







OPEN

DATA DESCRIPTOR

# Home monitoring with connected mobile devices for asthma attack prediction with machine learning

Kevin C. H. Tsang<sup>1,2</sup>  , Hilary Pinnock<sup>1</sup>, Andrew M. Wilson<sup>1,3,4</sup>, Dario Salvi<sup>5</sup>   
& Syed Ahmar Shah<sup>1,2</sup> 


Monitoring asthma is essential for self-management. However, traditional monitoring methods require high levels of active engagement, and some patients may find this tedious. Passive monitoring with mobile-health devices, especially when combined with machine-learning, provides an avenue to reduce management burden. Data for developing machine-learning algorithms are scarce, and gathering new data is expensive. A few datasets, such as the Asthma Mobile Health Study, are publicly available, but they only consist of self-reported diaries and lack any objective and passively collected data. To fill this gap, we carried out a 2-phase, 7-month AAMOS-00 observational study to monitor asthma using three smart-monitoring devices (smart-peak-flow-meter/smart-inhaler/smartwatch), and daily symptom questionnaires. Combined with localised weather, pollen, and air-quality reports, we collected a rich longitudinal dataset to explore the feasibility of passive monitoring and asthma attack prediction. This valuable anonymised dataset for phase-2 of the study (device monitoring) has been made publicly available. Between June-2021 and June-2022, in the midst of UK's COVID-19 lockdowns, 22 participants across the UK provided 2,054 unique patient-days of data.

## Background

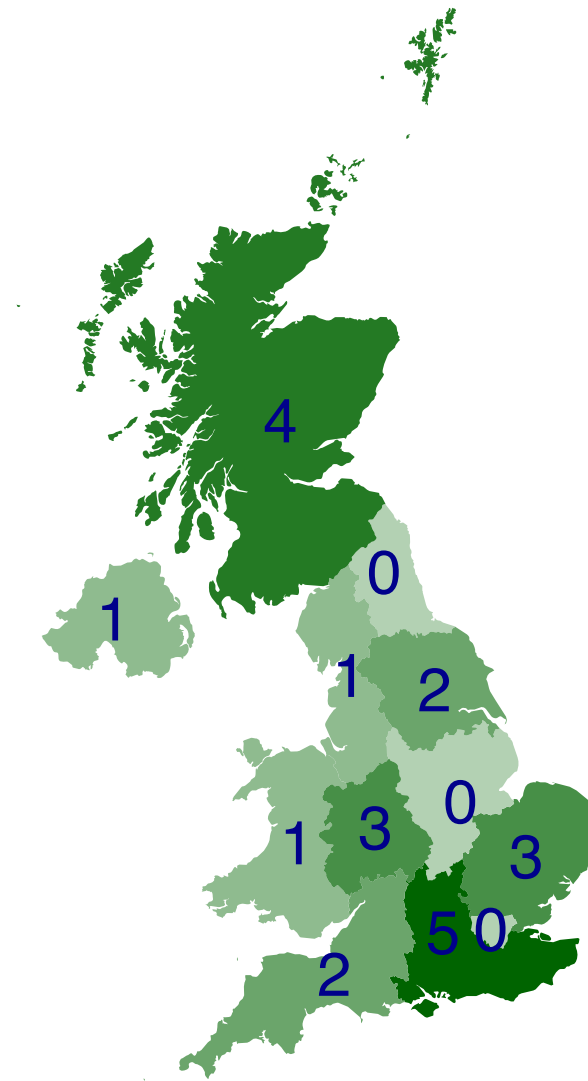
Asthma is a long-term condition that affecting around 5.4 million people in the UK and its impact on daily life can vary from day-to-day<sup>1</sup>. Since there is no known cure for asthma, self-management is a key part of patient care; this involves detecting deterioration and taking appropriate action to maintain control and prevent the threatened attack<sup>2</sup>. Traditional self-management action plans use symptom scores, sometimes supplemented by peak expiratory flow measurements, to determine a patient's asthma condition<sup>3-5</sup>. Keeping track of relief inhaler usage can also help measure asthma control<sup>6</sup>.

However, patients may regard this level of monitoring as tedious as it involves high levels of active engagement on their part. This is especially true when they are feeling well, because traditional monitoring is active and may yield repetitive readings. One method of alleviating some management burden is to reduce the manual labour associated with monitoring of asthma condition<sup>7</sup>. Passive monitoring (i.e. collecting data with minimum active user engagement), such as wearing a smartwatch, can provide an avenue to reducing the burden of monitoring if it is correctly used to give timely alerts<sup>8</sup>.

Three recent studies that have investigated the use of mHealth for asthma management were the Asthma Mobile Health Study (AMHS)<sup>9,10</sup>, myAirCoach<sup>11</sup>, and Biomedical REAL-Time Health Evaluation (BREATHE)<sup>12</sup>. The AMHS is the largest mHealth study for asthma conducted to date, containing large amounts of cross-sectional and longitudinal data. However, apart from the possibility of linking location data with historical weather reports, the dataset did not include any additional objective data (e.g., data from wearables, peak flow meters, or smart inhalers). To our knowledge, the dataset from myAirCoach and BREATHE are not publicly available and they have not yet been used to test any machine learning-based algorithms for asthma attack prediction<sup>12-14</sup>.

<sup>1</sup>Asthma UK Centre for Applied Research, Usher Institute, University of Edinburgh, Edinburgh, UK. <sup>2</sup>Medical Informatics, Usher Institute, University of Edinburgh, Edinburgh, UK. <sup>3</sup>Norwich Medical School, University of East Anglia, Norwich, UK. <sup>4</sup>Norwich University Hospital Foundation Trust, Colney Lane, Norwich, UK. <sup>5</sup>Internet of Things and People Research Centre, Malmö University, Malmö, Sweden. e-mail: [kevin.tsang@ed.ac.uk](mailto:kevin.tsang@ed.ac.uk); [ahmar.shah@ed.ac.uk](mailto:ahmar.shah@ed.ac.uk)

## AAMOS-00 Participants



**Fig. 1** AAMOS-00 participants locations. Participants were located across all four nations of the UK.

The AAMOS-00 observational study was carried out with the aim to provide a rich multi-dimensional dataset to develop better asthma attack prediction algorithms<sup>15</sup>. It combined market-available devices and application programming interfaces (APIs) to investigate the feasibility of a single mHealth system which pulls together data from sources available to asthma patients. Furthermore, the study included objective data sources (that can be collected passively) to complement current methods in asthma monitoring for self-management.

