Y.C. developed the initial study design/protocol, including electronic informed consent and the statistical plan, IRB submission and approval, app design and implementation, and budget management, ensured proper study execution, provided clinical support, refined surveys, and assisted in data interpretation, manuscript writing and revision and preparation for submission. P.W. contributed to study design and survey refinement, led statistical support and provided oversight for all data analysis and interpretation, generated figures and tables, and was a major contributor to manuscript writing and revision. L.R. assisted in the initial study design/protocol and IRB preparation and submission, led in the design of surveys, provided clinical support, participated in manuscript writing, and served as NJH liaison. N.T. provided statistical support, including data analysis and interpretation, generated figures and tables, participated in study design and survey refinement, provided major contributions to manuscript writing and revision, and served as a graphic artist liaison. M.Z. assisted in electronic informed consent design, led subsequent IRB submission and provided support, refined surveys, and was a major contributor to manuscript writing and preparation for submission. S.G.H. served as LifeMap Solutions scientific lead and provided support for AHA design, implementation, and functionality, served as a liaison to other technology partners, refined surveys, and assisted in data interpretation and manuscript writing. N.G. led the latter part of the study execution, provided subsequent IRB support, provided clinical support, refined surveys, and assisted in data interpretation and manuscript writing. E.R.S. provided statistical support and data analysis and interpretation, assisted in generating figures and tables, provided subsequent IRB support, and participated in manuscript writing. S.V. was involved in study execution and manuscript writing. M.B. assisted in data analysis and interpretation, and generated figures and tables. E.K. contributed to app design and implementation, as well as initial IRB document preparation. R.E. was the LifeMap Solutions technical lead, participated in app design and implementation, and helped ensure data integrity. R.W. assisted in study design and data interpretation. C.A.P. contributed to study design and data interpretation. J.T.D. contributed to study design input and manuscript revision. E.E.S. participated in study design, oversaw study execution, interpreted data, and participated in manuscript writing and revision.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the online version of the paper.

1. Topol, E.J., Steinhubl, S.R. & Torkamani, A. Digital medical tools and sensors. *J. Am. Med. Assoc.* **313**, 353–354 (2015).
2. Ritter, S. Apple's research kit development framework for iPhone apps enables innovative approaches to medical research data collection. *J. Clin. Trials* **5**, 2–3 (2015).
3. Jardine, J., Fisher, J. & Carrick, B. Apple's ResearchKit: smart data collection for the smartphone era? *J. R. Soc. Med.* **108**, 294–296 (2015).
4. Morgan, A., Mooney, S., Aronow, B. & Brenner, S. Precision Medicine: Data and Discovery for Improved Health and Therapy". Pacific Symposium on Biocomputing. Accessed May 7, 2016. http://psb.stanford.edu/psb-online/proceedings/psb16/intro-pm.pdf (2016).
5. Savage, N. Mobile data: Made to measure. *Nature* **527**, S12–S13 (2015).
6. Akinbami, O.J., Moorman, J.E. & Liu, X. Asthma prevalence, health care use, and mortality: United States, 2005–2009. (US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics Washington, DC, 2011).
7. Dorsey, E.R. *et al.* The use of smartphones for health research. *Acad. Med.* **92**, 157–160 (2017).
8. Menne, M.J. *et al.* Global historical climatology network-daily (GHCN-Daily), Version 3. (NOAA National Climatic Data Center, 2012).
9. Gul, R.B. & Ali, P.A. Clinical trials: the challenge of recruitment and retention of participants. *J. Clin. Nurs.* **19**, 227–233 (2010).
10. Lovato, L.C., Hill, K., Hertert, S., Hunninghake, D.B. & Probstfield, J.L. Recruitment for controlled clinical trials: literature summary and annotated bibliography. *Control. Clin. Trials* **18**, 328–352 (1997).
11. Lovejoy, B. ResearchKit did in 24 hours what would normally take 50 medical centers a year—Stanford University. *9to5 Mac.* Accessed Aug 15, 2016. https://9to5mac.com/2015/03/12/researchkit-usage/ (2015).

