BMJ Open

Prevalence of nocturnal cough in asthma and its potential as a marker for asthma control (MAC) in combination with sleep quality: protocol of a smartphone-based, multicentre, longitudinal observational study with two stages

Peter Tinschert,[1] Frank Rassouli,[2] Filipe Barata,[3] Claudia Steurer-Stey,[4,5] Elgar Fleisch,[1,3] Milo Alan Puhan,[4] Martin Brutsche,[2] Tobias Kowatsch[1]

To cite: Tinschert P, Rassouli F, Barata F, *et al*. Prevalence of nocturnal cough in asthma and its potential as a marker for asthma control (MAC) in combination with sleep quality: protocol of a smartphone-based, multicentre, longitudinal observational study with two stages. *BMJ Open* 2019;**9**:e026323. doi:10.1136/bmjopen-2018-026323

► Prepublication history for this paper is available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2018-026323).

Received 27 August 2018
Revised 5 November 2018
Accepted 16 November 2018

Check for updates

© Author(s) (or their employer(s)) 2019. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to
Peter Tinschert;
peter.tinschert@unisg.ch

ABSTRACT

Introduction Nocturnal cough is a burdensome asthma symptom. However, knowledge about the prevalence of nocturnal cough in asthma is limited. Furthermore, prior research has shown that nocturnal cough and impaired sleep quality are associated with asthma control, but the association between these two symptoms remains unclear. This study further investigates the potential of these symptoms as markers for asthma control and the accuracy of automated, smartphone-based passive monitoring for nocturnal cough detection and sleep quality assessment.

Methods and analysis The study is a multicentre, longitudinal observational study with two stages. Sensor and questionnaire data of 94 individuals with asthma will be recorded for 28 nights by means of a smartphone. On the first and the last study day, a participant's asthma will be clinically assessed, including spirometry and fractionated exhaled nitric oxide levels. Asthma control will be assessed by the Asthma Control Test and sleep quality by means of the Pittsburgh Sleep Quality Index. In addition, nocturnal coughs from smartphone microphone recordings will be labelled and counted by human annotators. Relatively unrestrictive eligibility criteria for study participation are set to support external validity of study results. Analysis of the first stage is concerned with the prevalence and trends of nocturnal cough and the accuracies of smartphone-based automated detection of nocturnal cough and sleep quality. In the second stage, patient-reported asthma control will be predicted in a mixed effects regression model with nocturnal cough frequencies and sleep quality of past nights as the main predictors.

Ethics and dissemination The study was reviewed and approved by the ethics commission responsible for research involving humans in eastern Switzerland (BASEC ID: 2017–01872). All study data will be anonymised on study termination. Results will be published in medical and technical peer-reviewed journals.

Trial registration number NCT03635710; Pre-results.

Strengths and limitations of this study

- ► The planned study has the potential to generate a novel, insightful and externally valid dataset for studying nocturnal symptoms in asthma.
- ► Nocturnal cough and sleep quality may be valid markers for asthma control, which could help to identify windows of opportunities for pharmacological and non-pharmacological interventions.
- ► In principle, smartphones are powerful devices for the automatic monitoring of asthma symptoms, and this study will provide evidence whether symptom detection accuracies actually meet clinical requirements under real-life conditions.
- ► While the study duration of 29 days per participant allows for the investigation of prevalence and trends in nocturnal cough and sleep quality, it is not sufficient to account for long-term symptom patterns and trends (eg, for seasonal effects in allergic asthma).
- ► Due to the heterogeneity of asthma, the potential clinical usefulness of the investigated markers may differ considerably between subgroups of individuals with asthma.

INTRODUCTION

Asthma, a chronic respiratory disease, is one of the most prevalent chronic conditions. According to the WHO, 235 million people suffer from asthma with 383 000 asthma-related deaths in 2015[1]. Common symptoms include wheezing, shortness of breath, chest tightness and coughing.[2] Coughing is perceived by patients as a troublesome symptom,[3] predicts asthma severity[4] and indicates a worse prognosis.[5]

In asthma, coughing and other symptoms tend to get worse at night and often cause awakenings.[2] It has been shown that 'asthma control', the extent to which asthma

BMJ Open: first published as 10.1136/bmjopen-2018-026323 on 7 January 2019. Downloaded from http://bmjopen.bmj.com/ on 20 January 2019 by guest. Protected by copyright.

symptoms are controlled by treatment,[2] is statistically associated with sleep quality.[6] While the association between sleep quality and asthma control is well established, the role of nocturnal cough is less clear.

A first cross-sectional study has indicated that nocturnal cough frequency might be a valid marker for asthma control,[7] rendering it a potentially useful parameter for disease self-monitoring. However, due to the design of the above-mentioned study, questions regarding longitudinal trends of nocturnal cough in asthma remain unanswered. It is also unclear whether nocturnal cough frequency holds any value for asthma control prediction. Moreover, predictions of asthma control on subsequent days based on nocturnal cough frequency could be independent of, moderated by or mediated through sleep quality.

The ongoing study addresses these research gaps by investigating the prevalence of nocturnal cough in asthma over the course of 4 weeks and by exploring the interplay of nocturnal cough frequency and sleep quality for asthma control levels on subsequent days. In this way, the study could indicate whether nocturnal cough and sleep quality are useful parameters for signalling windows of opportunity during which timely pharmacological and non-pharmacological interventions could be administered to prevent asthma deteriorations and attacks.

In addition to these medical objectives, the study aims to investigate the potential of commercial smartphones in the automatic detection of nocturnal cough and assessment of sleep quality. Enabled by the recent progress in the domain of machine learning-based symptom detection,[8 9] self-learning algorithms will be applied on the study's smartphone sensor data to develop models for detecting and assessing the above-mentioned symptoms. The ongoing study explores whether these models are sufficiently accurate for potential medical applications.