## Methods

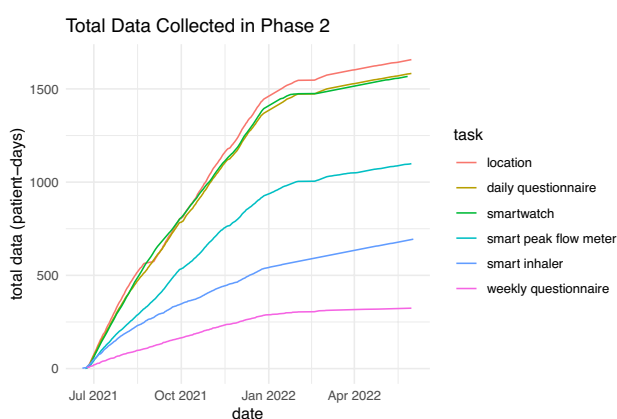
**Data summary.** A total of 22 patients participated in phase 2 of the study. Across 12 months, between 24<sup>th</sup> June 2021 and 2<sup>nd</sup> June 2022, 2,054 patient-days of data in phase 2. All participants of phase 2 agreed to share their anonymised data<sup>16</sup>. The average retention was 123 days (67%) in phase 2. The participants in phase 2 were mostly female (77%), average age of 40 years, mostly white (95%), and most (95%) had experienced an asthma attack in the past 12 months<sup>16</sup> (see Table 1).

A course of oral corticosteroids (OCS) for an asthma attack in the past 12 months was an inclusion criterion for the study. This was asked twice, once before accessing the consent forms, and once at the start of the AAMOS-00 study (reported in Table 1). A potential explanation for observing one patient who did not have a course of OCS in the past 12 months, was that the patient had a course of OCS nearly a full 12 months ago, then had some delay with downloading Mobistudy and starting data collection. Thus, the number of patients who had a course of OCS for an asthma attack in the past 12 months was not 100%.

Participants were located across all four nations of the UK (England, Northern Ireland, Scotland, and Wales), with a majority (73%) from England<sup>16</sup> (see Fig. 1).

Characteristics	AAMOS-00 Phase 2 (n = 22)
<b>Sex, n (%)</b>	
Female	17 (77%)
Male	5 (23%)
<b>Age, median (IQR)</b>	40.2 years old (15.7 years old)
<b>BMI, mean (SD)</b>	27.7 kg/m <sup>2</sup> (5.7 kg/m <sup>2</sup> )
<b>Race (White), n (%)</b>	21 (95%)
<b>Smoker, n (%)</b>	
Never (<100 cigarettes)	17 (77%)
Previous	5 (23%)
<b>Had hospitalisations in past 12 months, n (%)</b>	6 (27%)
<b>Had a course of OCS for asthma attack in past 12 months, n (%)</b>	21 (95%)
<b>RCP3 in past month, mean</b>	2.4

**Table 1.** AAMOS-00 phase 2 participant characteristics. Royal College of Physicians “3 Questions” (RCP3) score ranges from 0 to 3, 0 indicating good control, 1–3 indicating poor control<sup>17</sup>.



**Fig. 2** Total data collected.

In phase 2, a total of 1,583 daily questionnaires<sup>16</sup> and 324 weekly questionnaires<sup>16</sup> were collected. There were 694 patient-days of smart inhaler<sup>16</sup>, 1,567 patient-days of smartwatch data<sup>16</sup>, 1,099 patient-days of peak expiratory flow (PEF) recordings via the smart peak flow meter<sup>16</sup>, and 1,657 patient-days of sent locations<sup>16</sup> (see Fig. 2). The total patient-days of relief puffs was relatively low since patients who spent a day without using the relief inhaler did not add to this count.

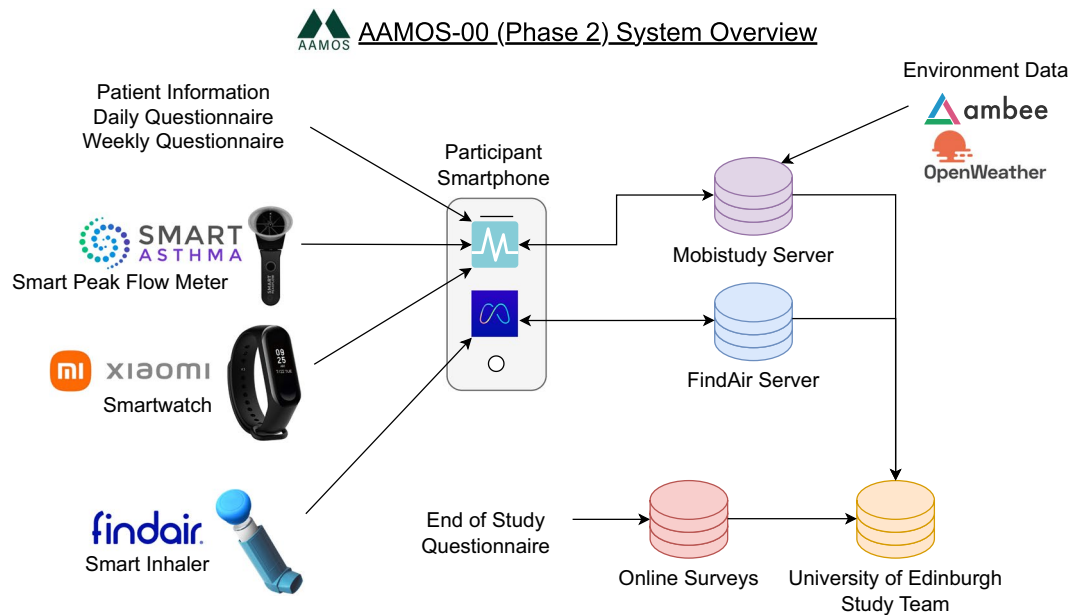
**Study design.** The AAMOS-00 observational study conducted data collection over 14 months from April 2021 to June 2022. The study had two phases.