12. Smith, A. US Smartphone use in 2015. Accessed September 5 2016: http://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015/ (Pew Research Center, 2015).
13. Arora, S. *et al.* Describing the evolution of mobile technology usage for Latino patients and comparing findings to national mHealth estimates. *J. Am. Med. Inform. Assoc.* **23**, 979–983 (2016).
14. Wei, J., Hollin, I. & Kachnowski, S. A review of the use of mobile phone text messaging in clinical and healthy behaviour interventions. *J. Telemed. Telecare* **17**, 41–48 (2011).
15. Dolan, B. How mobiles make clinical trials bigger, faster and more efficient. *Mobile Health News.* Accessed June 11, 2016. http://www.mobihealthnews.com/28198/how-mobiles-make-clinical-trials-bigger-faster-and-more-efficient (2013).
16. Miseta, E. Mobile devices in clinical trials: tested, effective, proven!. *Clinical Leader.* Accessed on May 5, 2016. https://www.clinicalleader.com/doc/mobile-devices-in-clinical-trials-tested-effective-proven-0002 (2016).
17. Henderson, L. Cell phone key to REMOTE data collection. *Applied Clinical Trials.* Accessed on May 3, 2016. http://www.appliedclinicaltrialsonline.com/cell-phone-key-remote-data-collection (2011).
18. Coons, S.J. *et al.* Capturing patient-reported outcome (PRO) data electronically: the past, present, and promise of ePRO measurement in clinical trials. *Patient* **8**, 301–309 (2015).
19. Podsakoff, P.M. & Organ, D.W. Self-reports in organizational research: problems and perspectives. *J. Manage.* **12**, 531–544 (1986).
20. Murphy, K.R. *et al.* Asthma management and control in the United States: results of the 2009 asthma insight and management survey. *Allergy Asthma Proc.* **33**, 54–64 (2012).
21. Halpern, S.D., Karlawish, J.H., Casarett, D., Berlin, J.A. & Asch, D.A. Empirical assessment of whether moderate payments are undue or unjust inducements for participation in clinical trials. *Arch. Intern. Med.* **164**, 801–803 (2004).
22. Russell, M.L., Moralejo, D.G. & Burgess, E.D. Paying research subjects: participants' perspectives. *J. Med. Ethics* **26**, 126–130 (2000).
23. King, D., Greaves, F., Exeter, C. & Darzi, A. 'Gamification': influencing health behaviours with games. *J. R. Soc. Med.* **106**, 76–78 (2013).
24. Miller, A.S., Cafazzo, J.A. & Seto, E. A game plan: Gamification design principles in mHealth applications for chronic disease management. *Health Informatics J.* **22**, 184–193 (2016).
25. Centers for Disease Control and Prevention. Behavioral risk factor surveillance system: summary data quality report. Accessed Aug. 14, 2016. https://www.cdc.gov/brfss/annual_data/2013/pdf/2013_dqr.pdf (CDC, 2013).
26. Alkhaldi, G. *et al.* The effectiveness of technology-based strategies to promote engagement with digital interventions: a systematic review protocol. *JMIR Res. Protoc.* **4**, e47 (2015).
27. Wantland, D.J., Portillo, C.J., Holzemer, W.L., Slaughter, R. & McGhee, E.M. The effectiveness of Web-based vs. non-Web-based interventions: a meta-analysis of behavioral change outcomes. *J. Med. Internet Res.* **6**, e40 (2004).
28. Ryan, D. *et al.* Clinical and cost effectiveness of mobile phone supported self monitoring of asthma: multicentre randomised controlled trial. *Br. Med. J.* **344**, e1756 (2012).
29. Bloss, C.S. *et al.* A prospective randomized trial examining health care utilization in individuals using multiple smartphone-enabled biosensors. *PeerJ* **4**, e1554 (2016).
30. D'Arcy, S. *et al.* A method to assess adherence in inhaler use through analysis of acoustic recordings of inhaler events. *PLoS ONE* **9**, e98701 (2014).
31. Su, J.G. *et al.* Feasibility of deploying inhaler sensors to identify the impacts of environmental triggers and built environment factors on asthma short-acting bronchodilator use. *Environ. Health Perspect.* **125**, 254–261 (2017).
32. Hai, A. *et al.* Adherence monitoring and E-health: how clinicians and researchers can use technology to promote inhaler adherence for asthma. *J. Allergy Clin. Immunol. Pract.* **1**, 446–454 (2013).
33. van Boven, J.F., Trappenburg, J.C., van der Molen, T. & Chavannes, N.H. Towards tailored and targeted adherence assessment to optimise asthma management. *NPJ Prim. Care Respir. Med.* **25**, 15046 (2015).
34. Howard, S., Lang, A., Patel, M., Sharples, S. & Shaw, D. Electronic monitoring of adherence to inhaled medication in asthma. *Curr. Respir. Med. Rev.* **10**, 50–63 (2014).
35. Merchant, R.K., Inamdar, R. & Quade, R.C. Effectiveness of population health management using the propeller health asthma platform: a randomized clinical trial. *J. Allergy Clin. Immunol. Pract.* **4**, 455–463 (2016).
36. Bot, B.M. *et al.* The mPower study, Parkinson disease mobile data collected using ResearchKit. *Sci. Data* **3**, 160011 (2016).
37. Wilbanks, J. & Friend, S.H. First, design for data sharing. *Nat. Biotechnol.* **34**, 377–379 (2016).
38. Trister, A.D., Dorsey, E.R. & Friend, S. Smartphones as New Tools in the Management and Understanding of Parkinson's Disease. 2016NJP-Parkinson's Disease. Accessed Aug 1 2016. http://www.nature.com/articles/npjparkd20166.