METHODS AND ANALYSIS

The study is designed as a multicentre, longitudinal observational study with two stages. The study stages are identical in terms of included participants and data collection procedure. However, they differ in terms of objectives, endpoints and statistical analysis: the first stage is mainly concerned with descriptive analysis of nocturnal cough, while the second stage focuses on the prediction of asthma control based on nocturnal cough frequency and sleep quality.

Aims and objectives

The study has three objectives.

Objectives of the first stage

In the first study stage, the prevalence of nocturnal cough and trends in nocturnal cough frequency will be explored descriptively (1). As a second objective, the study data will be used to develop and evaluate machine learning models for nocturnal cough detection and sleep quality assessment (2).

Objectives of the second stage

In the second stage, the objective is to examine nocturnal cough frequency as a marker for asthma control in combination with sleep quality (3). In other words, the findings of Marsden *et al*[7] will be expanded by collecting longitudinal data, adding sleep quality to the prediction model, measuring additional outcomes such as asthma quality of life and cough status and accounting for environmental factors such as weather conditions, pollution and outside temperature.

Study setting

The study duration per participant will be 29 days with a 3-day run-in phase from the second to fourth study day. Two study centres in Switzerland are participating in the ongoing study: the Clinic for Pulmonology and Sleep Medicine of the Cantonal Hospital St. Gallen and the mediX Group practice Zurich in cooperation with the University of Zurich. On the first and last day of the study, participants will have an in-person appointment at one of the study centres. In the time between appointments, patient-reported outcomes and nocturnal sensor data will be collected by means of a smartphone.

The first participant was included in February 2018. At the time of manuscript submission, 37 participants have completed the study. Data collection is expected to conclude in the first quarter of 2019.

Eligibility criteria

The population investigated in this study are adults with asthma. Inclusion criteria are not restrictive to ensure that findings are generalisable to the target population of adults with asthma: participants must have self-reported physician-diagnosed asthma, be 18 years or older, provide written informed consent and know how to operate a smartphone (self-reported in a yes–no question).

Participants, to whom the following criteria apply, are excluded from participation: participants who are unlikely to successfully complete the daily study tasks on a regular basis (ie, participants with cognitive impairments due to diseases such as severe depression, dementia and Alzheimer's disease), participants for whom obtaining reliable nocturnal measurements is not feasible due to the study's rule-based smartphone recording system (ie, participants with severe insomnia and workers with varying day shift and night shifts) and participants for whom a solely acoustic-based allocation of nocturnal coughs is highly error prone (ie, participants who share the bedroom with a person from the same sex). For a more differentiated rationale behind the latter exclusion criterion, please refer to the discussion.

Outcomes

An overview of all study measures is presented in table 1 in the data collection section.

Endpoints of the first stage

The primary endpoint is coughs per night (c/n), obtained through smartphone audio-recordings, manually

BMJ Open: first published as 10.1136/bmjopen-2018-026323 on 7 January 2019. Downloaded from http://bmjopen.bmj.com/ on 20 January 2019 by guest. Protected by copyright.

Table 1 Overview of measures throughout the 29-day study duration

	Frame of reference	Time of measurement	Obtained by
Patient-reported outcomes			
Medical questionnaires			
ACT	Last 4 weeks	d_1, d_{29}	Physician*
ACT	Last week†	d_8, d_{15}, d_{22}, d_{29}	Smartphone app
PSQI	Last 4 weeks	d_{29}	Study nurse*
PSQI	Last night†	$d_1 - d_{29}$	Smartphone app
AQ20	Momentary	d_1, d_{29}	Smartphone app‡
LCQ	Last 2 weeks	d_1, d_{29}	Smartphone app‡
VAS (severity of nocturnal cough)	Last night	$d_1 - d_{29}$	Smartphone app
VAS (severity of cough by day)	Last day	$d_1 - d_{29}$	Smartphone app
VAS (asthma symptom strength)	This morning	$d_1 - d_{29}$	Smartphone app
Other questionnaires			
Usage habit smartphone app	Momentary	d_8, d_{15}, d_{22}, d_{29}	Smartphone app
Technical evaluation questionnaire	Last 4 weeks	d_{29}	Study nurse
Medical control questions			
Occurrence of cold/rhinitis	Last week	d_8, d_{15}, d_{22}, d_{29}	Smartphone app
Asthma related events (exacerbations and physician visits)	Last week	d_8, d_{15}, d_{22}, d_{29}	Smartphone app
Technical control questions			
Smartphone position overnight	Last days	d_3, d_{14}, d_{24}	Smartphone app
Smartphone distance overnight	Last days	d_3, d_{14}, d_{24}	Smartphone app
Objective measures			
Lung function assessments			
FEV_1		d_1, d_{29}	Spirometry
FVC		d_1, d_{29}	Spirometry
FEV_1/FVC		d_1, d_{29}	Spirometry
FeNO		d_1, d_{29}	NIOX handheld device
Smartphone data			
Microphone		$n_1 - n_{28}$§	Smartphone app
Accelerometer		$n_1 - n_{28}$§	Smartphone app
Proximity sensor		$n_1 - n_{28}$§	Smartphone app
Ambient light sensor		$n_1 - n_{28}$§	Smartphone app
GPS		$n_1 - n_{28}$§	Smartphone app
Bluetooth		$n_1 - n_{28}$§	Smartphone app
Usage data (screen lock status, connection status and battery status)		$n_1 - n_{28}$§	Smartphone app