Phase 1 was one month of daily and weekly questionnaire monitoring, which was used to select participants who were likely to adhere to monitoring for a long duration in phase 2. The seven-item daily questionnaire combined the AMHS<sup>9</sup> daily questionnaire with the Royal College of Physicians “3 Questions” (RCP3)<sup>17</sup> questions to measure daily asthma control, triggers encountered, and medication usage (Supplementary Information). The AAMOS-00 study’s 11-item weekly questionnaire had incorporated aspects of the RCP3<sup>17</sup>, Asthma Control Questionnaire (ACQ)<sup>4</sup>, and AMHS<sup>9</sup> weekly questionnaire to capture the asthma control and symptoms displayed in the past week, and any unscheduled care (Supplementary Information).

Phase 2 was six months of smart device monitoring, where participants received three smart monitoring devices (smartwatch, smart peak flow meter, and smart inhaler) to collect data daily in addition to continuing daily questionnaires. Furthermore, daily location was used to link with weather, air quality, and pollen reports in the local area. The data collection was carried out via the Mobistudy app, which integrated all data collection apart from the FindAir smart inhaler (see the “Mobile Monitoring Technology” section for more details about the implementation and Fig. 3 for an overview of the system architecture).

Study participants of phase 2 had a total of four daily tasks (daily questionnaire, send location, morning and evening peak flow reading), two weekly tasks (smartwatch data upload and weekly questionnaire), and a passive monitoring task (smart relief inhaler usage).

At the end of phase 2, we asked participants to complete a questionnaire about the acceptability and usability of the system (Supplementary Information), to assess the current implementation and help future development. The three-part questionnaire used three validated questionnaires: the System Usability Scale (SUS)<sup>18</sup> assessed usability, the mHealth Technology Engagement Index (mTEI)<sup>19</sup> assessed personal motivation to use technology



**Fig. 3** AAMOS-00 system overview. Centred around the participant's own smartphone, three smart monitoring devices (smart peak flow meter, smartwatch, smart inhaler) collected objective data and two API services provided information about local environment (weather, air quality, pollen) based on the participant's location.

for self-management, and the User version of Mobile Application Rating Scale (uMARS)<sup>20</sup> assessed app quality and perceived impact. We adapted some questions to better reflect the AAMOS-00 study system.

**Recruitment.** Participants across the UK were recruited via social media, invitation letters from Norfolk and Norwich University Hospitals (NNUH), and email invitations. Participants were adults with asthma who had experienced an asthma attack in the past 12 months (definition: ATS/ERS Task Force 2009)<sup>21</sup> and were prescribed with a pressurised metered dose relief inhaler that was compatible with FindAir ONE.

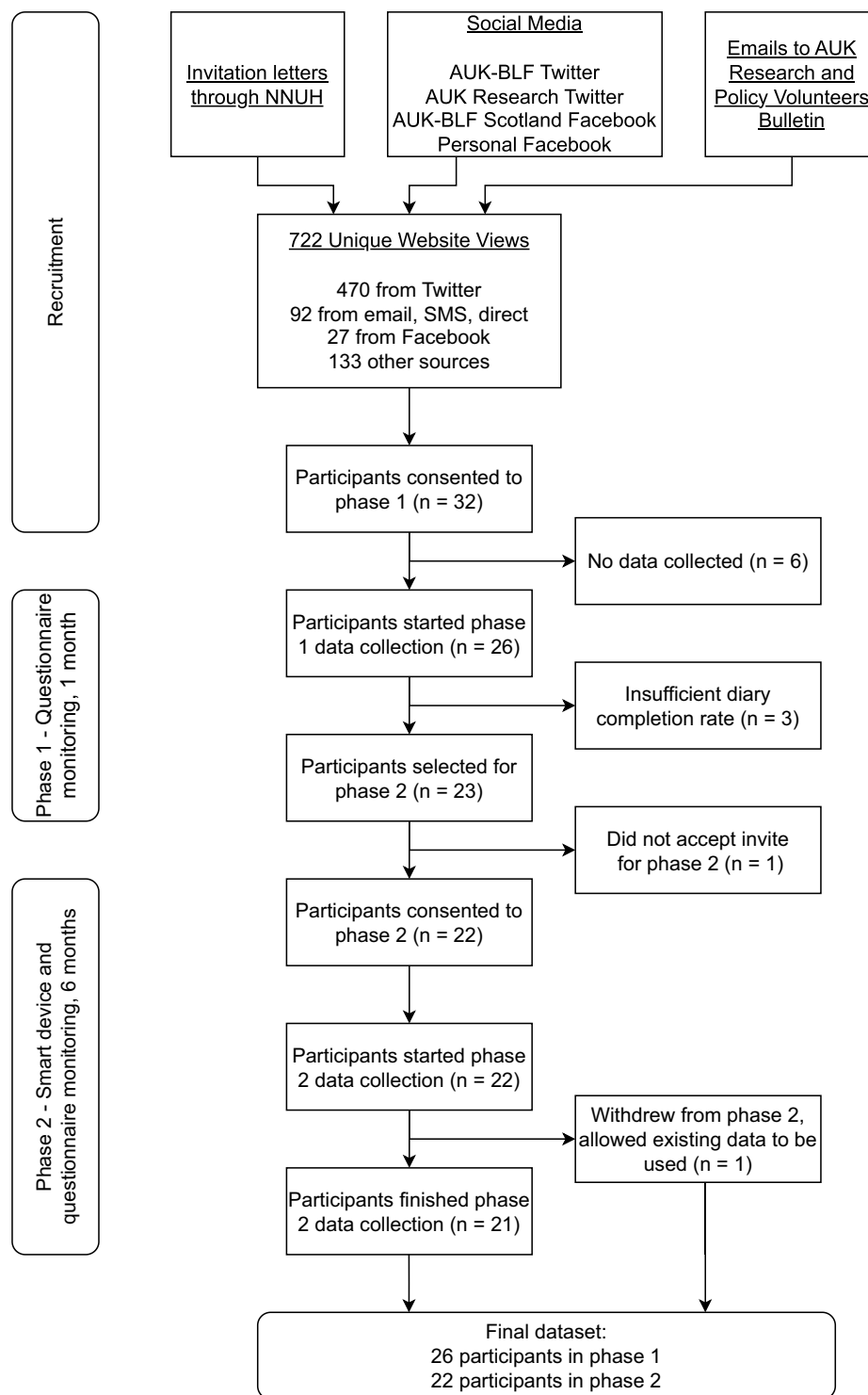
Social media recruitment consisted of disseminating invitations to the public via Twitter and Facebook via the Asthma + Lung UK and Asthma UK Centre for Applied Research (AUKCAR) accounts, which total around 175,000 followers. The Norfolk and Norwich University Hospital helped identify potentially eligible patients for the study and had sent them invitation letters. Email invitations were sent via the Asthma UKs Research and Policy Volunteers Bulletin, which is a channel to circulate research opportunities conducted by the Asthma + Lung UK to volunteers. Patients who were interested in joining the study were directed to the recruitment website, where they found the participant information sheets and the online consent form hosted on Online Surveys.