ONLINE METHODS

Setting for Apple ResearchKit and AHA. The Icahn School of Medicine at Mount Sinai collaborated with Sage Bionetworks, LifeMap Solutions (New York), and Apple to develop and launch the Asthma Health App and Asthma Mobile Health Study on March 9, 2015. This study was approved by the Institutional Review Board of the Icahn School of Medicine at Mount Sinai.

Code availability. 'The asthma health mobile app version 1.011 was built using Apple's ResearchKit framework (http://researchkit.org/), which is open source and available on GitHub (https://github.com/researchkit/researchkit). AppCore (https://github.com/ResearchKit/AppCore) is a layer built on top of ResearchKit that was shared among the five initial ResearchKit apps. The Bridge iOS SDK (https://github.com/Sage- Bionetworks/Bridge-iOS-SDK) provides integration with Sage Bionetworks' Bridge Server, a back-end data service designed for collection of participant donated study data (https://sagebionetworks.jira.com/wiki/display/BRIDGE/Bridge+REST+API)[36]. Code used in this study is also available as **Supplementary Software**.

Participant recruitment and enrollment. The AHA is available in the iTunes App Store searchable using a combination of keywords including 'asthma', 'health', and/or 'Mount Sinai'. Links to the AHA store page were embedded within relevant web sites of Mount Sinai, LifeMap Solutions, and Apple Inc. Prospective participants can download the app using their own App Store credentials. After downloading and opening the AHA, prospective participants are first presented with an inclusion/exclusion criteria questionnaire. Participants who meet the eligibility criteria proceed to electronic informed consent screens. To ensure participants understand the risks, benefits, and options of study participation, they must then pass a comprehensive quiz before creating an account. Following the creation of an account, prospective participants are asked to verify their e-mail address by clicking on a link in their e-mail to confirm their enrollment. The AHA user experience and recruitment process is detailed in **Figure 1**. Please see **Supplementary Note** for information on marketing for the study.

Eligibility criteria for participants. Participants were eligible to enroll in the study if they were aged 18 or over, lived in the United States, had a diagnosis of asthma with physician-prescribed asthma medications, and had an iPhone with a data plan. Pregnant women, non-English speakers, and those who could not document understanding of the consent based on a series of key questions were ineligible for the study.

Study design, data flow, and security. Supplementary Figure 7a provides a simplified layout of the initialization process in the Asthma Mobile Health Application study. Data entered by participants were collected and deidentified using advanced data security technologies developed by our partners at Sage Bionetworks. Data from the study were not shared with Apple or with non-study personnel. Please see **Supplementary Figure 7b** for detailed description of the backend design on health data encryption and securely stored.