*In case of non-resolvable appointment scheduling issues on day 29, the participant is instructed to fill out an online version of the questionnaire to ensure parallel measurements between participants.
†The temporal frame of reference for these questionnaires is shortened in comparison with the respective original publication to produce more distinct measurements for each participant, thereby increasing the data resolution for statistical modelling. The question framing and answer options are adapted to the revised frame of reference.
‡If a participant does not respond to the smartphone app questionnaire in the morning of day 29, the questionnaire will be administered by the study nurse during the participant's study centre visit.
§n_1 is the night between day 1 (d_1) and day 2 (d_2) and so forth.
ACT, Asthma Control Test; AQ20, Asthma Questionnaire 20; FeNO, fractionated exhaled nitric oxide; FEV_1, forced expiratory volume in 1 s; FVC, forced vital capacity; GPS, global positioning system; LCQ, Leicester Cough Questionnaire; PSQI, Pittsburgh Sleep Quality Index; VAS, visual analogue scale.

labelled by trained human annotators. Adhering to the recommendations of the European Respiratory Society, cough is defined in this study as 'a three-phase expulsive motor act characterized by an inspiratory effort (inspiratory phase), followed by a forced expiratory effort against a closed glottis (compressive phase) and then

BMJ Open: first published as 10.1136/bmjopen-2018-026323 on 7 January 2019. Downloaded from http://bmjopen.bmj.com/ on 20 January 2019 by guest. Protected by copyright.

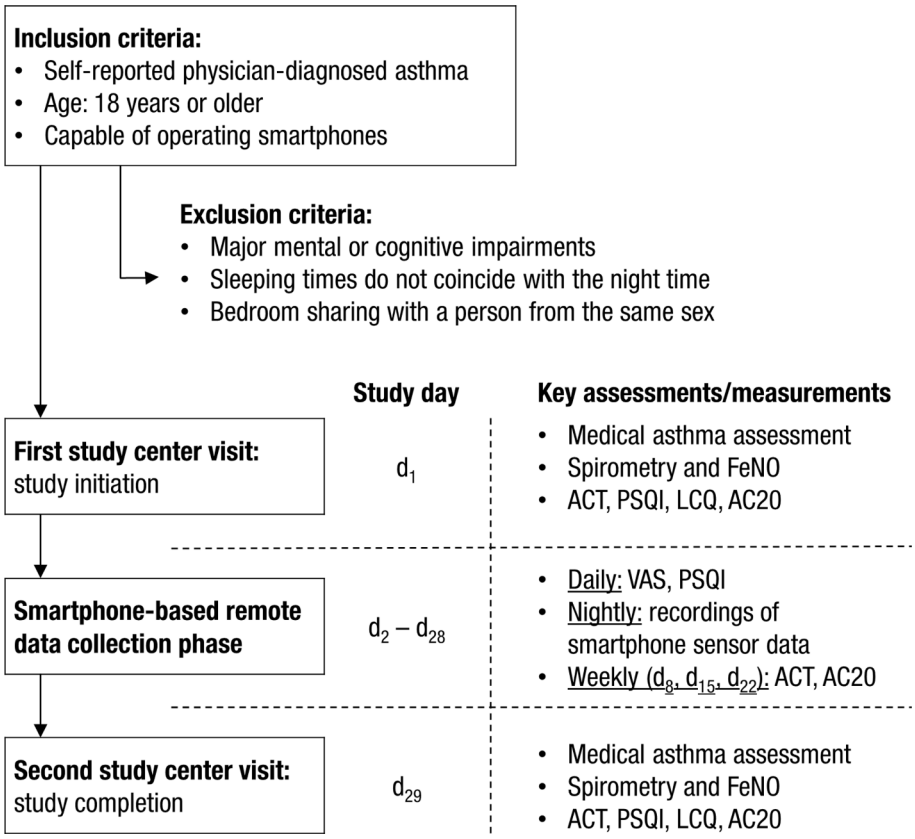

Figure 1 Study flow chart from the participant's perspective. Please note that more detailed information on the assessments and measurements is provided in the measures subsection of the data collection section. AC20, Asthma Questionnaire-20; ACT, Asthma Control Test; FeNO, fractionated exhaled nitric oxide; LCQ, Leicester Cough Questionnaire; PSQI, Pittsburgh Sleep Quality Index; VAS, visual analogue scales).

by opening of the glottis and rapid expiratory airflow (expulsive phase)'.[10]

Secondary endpoints are cough-specific health status (eg, the presence and severity of cough and its effects on quality of life, measured twice by the Leicester Cough Questionnaire[11] and daily by a visual analogue scale for cough severity[12]) and sleep quality as measured by the Pittsburgh Sleep Quality Index (PSQI).[13]

Endpoints of the second stage

In the second stage, the primary endpoint is patient-reported asthma control measured by the Asthma Control Test (ACT).[14]

The secondary endpoint of this stage is asthma-specific quality of life, as measured by the Asthma Questionnaire-20.[15]

Participant timeline

The study flow chart from a participant's perspective is illustrated in figure 1. Participants have to attend two study centre visits at the beginning (d_1) and end (d_{29}) of the study. Furthermore, participants have to adhere fully to the protocol of the smartphone-based remote data collection phase (d_1–d_{28}) on at least 23 days. Otherwise, they will be withdrawn from the study. More specifically, participants are instructed to keep the study smartphone in their bedroom at night and ensure that the smartphone either has a battery level of at least 80% or recharges at night. Participants are also asked to complete the daily questionnaire (preferably in the morning but before 21:00 at the latest).