Over the 10-month recruitment period from 15<sup>th</sup> February 2021 to 1<sup>st</sup> December 2021, 32 participants were recruited to phase 1 of the study (see Fig. 4). After one-month of data collection with daily questionnaires, 23 participants who had completed at least half of the requested daily questionnaires (14 of 28 days) were selected and invited to join the device monitoring portion of the study (phase 2). One participant declined the invitation. Twenty-two participants collected data for six months and one participant pulled out of the study during phase 2, citing frustration with the technology.

**Mobile monitoring technology.** *Mobistudy.* Mobistudy<sup>22</sup> is an open-source platform for facilitating mobile-based studies, it is managed by Malmö University, Sweden. The platform has three key components: a mobile app for participants (available for Android and iOS), a representational state transfer (REST) API server, and an online web portal for researchers to view the data in real time. The platform supports multiple studies and participants of the AAMOS-00 study were given a study invite code to join the AAMOS-00 study within Mobistudy.

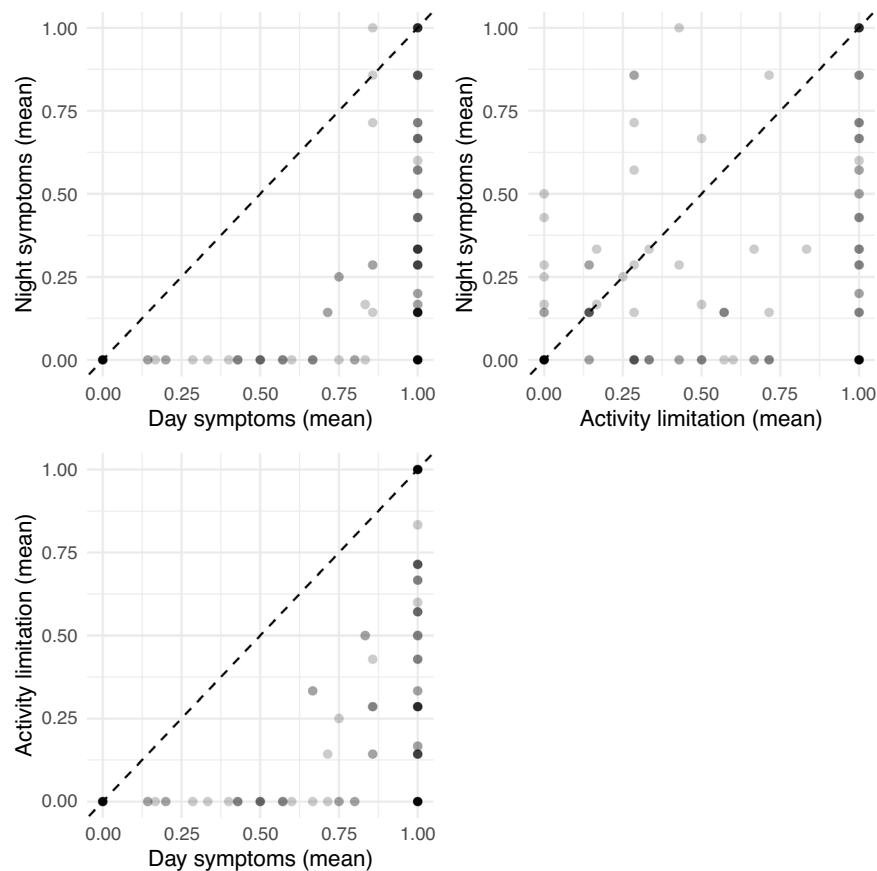
The Mobistudy mobile app was central to data collection where each daily and weekly assessment (questionnaires, peak flow measurement, smartwatch data upload, and sending location) appeared as an individual task of the home page on the participant's app.

*Smart Asthma Smart Peak Flow Meter.* The Smart Peak Flow Meter by Smart Asthma ([www.smartasthma.com](http://www.smartasthma.com)) is an affordable (£34.99) mHealth peak flow meter that measures peak flow (PEF) with the help of the processing in the smartphone. There are two modes of connection between the device and the smartphone, by direct audio jack connection or by a Bluetooth adapter. The device is a Class 2a medically certified device and has been tested and validated against a pulmonary waveform generator, deemed to pose a low to medium risk to patients and thus complies with the UK's and EU's safety and performance standards<sup>23–25</sup>. The device connected directly to the Mobistudy app, and the signal was translated into peak flow reading via the integrated library provided by Smart Asthma.



**Fig. 4** AAMOS-00 participants engagement. Phase 1 was one month of daily questionnaire monitoring. Phase 2 was six months of smart device and questionnaire monitoring.