Data sources for active data collection (surveys). Study surveys appeared on the participant's 'Activities' screen. Baseline surveys collected data on: 1) asthma history, including the frequency and time of symptoms and activity limitation; 2) asthma experience, including triggers and personal management plans; 3) medical history; and 4) demographics. Additionally, baseline healthcare utilization, asthma medications, self-reported controller adherence, quick relief/rescue medication use, peak flow, and other clinically relevant data were collected. Participants were also asked to set a goal for asthma control and complete the EQ-5D-5L (EuroQol version 5D-5L). After the intake process, daily and weekly surveys were administered, with EuroQol, 6-month milestone, and app feedback surveys occurring at less frequent intervals.

Data sources for passive data collection. Participants were asked for permission for the app to read various fields of HealthKit data. HealthKit data that already exists in the user's iPhone from other apps can be collected by app and relayed to the study data. Also, with the participant's consent, the nearest EPA air-quality reading along with that station's city and state were passively collected by the app when the participants viewed the dashboard tab. As of version 1.0.6 released May 5, 2016, the app began to send reports hourly whenever the user's location changed.

Quantitative variables for CDC and AHA baseline and clinical demographics. Demographic data was obtained from the 2013 Behavioral Risk Factor Surveillance System (BRFSS) (http://www.cdc.gov/asthma/most_recent_data.htm). Rates of hospitalization and emergency department visits come from the CDC/NCHS National Ambulatory Medical Care Survey, National Hospital Ambulatory Medicare Care Survey, National Hospital Discharge Survey, National Vital Statistics System, and National Health Interview Survey 2001–2009. CDC defined 'Uncontrolled asthma patients' as those who reported any of the following: (1) asthma symptoms more than 2 d a week in the past 30 d, (2) nighttime awakenings more than once a week in the past 30 d, or (3) short-acting β2-agonists use more than 2 d a week in the past 3 months (http://www.cdc.gov/asthma/asthma_stats/uncontrolled_asthma.htm). We compared Baseline users who supplied location data with national asthma prevalence statistics percent distribution by state (http://www.cdc.gov/asthma/brfss/2013/tableC1.htm). Percent prevalence was $\log_{10}$ transformed and states or territories with fewer than five AHA users were omitted from analysis. The correlation coefficient was calculated based on $n = 49$ states representing 4,612 Baseline users.

Quantitative variables for GINA (Global Initiative for Asthma) symptom control. 1) daytime symptoms occur less than twice per week (<8 times per month); 2) no occurrence of nocturnal awakenings (0 per month); 3) quick relief of symptoms occurs fewer than twice per week (<8 puffs per month); and 4) no activity limitation due to asthma symptoms (0 per month). Asthma is considered 'Uncontrolled' if four of the above statements are true, 'Partly controlled' if 2–3, and 'Well controlled' if 0–1 (http://ginasthma.org/).

Statistical methods for survival analysis of daily survey participation. We calculated the number of days between enrollment and the completion of the last daily survey question for each Robust cohort participant using data collected between the launch of the study, March 9, 2015, and September 8, 2015. To avoid bias from using the September 9, 2015, cutoff, we only considered participants with at least 90 d of study enrollment (participants' enrollment before June 9, 2015). This leads to 537 Robust cohort participants. Participants were treated as censored if a daily or weekly survey question was answered within 2 weeks of September 9, 2015.

The Cox proportional hazards model implemented in the Lifelines python package (version 0.8.0.0) was used to identify features associated with longer daily survey participation. We created the following predictor categories to increase sample size: education (no college, some college, college graduate) and health insurance (no health insurance, or private/public health insurance). All reported statistically significant associations exhibited a monotonic relationship with the enrollment day of the last daily survey question completed in the absence of grouping. We excluded 'decline to answer' responses and participants with missing predictor or outcome data. Features were not standard scaled to preserve interpretability of the hazard ratios; however, standard scaling did not change the direction or statistical significance of the reported associations. We used the KaplanMeierFitter function in the Lifelines python package to plot the survival curves for three strata of self-reported education levels and two strata of self-reported age of asthma diagnosis. The Seaborn python package (version 0.6.0) was used to generate heat maps for the predictor correlation matrix, which was calculated using the cor function of the Pandas package (version 0.17.1).