The length of the daily questionnaire differs between days. On 21 of 29 days, participants have to answer 13 questions (approximately 4 min effort). On three occasions, participants will receive 15 questions (d_3, d_{14}, d_{24}; 5 min effort). At the end of each study week, participants are asked to fill out an extended questionnaire with 43 questions (d_8, d_{15}, d_{22}; 15 min effort). Most questionnaires on the first and last study day will also be administered by means of the smartphone (ie, 52 questions on d_1 and 62 questions on d_{29} with 20 and 25 min effort, respectively). Please refer to the data collection section for a detailed description of the questionnaire schedule.

Sample size

The study is powered with regard to the primary endpoint of the second stage. The required sample size was calculated based on the results of Marsden *et al*.[7] That is, a correlation of r=0.30 between nocturnal cough frequency and asthma control was set as the expected strength of correlation. In order to ensure a power (1-ß) of 80% with a two-tailed type 1-error probability (α) of 5% for rejecting a zero bivariate correlation,[16] 85 participants

BMJ Open: first published as 10.1136/bmjopen-2018-026323 on 7 January 2019. Downloaded from http://bmjopen.bmj.com/ on 20 January 2019 by guest. Protected by copyright.

are required. In terms of parameter accuracy estimation,[17] this sample size would result in a 95% CI for the above-mentioned correlation of [0.09, 0.48].

A dropout rate of 10% is assumed, thus 94 participants will be recruited. To the best knowledge of the authors, no research that investigated the relationship of nocturnal cough, sleep quality and asthma control in a longitudinal setting is available hitherto. This prohibits a more precise evidence-based estimation of the required sample size. Consequently, the required sample size calculation does not take sleep quality and the longitudinal data structure into account. However, due to the comparable strength of association between sleep quality and asthma control (r=0.31–0.36)[6] and nocturnal cough and asthma control (r=0.30),[7] the required sample size holds for this relationship, too.

Recruitment

Personal and impersonal recruitment is employed. Eligible participants from the participating study centres will be referred directly to the study nurse. Multiple channels and methods are used for impersonal recruitment: emails are sent out to mailing lists from Swiss participant organisations for lung diseases and nearby universities, flyers are displayed in local medical offices, pharmacies and universities and online ads are posted to social media websites and digital bulletin boards. Additionally, an email is sent out to clients from a Swiss health insurance provider who have submitted at least one asthma medication receipt for reimbursement in the last 6 months. Please note that no personal information of any participant will be disclosed to the insurance provider. Only aggregated data are made available to the insurance provider (eg, the presumptive number of clients who have responded to mail recruitment).

The recruitment material contains a link to an online screening survey, which consists of the participant information form, a video introducing the study and a short screening questionnaire to test for study eligibility. If potential participants state their interest for study participation and pass the screening survey, their contact details are sent to the study nurses for appointment scheduling.

Participants do not receive monetary incentives for study participation. However, travelling costs are reimbursed up to SFr50 and a lottery will be held on study termination, with 10 prizes worth SFr300 each. Furthermore, a personalised summary of patient-reported study data is made available to each participant after successful study participation.

Data collection

Methods

Health professionals will perform medical assessments in the study centres on the first and last day of the study. For the time between physician appointments, participants will be equipped with a smartphone (Samsung Galaxy A3 2017, SM-A320FL) on which 'Clara', the chat-based study app, is installed. This app is a study-specific adaption of the mobile app[18] for the open source behavioural intervention platform MobileCoach (www.mobile-coach.eu).[19] It records a participant's sensor data in the night (eg, audio data via a smartphone's microphone) and delivers the daily questionnaires to the patients. Figure 2 illustrates the app's user experience. More information on the app is provided in the retention section.

Measures

Table 1 summarises the measures including temporal framing for patient-reported outcomes, time of measurement and by which means they are obtained (eg, medical personnel or by the mobile app). All smartphone measures are collected on the study smartphone handed out to participants on the first study day. Measures can be divided into patient-reported outcomes (eg, medical questionnaires) and objective measures (ie, lung function assessments and nocturnal smartphone sensor data).

Nocturnal smartphone sensor data are recorded by default from 23:00 onwards for a recording length of 9 hours and 40 min to account for the sleeping patterns of the general Swiss population (ie, inferring from a citizen science study,[20] 95% of nocturnal sleep sessions are shorter than this default recording length). Furthermore, participants can manually start (after 21:00) and stop (after 04:00) the sensor data recording, thereby overwriting the default recording settings. If a participant starts the recording manually after the recording has already been started automatically (ie, after 23:00), the default recording length is extended so that it encompasses 9 hours and 40 min from the moment of manual initiation.

In addition to the measures in table 1, a comprehensive medical assessment of a participant's asthma is performed on the first study day. The following medical parameters are determined in this assessment: asthma type (eg, allergic/non allergic asthma, asthma and obesity), asthma history (eg, onset, asthma symptoms and exacerbations in the past), asthma severity and control (by means of Global Initiative for Asthma (GINA) stage[2]), asthma medication that is prescribed and taken and smoking behaviour. The medical assessment on the last study day examines whether changes in the above-mentioned parameters occurred during the course of the study.

Retention

Adherence to the data collection protocol is promoted through multiple strategies. First, participants meet with the study nurse during their first study centre visit to discuss how they can implement the instructions in their daily life and overcome possible barriers. Second, the study app contains a reminder system. If participants do not interact with the study app (and thus fail to comply with their study tasks), they will receive an SMS reminder to their private smartphone. If non-adherence occurs for two consecutive days, the participant receives a reminder phone call from the study nurse. Third, the study app is specifically designed to promote participant retention.