*Xiaomi Miband3 smartwatch.* The MiBand3 by Xiaomi ([www.mi.com](http://www.mi.com)) is an affordable (£25.00) smartwatch that can be used to monitor activity and heart rate. The CE (Conformité Européenne) marked device is light-weight and includes a touch screen where navigate to different screens such as total steps today, battery, exercise mode, and current heart rate. Four minute-by-minute signals were collected with the MiBand3, heart rate reading, activity type, activity intensity, and total steps. The data upload from the smartwatch used a Bluetooth connection to Mobistudy and an integrated library.



**Fig. 5** Comparison of daily answers to questions for asthma control averaged over a week (0 = no symptoms recorded on any day of the week, 1 = symptoms present on all days of the week). In line with previous understanding<sup>52</sup>, the day symptoms were almost always higher than night symptoms and activity limitation, whereas nocturnal symptoms and activity limitation were relatively independent.

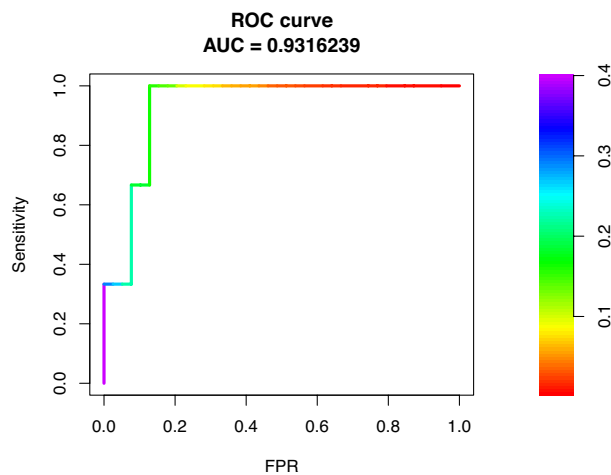
*FindAir ONE smart inhaler.* The FindAir ONE by FindAir ([www.findair.eu](http://www.findair.eu)) is a smart inhaler (also known as a Bluetooth cap) for pressurized metered-dose inhalers (pMDI) inhalers, to track when the inhaler is used, priced at €59.00 per year. The CE marked device has a battery life of 12 months after first use and is un-rechargeable. Once attached to the inhaler, the smart inhaler automatically logs actuations when the inhaler is used. When the device is connected to the smartphone via Bluetooth, the stored data is transferred to the smartphone and then FindAir's server. We used a secure REST API connection to transfer data between the FindAir server to the study's server (see Fig. 3).

*Open Weather Maps and Ambee.* Based on the location of the participant when completing the location task, the local weather, air quality, and pollen count were fetched using Open Weather Maps<sup>26</sup> and Ambee's<sup>27</sup> APIs (see Fig. 3). The information included weather, temperature, humidity, cloud cover, wind, air quality index (AQI)<sup>28</sup>, and grass, tree, and weed pollen count.

*Online Surveys.* The study consent forms and the exit questionnaire about usability and acceptability were hosted on Online Surveys (<https://www.onlinesurveys.ac.uk/>), which is an online service where participants can visit a webpage to complete the questionnaires and the data is then securely held by their servers. Afterwards, the responses were transferred to the study's servers.

*Data anonymisation.* The directly identifiable information fields<sup>29</sup> were removed or replaced. These included name, dates (including date of birth and date of data entry), location, height, weight, medication used, and user key.

- The names of participants were removed.
- The age in years was calculated at the end date of the study, which replaced the participant's date of birth information. Furthermore, the age was reduced in granularity via the use of age ranges.
- Likewise, only the age range of the age of asthma diagnosis was made available. The ranges were early childhood (0 to 6 years old), late childhood (7 to 11 years old), adolescence (12 to 18 years old), and late onset (19+ years old)<sup>30</sup>.



**Fig. 6** ROC curve of random forest classifier. AUC = 0.93 and AUPRC = 0.55 suggesting a strong signal was present. Data had 15 positive class and 159 negative class.

- The daily locations of participants were not made available, but the local environmental data (weather, air quality, and pollen count) collected during the study were made available. A single location of participants at the UK region level was not identifiable and was made available. The information would be sufficient to link localised historic weather data.
- Body mass index (BMI), calculated from height and weight, and theoretical maximum PEF, calculated from height, age, and sex, were known risk factors of asthma attacks<sup>3,4</sup> and important measures. The BMI range and theoretical maximum PEF rounded to the nearest 5 were made available.
- The list of medication used by patients were removed.
- The participant user keys were replaced with a new random number between 100 and 999.
- All dates of data entry in the dataset were removed. The dates of data entry were transformed to the number of days after each participant started phase 2.

Patient sex and race were made available, because there are known sex and ethnic differences in asthma<sup>31–33</sup>. They were indirectly identifiable information and considered to have a low risk of deanonymisation.

**Data pre-processing.** The published dataset was produced by combining raw JSON, CSV, and XSLX files from the aforementioned data sources. All the data pre-processing was conducted using R (v4.2.1)<sup>34</sup> and the following packages: fuzzyjoin<sup>35</sup>, gtools<sup>36</sup>, janitor<sup>37</sup>, jsonlite<sup>38</sup>, lubridate<sup>39</sup>, plyr<sup>40</sup>, PostcodesioR<sup>41</sup>, qdapTools<sup>42</sup>, rapportools<sup>43</sup>, readxl<sup>44</sup>, tidyverse<sup>45</sup>, zoo<sup>46</sup>.

**Ethics.** Ethics approval was provided by the East of England - Cambridge Central Research Ethics Committee. IRAS (Integrated Research Application System) project ID: 285505 with governance approval from ACCORD (Academic and Clinical Central Office for Research and Development), project number: AC20145. The study sponsor is ACCORD, the University of Edinburgh. The anonymised data has been made available with the consent of each participant.

**Limitations.** A major limitation of the AAMOS-00 study was the narrow inclusion criteria, which selected asthma patients who had an interest in monitoring and had experienced a severe asthma attack in the past 12 months. Although the dataset contains limited patients, there are over 2000 unique patient days of longitudinal multi-dimensional data. Speaking with patient and public involvement (PPI) members, we believed the average retention to daily monitoring in the general population would be substantially lower than what we observed in this study. However, the average retention in phase 2 was 123 days, which included daily tasks, this was much longer than our initial estimate.