We first performed univariate survival analyses for nine covariates based on 537 Robust users. We then carried out a multivariate survival analysis of daily survey participation to adjust for collinearity (**Supplementary Fig. 2a**). The direction, magnitude, and statistical significance of the relationship between age and study entry date with study participation were Robust to different variable collapsing strategies and stratification by study entry month.

Statistical methods for individual response rate regression. We focused on the same set of 537 participants considered in the survival analysis of the retention time. We calculated the 'individual response rate', defined as the number of days with at least one daily survey question completed divided by the number

doi:10.1038/nbt.3826

of days enrolled through September 9, 2015. We then used ordinary least-squares regression (StatsModels version 0.6.1) to identify associations between the logit-transformed individual response rate.

Statistical methods for daily survey data. Daily survey responses collected on the same calendar day were combined in the following way. Any 'True' answer to questions about use of: control medication, night symptoms, day symptoms, or quick relief inhaler was used. The maximum reported value for questions about peak flow and quick relief puffs was used. The union of the reported asthma triggers reported on the same calendar day was used. A heatmap was generated using the matplotlib python library to show the self-reported data for day symptoms, night symptoms, and quick relief inhaler usage in daily surveys (**Supplementary Fig. 3**).

Statistical methods for concordance between GINA control based on baseline surveys and prospectively collected daily symptom reports during the study. The association between baseline GINA control group and prospectively collected daily symptom reports (% day symptoms, $n = 2{,}295$; % night symptoms, $n = 2{,}295$; quick relief usage frequencies, $n = 2{,}295$; and controller usage frequencies, $n = 2{,}285$) were evaluated with Kruskal–Wallis groupwise rank sum test. In addition, for each user, peak flow measures more than 900 and less than 60 were removed and were averaged across the study period. Then, a multiple linear regression model was used to evaluate the association between mean peak flow of each user during the study period and their baseline categorical GINA control groups, sex, height, and age ($n = 183$).

Statistical methods for concordance of survey responses within time-series. For each Robust user, we evaluated pair-wise Pearson correlations of user daily survey responses for peak flow, day and night symptoms (yes/no), and quick-relief usage (yes/no). We first filtered out users with fewer than ten pairs of observations. For each user, peak flow measures more than 900 and less than 60 were removed. Users whose survey responses had a s.d. of zero across the time-series for either pair of observations, or where the observations for each pair were exactly matching were also removed. The resulting distribution of Pearson correlations evaluated for each user for each pairwise comparison is shown in **Supplementary Figure 4**.

Statistical methods for delineation of United States into northern and southern regions based on temperature. Several app features were introduced after launch, including collection of encrypted user location data, which started on April 21, 2015. To compare data on asthma triggers and compare it to objective environmental measures, we first categorized 'user-date locations' as northern or southern based on the clustering of local temperature profiles, where user-date location refers to the latitude and longitude information of one user on one particular day. We obtained temperature data from the National Oceanic and Atmospheric Administration 4 (NOAA4), which collects daily maximum temperature from more than 53,000 US weather stations in the Global Historical Climatology Network (GHCDN).

To facilitate comparison with other environmental data sets, clustering was performed at the zip code level. First, each user-date location was linked to the nearest zip code and GHCDN weather station by minimum great circle distance using the 'fields' package in R. Latitude and longitude coordinates for US zip codes and weather stations were obtained from the NOAA4 and the US Census Bureau (https://www.census.gov/geo/maps-data/), respectively. User-date locations with nearest weather station or zip code more than 50 miles away were not included in subsequent analyses. In total, the 8,083 user locations collected during our study period map to 3,646 unique zip codes and 576 US weather stations.

Delineation of the US into northern and southern regions was based on hierarchical clustering of temperature profiles derived from the maximum daily temperature data corresponding to our 6-month study period. For each of 3,646 zip codes within 50 miles of any of the user locations collected during our study period, 37 features were constructed from the daily maximum temperature data collected at the corresponding US weather stations. Specifically, since not all of the user-date locations that link to the same zip code are linked to the same station, we first calculated the daily maximum temperature at each zip code by taking the mean daily maximum temperature over the set of stations linked to each zip code via user-date location as described above. We then derived our 37 features by taking the median of the daily maximum temperature at each zip code across 5-d intervals of our 185-d study period.