BMJ Open: first published as 10.1136/bmjopen-2018-026323 on 7 January 2019. Downloaded from http://bmjopen.bmj.com/ on 20 January 2019 by guest. Protected by copyright.

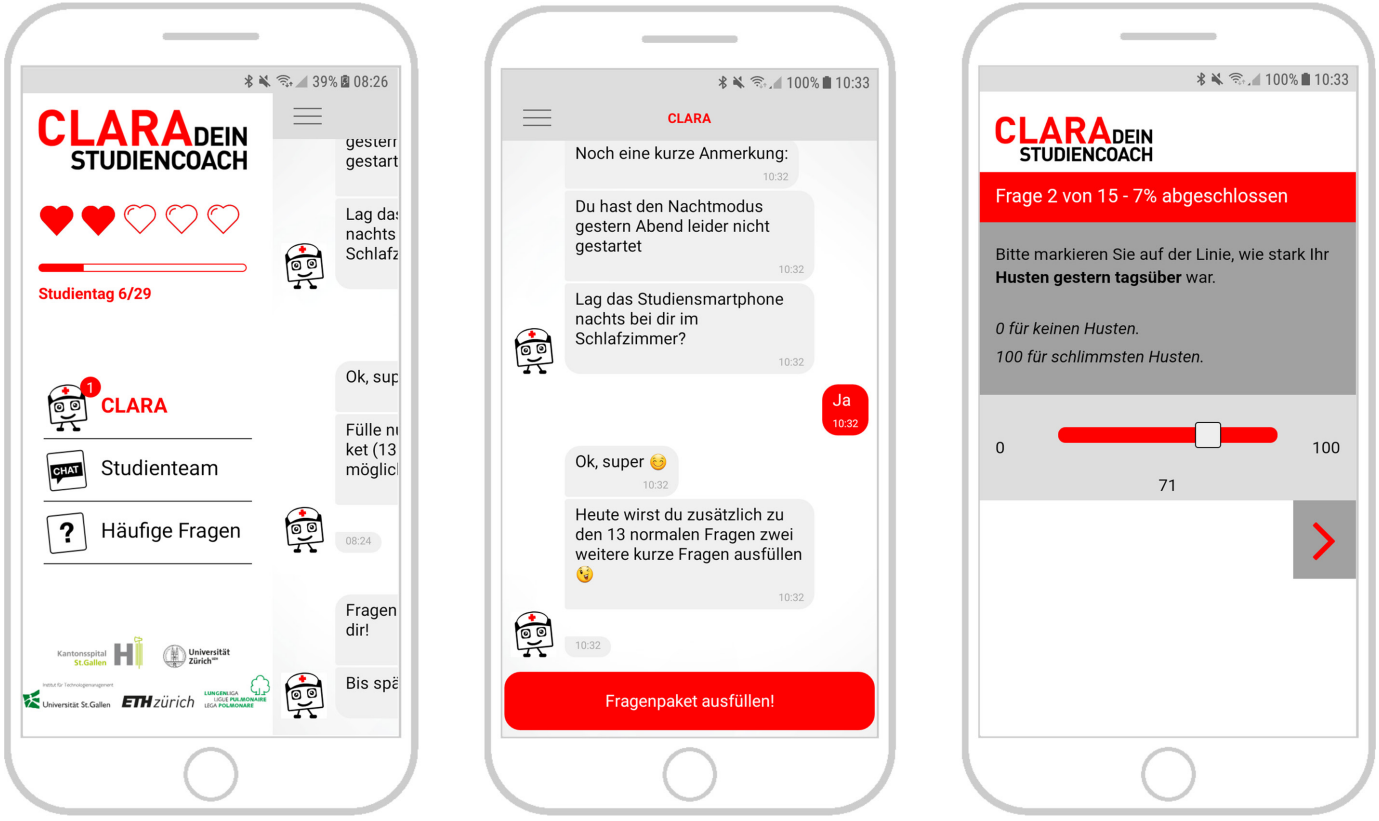

Figure 2 Screenshots of the 'Clara' study app. Sidebar for navigation (left), chat-based interface with predefined answer options (middle) and questionnaire module for patient-reported outcomes (right).

Participants are guided through the remote data collection phase of the study by a text-based healthcare chatbot.[21] More specifically, all study information, instructions and tasks are conveyed in a daily conversation with 'Clara', a virtual study nurse designed to establish a working alliance with the participant.[22] This interactive approach is known to increase adherence in clinical settings.[23] Furthermore, several nudging techniques[24] are incorporated in the app (figure 2 shows how the study makes use of loss aversion in participants to increase engagement, ie, by symbolising non-adherence as lost hearts).

Data management

All participant data gathered by the physician (eg, results of the asthma assessment) and study nurse (eg, lung function assessments) in the study centres will be transferred into an electronic format and stored online on the study server. Traceability of changes and version control will be managed by the version control software TortoiseSVN (https://tortoisesvn.net/).

Questionnaire data collected by the mobile app will be instantly saved on the study server. Nocturnal sensor data will be stored locally on the smartphone and backed up to external hard drives and secure online storage once a participant has completed the study. However, excerpts of nocturnal sensor data are uploaded continuously to the study server for quality monitoring purposes. ETH Zurich provides all study servers and online storage services used in the study.

To match different data sources, a unique user hash is randomly generated for data obtained by smartphone application for each participant, and then linked to the participant ID used in the study centres' case report forms. Please note that a participant's personal information is kept exclusively at the study centres in a separate document (ie, it is not included in the case report forms). Once the data analysis is complete, the document containing a participant's personal information is destroyed to anonymise the study data.