Due to the technical implementation of using a mobile phone app to accommodate the general population, the passive monitoring devices of location and smartwatch require daily and weekly active engagement with the app. This meant that this study could not explore the potential of completely passive monitoring devices that require no user intervention. Furthermore, the technical issues encountered by some patients could have affected the adherence to monitoring with the smart devices.

The AAMOS-00 study was conducted over several periods of national lockdowns in the UK due to the COVID-19 pandemic<sup>47</sup>. During this unique time period, the general public had spent more time indoors, had facemask wearing, kept social distancing, and had reduced road traffic<sup>48</sup>, which reduced the exposure to some triggers such as virus, pollutants, and outdoor allergens. The effect of lockdowns was a substantial reduction in asthma attack rates<sup>49</sup>. Comparing this data with other dataset collected during times without lockdowns could provide further insights into how a drastic societal change has affected asthma.

## Data Records

Researchers are able to download the anonymised data from phase 2 of the AAMOS-00 study<sup>16</sup> via Edinburgh DataStore (a digital repository of research data produced at the University of Edinburgh) (<https://datashare.ed.ac.uk/handle/10283/4761>). The dataset was designed for longitudinal analysis, linking the different signals by “user\_key” and “date” would give the most holistic view of the patients. The date begins at 1 for each patient, this represents the first day of data entry in phase 2 for each patient. Each increment is a calendar day.

The dataset contains a separate file for each data signal. The list of files provided are as follows:

- Data Documentation – contains a information about the dataset and the study
- Data Dictionary – contains the data definition for all variables
- Daily Questionnaire Data – includes the responses to the daily questionnaire (including symptoms, medication use, triggers encountered)
- End of Study Questionnaire Data – includes the usability and acceptability score collected at the end of the study
- Environment Data – includes the daily weather, pollen, air pollution data
- Patient Information – includes the patient information (including age, sex, region) collected at the start of the study
- Smart Peak Flow Meter Data – includes twice-daily measurements of peak flow measured via the smart peak flow meter
- Smart Inhaler Data – includes relief inhaler usage information (timestamp and medication name) collected via smart inhaler
- Smartwatch Data 1 (data entries 1-1,000,000) – includes the minute-by-minute data from the smartwatch (including activity type, intensity, steps taken, heart rate), set one
- Smartwatch Data 2 (data entries 1,000,001-2,000,000) – set two of the smartwatch data, continues from set one
- Smartwatch Data 3 (data entries 2,000,001-2,101,829) – set three of the smartwatch data, continues from set two
- Weekly Questionnaire Data – includes the responses to the weekly questionnaire (including asthma control, symptoms, and any unscheduled care)

Further details about all the variables in each data file can be found in the Data Dictionary<sup>16</sup> (Supplementary Table).

## Technical Validation

During the development of Mobistudy and the integration of external sensors, standard software engineering techniques for quality assurance were put in place, such as extending testing and, when feasible, automated tests. Sensors integrated algorithms to extract physiological measurements from raw data, either in their internal firmware (MiBand 3) or as software libraries (Smart Peak Flow meter), which were integrated into the app. All smartwatches underwent a software update, to ensure that the latest firmware version (v2.4.0.32) was installed.

Our research team, comprising clinicians, checked that all values from questionnaires and devices were clinically plausible. The smart peak flow meters provided readings within our expectation. Only two peak flow measurements were outside theoretical range of values (based on age, sex, and height)<sup>50</sup>. For both readings, a valid peak flow reading taken within minutes replaced the outlier value. From the smartwatch, the mean sleep duration per day was 7.6 hours (interquartile range of 3.1 hours), which was within the expected range. Also from the smartwatch, the mean heart rate was 82 BPM, which was high but within physiological range (normal resting heart rate is between 50 and 90 BPM)<sup>51</sup>.

The daily asthma control questions displayed expected characteristics. As reported by Pinnock *et al.*<sup>52</sup>, the day-time symptoms alone is less of an indication of poor control than nocturnal symptoms or activity limitation alone. The AAMOS-00 data supports the finding. In general, the day symptoms were almost always higher than night symptoms and activity limitation, whereas nocturnal symptoms and activity limitation were relatively independent (see Fig. 5). As a feature, this means that day symptoms were superseded by the information of night symptoms and activity limitation.

Reframing this data as binary classification problem, we attained a high area under the receiver operating characteristic (ROC) curve (AUC) and area under the precision-recall curve (AUPRC) to classify weeks where patients attended unscheduled asthma doctor appointments using daily data, suggesting the classifier performs well using this data. Data processing was used to extract the weekly mean value of daily data in the dataset (from the daily questionnaire, environment, smartwatch, smart peak flow meter, and smart inhaler). This formed a dataset with 15 observations in the positive class (attended unscheduled appointment during the week) and 159 observations in the negative class. Then using an 80%-20% training-test split, we trained a random forest classifier and achieved good performance (AUC = 0.93 and AUPRC = 0.55) (see Fig. 6).

## Usage Notes

The dataset is licensed under the Creative Commons Attribution 4.0 International Public License (CC BY 4.0).

To download the dataset, visit the DataShare page (<https://datashare.ed.ac.uk/handle/10283/4761>).

R scripts have been provided to assist the usage of this dataset, including joining data tables, data wrangling, and an example binary classification problem.



## Code Availability

The Mobistudy version 0.2.6 used in the AAMOS-00 study can be found at <https://github.com/Mobistudy>. The software to translate the smart peak flow meter signal into peak flow was integrated with Mobistudy using a Cordova plugin, the code can be found on GitHub: <https://github.com/kevinchtsang/cordova-plugin-spf>. The smartwatch was integrated into Mobistudy based on the open-source work by Volodymyr Shymanskyi (<https://github.com/vshymanskyi/miband-js>), José Rebelo, and Gadgetbridge ([www.gadgetbridge.org](http://www.gadgetbridge.org)).