Hierarchical clustering was performed using the 'hclust' package in R. Specifically, we first calculated the pairwise Euclidean distances between each feature–zip code and then performed hierarchical clustering on the resulting distance matrix using the complete linkage method. For further analysis, user zip codes were assigned to the north or south based on the first bifurcation of the resulting dendrogram (**Supplementary Fig. 1b**). **Supplementary Figure 1c** shows the geographical regions associated with each cluster.

Statistical methods for analysis of asthma triggers. Users were asked to provide triggers of their asthma symptoms at the start of enrollment, and then on a daily basis throughout the study period. To compare time-series trigger data in different regions, we assigned each user-date location to the north or south based on the clustering procedure described above. For users who remained within the same region throughout the period observed, we extrapolated their regional assignment (north/south) to their full enrollment period. This resulted in 24,720 user-date locations corresponding to 545 unique users from the Robust user cohort, for further analysis. Trigger distributions were calculated based on the total number of triggers reported for a given time period (season, 5 d or 1 d). Periods for which the total number of user-date data points were fewer than 10 were treated as missing. A comprehensive depiction of time-series data for the northern and southern regions at baseline and throughout the study period is shown in **Supplementary Figure 5**, where the ordering of triggers from top to bottom is based on the percentage rank at baseline. Each curve was generated using the R function smooth.spline. Note that data for the last 16 d are omitted in **Supplementary Figure 5** because a very limited number of users in the southern region responded to the daily survey in those days.

Trigger distributions were compared with environmental data from several sources. Pollen count data were obtained from https://www.pollen.com, which provides monthly average pollen counts for zip codes throughout the United States. We obtained pollen count data for 49 zip codes corresponding to 53 users from the Robust user cohort with 5,601 user-date locations in the north and 16 zip codes corresponding to 11 users from the Robust user cohort with 123 user-date locations in the south. Zip codes were selected to ensure an adequate number of user-date locations ($n \geq 10$) for calculating the trigger distribution in the final half of our study period, where data for southern users are sparse. For the southern and northern regions, we calculated the average and the standard error of the pollen level across the corresponding zip codes for each region. The data are shown in **Figure 4b** as the dashed lines. In addition, in **Figure 4b**, the daily percentage of users reporting pollen as a trigger was plotted along time, where the running mean based ($n = 10$) and Bollinger bands (s.d. = 1) are shown.

In **Figure 4c**, the percentage of Robust users reporting extreme heat as their asthma trigger is compared with maximum daily temperature data obtained from the NOAA4. The percentage of users reporting extreme heat as an asthma trigger was calculated for 37 5-d intervals based on weather data from 492 weather stations corresponding to 432 users from the Robust user cohort with 19,031 user-date locations in the north and 108 weather stations corresponding to 113 users from the Robust user cohort with 8,347 in the south. **Figure 4c** shows running mean curves and Bollinger bands for maximum temperature and extreme heat trigger distributions based on a smoothing window of 11 and a s.d. of 1.

To investigate the effects of wildfire incidents occurring during our study period on AHA users, we collected data on wildfire locations and their start dates from the InciWeb incident information service (https://inciweb.nwcg.gov). We then searched for AHA users within a 200-mile radius of each fire. Using this procedure, we were able to find a total of 37 Robust users in proximity to Washington state wildfire locations, including the following: 35 users (499 userdate locations) near the Sleepy Hollow wildfire (6/28), 27 users (385 user-date locations) near the Wolverine Fires (7/29), 25 users (380 user-date locations) near the North Star wildfire (8/13), 29 users (398 user-date locations) near the Chelan Complex wildfire (8/14), 26 users (381) near the Turn Block wildfire (8/14), and 26 (381) near the Okanogan wildfire (8/15). In **Figure 4d**, we illustrate the percentage of triggers due to air quality complaints

for our study period along with EPA logs of air quality data (AQI) (http://www3.epa.gov/airdata/ad_data.html), where the maximum air quality index curve is based on a running average (n = 3). Bollinger bands based on a s.d. of 1 are shown.

Statistical methods for activity limitation. In the weekly survey, users are asked to report whether they have experienced activity limitation in the past week. We compared the proportion of users reporting activity limitation in the first week and the last week of their enrollment for 1,926 users from the Robust user cohort who were enrolled for longer than 90 d using the Wilcoxon signed rank test. We performed the same analysis restricting to the subset of users who were enrolled exclusively in the summer months (n = 331).