Analysis

Data preparation

In order to obtain c/n, which is the primary endpoint of the first stage and a fundamental part of the analysis in the second stage, the raw audio data have to be annotated for cough events. Due to the study setting, human annotators will label cough exclusively based on audio data. Every annotator will receive labelling training before being granted access to the audio data.

Annotators will be asked to label the explosive phase of a cough and, if multiple coughs occur in succession, the corresponding cough epoch[10] (see figure 3 for an illustrative example for the labelling). By storing the meta-information of the cough labels (ie, time at which the cough occurred in a given night and cough duration), the time spend coughing per night can be calculated (also referred to as 'cough seconds'). In this way, the study's labelling process covers three out of the four ways to quantitatively

BMJ Open: first published as 10.1136/bmjopen-2018-026323 on 7 January 2019. Downloaded from http://bmjopen.bmj.com/ on 20 January 2019 by guest. Protected by copyright.

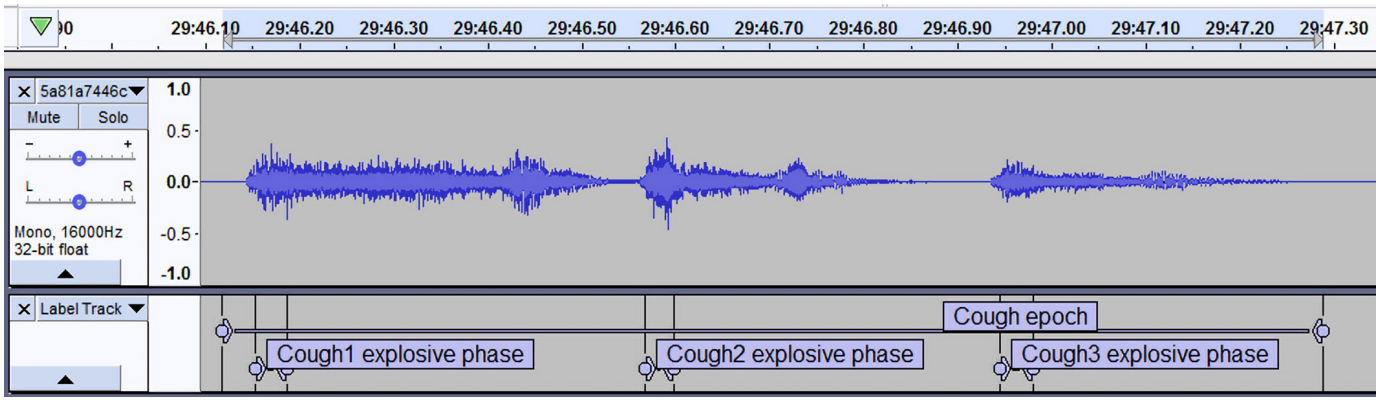

Figure 3 Labelling example for cough explosive phases and epochs. Screenshot taken in Audacity 2.2.2. (https://www.audacityteam.org).

describe coughing that are proposed by the European Respiratory Society.[10]

Furthermore, annotators will be asked to label the sex of the person for each cough. If uncertainties occur during the labelling process (eg, unclear sex of a cougher or whether a sound was a cough), annotators will mark the corresponding section for review. A subset of the recorded nights will be labelled by all annotators, so that inter-rater agreement can be estimated.

For subsets of audio data where no human speech is included, coughs may be labelled by means of a crowdsourcing approach (eg, Amazon Mechanical Turk; https://www.mturk.com/).

Planned statistical analysis in the first stage

In the first study stage, c/n will be analysed descriptively and visually. Additionally, autoregressive and moving average models will be used to model trends in c/n over multiple nights.

Additionally, machine-learning models will be developed to: (1) automatically detect nocturnal cough and (2) assess sleep quality based on the smartphone sensor data. For (1), manually annotated coughs constitute the ground truth and parameters obtained from the smartphone audio-recordings serve as algorithm features. For (2), PSQI scores are the ground truth and all available smartphone sensor data may be used as algorithm features. Both algorithms will be designed using supervised machine learning techniques and evaluated in accordance with standard practice of algorithm development.[25]

Planned statistical analysis in the second stage

In the second stage, a mixed effects model will be estimated to analyse the diagnostic efficacy of nocturnal cough and sleep quality as markers for asthma control and asthma-related quality of life. The mixed effects model is selected to account for the nested data structure (ie, c/n is nested in participants). Asthma control, as measured by the ACT, will be the predicted outcome. In a second model, asthma-related quality of life as measured by the AC20 will be used as the outcome variable instead. Predictor variables in both models will be c/n, PSQI scores and the interaction of c/n and PSQI scores. The models will take multiple control variables into account (eg, type of asthma, medication intake, concomitant diseases and smoking behaviour). By modelling the data with a random intercept and random slope, the heterogeneous manifestation of asthma symptoms in participants will be represented in the models (eg, it seems plausible that c/n might have a higher predictive power for asthma control and/or quality of life for some participants; compare cough-variant asthma[26]). In order to account for repeated measurements, the statistical models will rely on an autoregressive covariance structure.[27]

In a further exploratory analysis, a modified version of the asthma control mixed effects prediction model will estimated in which the outputs of the two machine learning models will serve as the predictors instead of the corresponding ground truths (ie, manually labelled coughs and PSQI scores as described in the last paragraph). The prediction performance of the original and modified model version will be compared so that the utility of a machine learning based symptom detection approach can be evaluated.

Please note that since the present study is rather explorative in its nature, deviations from the planned statistical analysis may be warranted based on the structure and distributions of the collected data.