R scripts to illustrate joining the data tables and forming a binary classification problem can be found at <https://github.com/kevinchtsang/AAMOS-00-Starter>.

Received: 24 October 2022; Accepted: 15 May 2023;

Published online: 08 June 2023

## References

1. Asthma UK. *Asthma Facts and Statistics*. <https://www.asthma.org.uk/about/media/facts-and-statistics/> (2021).
2. Pinnock, H. *et al.* Systematic Meta-Review of Supported Self-Management for Asthma: A Healthcare Perspective. *BMC Med.* **15**, 64 (2017).
3. Scottish Intercollegiate Guidelines Network (SIGN)/ British Thoracic Society (BTS). *SIGN 158 British Guideline on the Management of Asthma*. (Edinburgh: SIGN, 2019).
4. Juniper, E. F., O'Byrne, P. M., Guyatt, G. H., Ferrie, P. J. & King, D. R. Development and Validation of a Questionnaire to Measure Asthma Control. *Eur. Respir. J.* **14**, 902–907 (1999).
5. Tsang, K. C. H., Pinnock, H., Wilson, A. M. & Shah, S. A. Application of Machine Learning to Support Self-Management of Asthma with mHealth. in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)* 5673–5677 <https://doi.org/10.1109/EMBC44109.2020.9175679> (IEEE, 2020).
6. 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).
7. May, C. R., Montori, V. M. & Mair, F. S. We Need Minimally Disruptive Medicine. *BMJ* **339**, 485–487 (2009).
8. Tsang, K. C. H., Pinnock, H., Wilson, A. M. & Shah, S. A. Application Of Machine Learning Algorithms for Asthma Management with mHealth: A Clinical Review. *J. Asthma Allergy* **15**, 855–873 (2022).
9. Chan, Y.-F. Y. *et al.* The Asthma Mobile Health Study, a Large-Scale Clinical Observational Study Using ResearchKit. *Nat. Biotechnol.* **35**, 354–362 (2017).
10. Chan, Y.-F. Y. *et al.* The Asthma Mobile Health Study, Smartphone Data Collected Using ResearchKit. *Sci. Data* **5**, 180096 (2018).
11. Honkoop, P. J. *et al.* MyAirCoach: The Use of Home-Monitoring And mHealth Systems to Predict Deterioration in Asthma Control and the Occurrence of Asthma Exacerbations; Study Protocol of an Observational Study. *BMJ Open* **7**, e013935 (2017).
12. Bui, A. A. T. *et al.* Biomedical REAL-Time Health Evaluation (BREATHE): Toward an mHealth Informatics Platform. *JAMIA Open* **3**, 190–200 (2020).
13. Khushial, R. J. *et al.* Effectiveness of myAirCoach: A mHealth Self-Management System in Asthma. *J. Allergy Clin. Immunol. Pract.* **8**, 1972–1979.e8 (2020).
14. myAirCoach. *D7.6 Data Management Plan (Second Year Update)*. <https://www.myaircoach.eu/content/d76-data-management-plan-second-year-update> (2016).
15. Tsang, K. C. H., Pinnock, H., Wilson, A. M., Salvi, D. & Shah, S. A. Predicting Asthma Attacks Using Connected Mobile Devices and Machine Learning: the AAMOS-00 Observational Study Protocol. *BMJ Open* **12**, e064166 (2022).
16. Tsang, K. C. H., Pinnock, H., Wilson, A. M., Salvi, D. & Shah, S. A. AAMOS-00 Study: Predicting Asthma Attacks Using Connected Mobile Devices and Machine Learning. *Edinburgh DataShare*. <https://doi.org/10.7488/ds/3775> (2022).
17. Pearson, M. & Bucknall, C. Measuring Clinical Outcome in Asthma: A Patient-Focused Approach. *R. Coll. Physicians London* (2000).
18. Brooke, J. SUS: A 'Quick and Dirty' Usability Scale. in *Usability Evaluation In Industry* vol. **189** 207–212 (CRC Press, 1996).
19. Dewar, A. R., Bull, T. P., Malvey, D. M. & Szalma, J. L. Developing a Measure of Engagement with Telehealth Systems: The mHealth Technology Engagement Index. *J. Telemed. Telecare* **23**, 248–255 (2017).
20. Stoyanov, S. R., Hides, L., Kavanagh, D. J. & Wilson, H. Development and Validation of the User Version of the Mobile Application Rating Scale (uMARS). *JMIR mHealth uHealth* **4**, e72 (2016).
21. Reddel, H. K. *et al.* An Official American Thoracic Society/European Respiratory Society Statement: Asthma Control and Exacerbations - Standardizing Endpoints for Clinical Asthma Trials and Clinical Practice. *Am. J. Respir. Crit. Care Med.* **180**, 59–99 (2009).
22. Salvi, D. *et al.* Mobistudy: Mobile-Based, Platform-Independent, Multi-Dimensional Data Collection for Clinical Studies. in *11th International Conference on the Internet of Things* 219–222 <https://doi.org/10.1145/3494322.3494363> (ACM, 2021).
23. VanZeller, C., Williams, A. & Pollock, I. Comparison of Bench Test Results Measuring the Accuracy of Peak Flow Meters. *BMC Pulm. Med.* **19**, 74 (2019).
24. Commission, E. CE marking. *European Commission* [https://ec.europa.eu/growth/single-market/ce-marking\\_en](https://ec.europa.eu/growth/single-market/ce-marking_en) (2022).
25. Gov, U. K. Medical Devices: Conformity Assessment and the UKCA Mark. *GOV.UK* <https://www.gov.uk/guidance/medical-devices-conformity-assessment-and-the-ukca-mark> (2020).
26. OpenWeatherMaps. *OpenWeatherMap*. <https://openweathermap.org/> (2021).
27. Ambee. *Ambee*. <https://www.getambee.com/> (2021).
28. Bartelds, H., Heich, H. & Léger, K. *CAQI Air Quality Index: Comparing Urban Air Quality across Borders-2012. Common Information to European Air* [www.citeair.eu](http://www.citeair.eu) (2012).
29. Hrynaskiewicz, I., Norton, M. L., Vickers, A. J. & Altman, D. G. Preparing Raw Clinical Data for Publication: Guidance for Journal Editors, Authors, and Peer Reviewers. *BMJ* **340**, c181–c181 (2010).
30. Trivedi, M. & Denton, E. Asthma In Children and Adults—What Are the Differences and What Can They Tell Us About Asthma? *Front. Pediatr.* **7**, 256 (2019).
31. Zein, J. G. & Erzurum, S. C. Asthma is Different in Women. *Curr. Allergy Asthma Rep.* **15**, 28 (2015).
32. Netuveli, G., Hurwitz, B. & Sheikh, A. Ethnic Variations in Incidence of Asthma Episodes in England & Wales: National Study of 502,482 Patients in Primary Care. *Respir. Res.* **6**, 120 (2005).
33. Busby, J. *et al.* Ethnic Differences in Severe Asthma Clinical Care and Outcomes: An Analysis of United Kingdom Primary and Specialist Care. *J. Allergy Clin. Immunol. Pract.* **10**, 495–505.e2 (2022).
34. Core R Team. R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing* vol. **2** <http://www.r-project.org> (2022).
35. Robinson, D. *fuzzyjoin: Join Tables Together on Inexact Matching*. <https://cran.r-project.org/package=fuzzyjoin> (2020).
36. Bolker, B., Warnes, G. R. & Lumley, T. *gtools: Various R Programming Tools*. <https://cran.r-project.org/package=gtools> (2022).
37. Firke, S. *janitor: Simple Tools for Examining and Cleaning Dirty Data*. <https://cran.r-project.org/package=janitor> (2021).