Sample size. Because the goals of the various analyses in the study differed significantly, different subsets of users (with specific characteristics) were carefully selected for each analysis. In **Figure 1**, we introduced a few different sample cohorts and sub-cohorts. In **Supplementary Table 1a**, we summarized sample sizes for the analyses corresponding to each of our main figures. We provide further details on the sub-cohort sample sizes used in all analyses below. Note that there were no pre-specified sample sizes in these analyses. All sample size numbers are based on observed data from the study.

1. In **Table 1** and **Supplementary Table 2**, demographic and clinical characteristics were summarized for Baseline, Robust, and Milestone users based on available data for each variable. See accompanying legends for sample sizes particular to each calculation.
2. Geographic distribution assessment was performed for 4,621 users from the Baseline user cohort who supplied their location information (**Fig. 1**).
3. The geographic distribution of n = 4,612 Baseline users was compared with national asthma prevalence statistics from the CDC for 49 US states (**Supplementary Fig. 1a**).
4. Association analysis to identify factors impacting the time of enrollment was performed based on Robust users with adequate data to determine biological sex (n = 719), baseline GINA category (n = 2,295), frequency of activity limitation (n = 2,308), and symptoms (n = 2,308) (**Fig. 2a–c**).
5. Detection of factors affecting user retention patterns and response rates was carried out based on 537 users from the Robust user cohort, who were enrolled in the study for >90 d and provided data for all the covariates considered in the analysis (**Supplementary Fig. 2d**,**e** and **Supplementary Table 4a**,**c**).
6. Association between patients' daily survey responses and their baseline GINA categories were evaluated based on subsets of the Robust user cohort reporting daily (n = 2,295) and nightly symptoms (n = 2,295), quick-relief inhaler usage (n = 2,295), and controller medicine usage (n = 2,285) (**Fig. 3a**,**d**).
7. Regression analysis of peak flow information was based on data from 183 users from the Robust user cohort who voluntarily submitted at least one peak flow measurements during the study period, and all covariates (e.g., age of onset, gender, height) in the analysis (**Fig. 3e**).
8. Concordance analysis of reported asthma symptoms, rescue inhaler use, and peak flow measurements within time-series was based on subsets of Robust user cohort, who provided their corresponding information. Specifically, correlations between daily/nightly symptoms and rescue inhaler use were based on n = 979 and n = 761 users, respectively. Correlations between daily and nightly symptoms and between peak flow and puff usage were based on n = 817 and n = 217 users, respectively. And correlations between daily/nightly symptoms and peak flow values were based on n = 235 and n = 173 users, respectively. (**Supplementary Fig. 4**).
9. Geographic distribution of asthma triggers were evaluated based on 545 users of the Robust user cohort, who, throughout their respective enrollment periods, could be consistently mapped to one of two geographical regions (north/south) based on time series location data (**Fig. 4a** and **Supplementary Fig. 1b**,**c**).
10. Assessment of changes of activity limitation during the study period was based on 1,926 users who are from the Robust user cohort and had been enrolled for >90 d (**Fig. 5a**). In addition, 331 of the 1,926 users, who supplied enough data during the summer, were used to assess the impact of app usage on users' activity limitation during summer months (**Fig. 5a** and **Supplementary Table 5a**,**b**).
11. 173 of the 175 users in the Milestone user cohort supplied complete data to derive their GINA categories at the enrollment and after 6-month's App usage. Data of these 173 were used to evaluate the impact of 6-month's App usage on users' GINA categories (**Fig. 5b** and **Supplementary Table 5d**).
12. Feedback and milestone survey results were evaluated for Milestone users who replied to questions about whether the app helped them to alleviate their troubles with asthma (n = 175), achieve their initial goals for asthma control (n = 172), and manage their asthma (n = 168), (**Fig. 5c**).
13. The relationship between GINA category and milestone survey feedback indicating whether the app helped to prevent visits to the emergency department or doctor was evaluated for 125 and 127 Milestone users, respectively (**Supplementary Table 5d**,**e**).

doi:10.1038/nbt.3826