Patient and public involvement

The development of the research objectives was informed by two focus groups with adults with asthma. The 15 participants were recruited by means of purposive sampling, ensuring that patients with uncontrolled, partially controlled and well-controlled asthma from a wide age range (21–66 years) participated in the discussions. Patients voiced the need for a monitoring method with objective and unobtrusive symptom measurements. Furthermore, passive nocturnal symptom monitoring by means of a smartphone was regarded a promising approach to satisfy this need.

BMJ Open: first published as 10.1136/bmjopen-2018-026323 on 7 January 2019. Downloaded from http://bmjopen.bmj.com/ on 20 January 2019 by guest. Protected by copyright.

Patients were not directly involved in the design of the study. However, insights from the focus groups have informed the design of the study app.

Study participants will receive a personalised summary of the collected patient-reported study data after successful study completion and can opt-in to receive the study results once they have been published.

ETHICS AND DISSEMINATION

Consent

Participants will provide informed written consent at their first visit of a study centre. Participants can withdraw their consent at any time during the study.

Access to data and confidentiality

Only the study team (ie, the study nurses, involved physicians and technical support staff) will have access to the study data before anonymisation. Smartphone sensor data will be stored separately from private participant information obtained in the case report form (CRFs). After the study is completed, all participant data will be matched, and the physical files containing the private information will be destroyed. Furthermore, all audio segments containing a participant's speech will be erased with the objective to anonymise all study data. For anonymised audio data, the cough labelling process may be scaled by means of a crowdsourcing approach.

Dissemination policy

Results of the study are planned to be published in medical and technical peer-reviewed journals. Consistent with the Open Research Data initiative of the Swiss National Science Foundation,[28] the study team may publish the anonymised study data in a scientific database if participants provide informed consent and all the applicable legal provisions are met. CSS insurance, the external funder of the study, has no active involvement in the study as stated in the funding contracts, for example, in terms of study design, data analysis or the decision to publish results.

DISCUSSION

The study follows three main objectives: (1) to ascertain the prevalence and trends of nocturnal cough in asthma; (2) to develop and evaluate smartphone-based machine learning models for automated nocturnal cough detection and sleep quality assessment; and (3) to investigate the potential of nocturnal cough frequency as a marker for asthma control in combination with sleep quality.

The study is characterised by the volume, variety and ecological validity of the collected data. Studies in the domain of nocturnal respiratory symptoms are often restricted to cross-sectional data,[6 7] or only consider limited data sources.[29] They often rely exclusively on survey data[30–32] or collect data in artificial settings, such as using obtrusive devices for cough detection.[33–35] External validity is further supported by the modest eligibility criteria for study participation: from a medical perspective, participants are only excluded if they suffer from a condition that renders successful study participation unlikely. Therefore, study results should be generalisable to the general population, which is often not the case in asthma-related trials.[36]

However, external and internal validity constitute conflicting priorities in study design.[37] The study's emphasis on external validity results in a threat to internal validity: in order to enable in situ measurements with a high temporal resolution, the study relies considerably on patient-reported outcomes and a purely sound-based cough annotation approach.

To mitigate constraints of patient-reported outcomes, participants will have a run-in phase of 3 days at the beginning of the study, including a training session on the first day, when they fill out daily recurring questions under supervision of the study nurse.

To enable accurate annotation of coughs, a participant's acoustic fingerprint will be collected on the first study day by recording intentional coughs and standardised voice samples. Furthermore, participants will indicate every morning whether they have slept alone or shared the bedroom with another person or a pet. By providing additional information to human cough annotators, accuracy of the labelling process is expected to increase. Moreover, an internal unpublished analysis conducted with data from an earlier study[38] has indicated that humans are able to assign coughs correctly to a person's sex only by means of sound. In this analysis, six annotators, who have not received labelling training or information about the participants, assigned 53 lab-recorded coughs to the correct sex with an average Cohen's kappa of 0.92 (interpretable as an almost perfect agreement[39]) after being allowed to listen to each cough only once. This finding is in line with prior research on sex differences in cough[40 41] and suggests that human annotators are in principle able to exploit sex differences in cough sounds to accurately match coughs to participants. Therefore, the study's exclusion of participants who share their bedroom with a person from the same sex is a further measure to enable a reasonably accurate labelling process.

Furthermore, participants regularly specify the smartphone's position and average distance to the bed at night. This information serves as a quality control measure for the smartphone microphone recordings. By being able to reproduce the conditions of seemingly flawed recordings (eg, recordings with almost no noise), informed decisions can be made post hoc as to whether a participant should be withdrawn from the study on grounds of problematic recording conditions. However, no such conditions could be identified in systematic field tests prior to the study.

Limitations

The study constitutes an extensive invasion of a participant's privacy and implicitly requires participants to be at least partially open to the use of consumer technology

BMJ Open: first published as 10.1136/bmjopen-2018-026323 on 7 January 2019. Downloaded from http://bmjopen.bmj.com/ on 20 January 2019 by guest. Protected by copyright.

for health related matters. Thus, a self-selection bias[42] of participants can be expected. However, it does not seem plausible to assume that attitudes towards study participation are associated with any aspect of a participant's asthma in a significant way. Therefore, the self-selection bias appears to be a negligible limitation for the validity of the study. However, the inclusion criterion of self-reported ability to use a smartphone might introduce a selection bias affecting the study's generalisability, especially due to poorer technology adoption in the elderly.[43]

A more severe limitation is due to the fact that symptom severity in asthma varies over time and is often driven by seasonal changes.[2] With a duration of 29 days, the proposed study does not allow for the investigation of long-term within-person symptom trends and patterns. The limited study duration also prohibits the prediction of relatively rare, but clinically more meaningful outcomes such as asthma attacks,[36] which occur on average approximately once every 2 years per patient.[44] For example, a study with the objective to predict the risk of an asthma attack in the next 30 days considered a period of 18 months[45] in order to obtain sufficient observations for statistical prediction.