38. Ooms, J. *The jsonlite Package: A Practical and Consistent Mapping Between JSON Data and R Objects*. Preprint at <https://arxiv.org/abs/1403.2805> (2014).
39. Grolemund, G. & Wickham, H. Dates and Times Made Easy with {lubridate}. *J. Stat. Softw.* **40**, 1–25 (2011).
40. Wickham, H. The Split-Apply-Combine Strategy for Data Analysis. *J. Stat. Softw.* **40**, 1–29 (2011).
41. Walczak, E. *PostcodesioR: API Wrapper Around 'Postcodes.io'*. <https://cran.r-project.org/package=PostcodesioR> (2021).
42. Rinker, T. W. {qdapTools}: Tools to Accompany the qdap Package. *University at Buffalo/SUNY* <http://github.com/trinker/qdapTools> (2015).
43. Blagotić, A. & Daróczy, G. *Rapport: a report templating system*. <https://cran.r-project.org/package=rapport> (2022).
44. Bryan, H. W. and J. readxl: Read Excel Files. <https://cran.r-project.org/package=readxl> (2022).
45. Wickham, H. *et al.* Welcome to the {tidyverse}. *J. Open Source Softw.* **4**, 1686 (2019).
46. Zeileis, A. & Grothendieck, G. zoo: S3 Infrastructure for Regular and Irregular Time Series. *J. Stat. Softw.* **14**, 1–27 (2005).
47. Institute for Government analysis. Timeline of UK Government Coronavirus Lockdowns and Measures, March 2020 to December 2021. <https://www.instituteforgovernment.org.uk/sites/default/files/timeline-coronavirus-lockdown-december-2021.pdf> (2020).
48. UK Government (Department for Transport). Impact of the Coronavirus (COVID-19) Pandemic on Flow Weighting for Congestion Data. <https://www.gov.uk/government/publications/road-congestion-and-travel-times-statistics-guidance/impact-of-the-coronavirus-covid-19-pandemic-on-flow-weighting-for-congestion-data> (2022).
49. Shah, S. A., Quint, J. K. & Sheikh, A. Impact of COVID-19 Pandemic on Asthma Exacerbations: Retrospective Cohort Study of Over 500,000 Patients in a National English Primary Care Database. *Lancet Reg. Heal. - Eur.* **19**, 100428 (2022).
50. Nunn, A. J. & Gregg, I. New Regression Equations for Predicting Peak Expiratory Flow in Adults. *BMJ* **298**, 1070 (1989).
51. Nanchen, D. Resting Heart Rate: What is Normal? *Heart* **104**, 1048–1049 (2018).
52. Pinnock, H. *et al.* Clinical Implications of the Royal College of Physicians Three Questions In Routine Asthma Care: A Real-Life Validation Study. *Prim. Care Respir. J.* **21**, 288–94 (2012).

## Acknowledgements

This work was supported by Asthma + Lung UK as part of the Asthma UK Centre for Applied Research grant number AUK-AC-2018-01. Support from Malmö University was co-funded by the Knowledge Foundation KK-stiftelsen. We thank all the participants of the AAMOS-00 study, this research would not be possible without their time and effort. We thank the AUKCAR PPI for their involvement in developing and analysing the study. We thank the Mobistudy team and Malmö University for their support with data collection. We thank Smart Respiratory Products Ltd for providing the Smart Peak Flow Meter and associated software. We thank FindAir for providing the FindAir ONE devices and FindAir's API. We thank Ambee for providing the pollen data. We thank the Asthma + Lung UK and AUKCAR social media teams, Malcolm Marquette and the Norwich and Norfolk University Hospital for their assistance in participant recruitment. We thank Dr Sarah Brown (Edinburgh Innovations, University of Edinburgh, UK) for organising the contracts required for the study. We thank Aryelly Rodriguez (Edinburgh Clinical Trials Unit, University of Edinburgh, UK) for statistical advice in producing the anonymised AAMOS-00 dataset.

## Author Contributions

K.C.H.T., H.P., A.M.W. and S.A.S. designed the study. S.A.S. is the study guarantor. K.C.H.T. and D.S. set up the data collection system. K.C.H.T. drafted the manuscript which was critically revised by H.P., A.M.W., D.S. and S.A.S. All authors approved the final version of the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41597-023-02241-9>.

**Correspondence** and requests for materials should be addressed to K.C.H.T. or S.A.S.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023