Another important limitation of the study stems from the conceptualisation of asthma as a disease. As a recent expert commission has argued, asthma is not a homogenous, clearly defined disease but instead an umbrella term for chronic lung disorders with distinct aetiologies, endotypes and phenotypes.[36] In this study, self-reported physician-diagnosed asthma was defined as the main inclusion criterion. Due to this broad and unspecific inclusion criterion, the study sample will likely include heterogeneous types of asthma. It is conceivable that the predictive power of nocturnal cough and sleep quality for asthma control may vary significantly between different types of asthma and therefore between participants. This issue can be addressed statistically by modelling the predictive power of nocturnal cough and sleep quality for asthma control as random effects. However, a statistical solution to a conceptual issue is not a satisfying remedy for clinical practice: if the predictive power varies considerably between individual participants, health professionals would not be able to make an informed decision at the point of care whether disease monitoring based on nocturnal cough and sleep quality would yield a clinical benefit for any given patient. To address this limitation, subsequent studies could focus on specific subgroups[46] to deliver clinically useful predictions: if subgroups can be identified for which a robust and consistent prediction applies, nocturnal cough and sleep quality could serve as valid digital biomarkers in certain subpopulations despite the heterogeneity of asthma. However, if the objective of subsequent studies is not to enable applicability but to ensure generalisability of results, larger sample sizes will be required to encompass asthma's heterogeneity. From a technological point of view, the generalisability of the machine learning models is constrained due to the fact that only one smartphone model is used in this study. It is unclear to which extent the accuracy of automated nocturnal cough detection and sleep quality assessment would vary between smartphones models which, for example, differ in terms of built-in sensors,[47] recording quality and number of microphones.[48] Additionally, other important factors, such as battery usage[49] and technological dependence on the smartphone's operating system and manufacturer,[50 51] need to be addressed before the smartphone-based symptom detection models could be implemented in clinical practice or in patient self-monitoring.

CONCLUSION

The proposed longitudinal study investigates the prevalence of nocturnal cough in asthma and its potential to predict asthma control individually and in combination with sleep quality. Although the study design prioritises external over internal validity, a selection bias due to poor technology adoption in seniors and the limited sample size might undermine generalisability. By collecting the data with standard smartphones, the study aims to lay the groundwork for scalable digital biomarkers for asthma.

Author affiliations
[1]Institute of Technology Management, University of St. Gallen, St. Gallen, Switzerland
[2]Lung Center, Cantonal Hospital St. Gallen, St. Gallen, Switzerland
[3]Department of Management, Technology and Economics, ETH Zurich, Zurich, Switzerland
[4]Institute of Epidemiology, Biostatistics and Prevention, University of Zurich, Zurich, Switzerland
[5]mediX Group Practice, Zurich, Switzerland

Acknowledgements The authors would like to thank Grace Xiao for proofreading the manuscript.

Contributors Study design: all authors. Study execution in the study centres: FR and CS-S. Technological support of the study execution: PT and FB. Development of study app: FB, PT and TK. Writing and editing of manuscript: PT. Critical review and revision of manuscript: MAP, TK, FB, CS-S, FR, MB and EF. Guarantor: FR. All authors reviewed and approved the manuscript before submission.

Funding This study is funded by CSS Insurance, Switzerland. The CSS insurance supported the recruitment of participants but had no role in study design, app design, data management plans, or in reviewing and approving the manuscript for publication.

Competing interests PT, FB, EF and TK are affiliated with the Center for Digital Health Interventions (www.c4dhi.org), a joint initiative of the Department of Management, Technology and Economics at ETH Zurich and the Institute of Technology Management at the University of St. Gallen, which is funded in part by the Swiss health insurer CSS. EF and TK are also cofounders of Pathmate Technologies, a university spin-off company that creates and delivers digital clinical pathways and has used the open source MobileCoach platform for that purpose, too. However, Pathmate Technologies is not involved in the study app described in this paper.

Patient consent for publication Not required.

Ethics approval The study protocol was reviewed and approved by the ethic commission responsible for research involving humans in eastern Switzerland (ie, Ethikkommission Ostschweiz; Business Administration System for Ethics Committees ID: 2017–01872).

Provenance and peer review Not commissioned; externally peer reviewed.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is

BMJ Open: first published as 10.1136/bmjopen-2018-026323 on 7 January 2019. Downloaded from http://bmjopen.bmj.com/ on 20 January 2019 by guest. Protected by copyright.

properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

REFERENCES

1. World Health Organization. Asthma 2017. http://www.who.int/en/news-room/fact-sheets/detail/asthma (Accessed 5 Jul 2018).
2. Global Initiative for Asthma. 2018. 2018 GINA Report, Global strategy for asthma management and prevention.
3. Osman LM, McKenzie L, Cairns J, *et al*. Patient weighting of importance of asthma symptoms. *Thorax* 2001;56:138–42.
4. de Marco R, Marcon A, Jarvis D, *et al*. Prognostic factors of asthma severity: a 9-year international prospective cohort study. *J Allergy Clin Immunol* 2006;117:1249–56.
5. Thomson NC, Chaudhuri R, Messow CM, *et al*. Chronic cough and sputum production are associated with worse clinical outcomes in stable asthma. *Respir Med* 2013;107:1501–8.
6. Luyster FS, Teodorescu M, Bleecker E, *et al*. Sleep quality and asthma control and quality of life in non-severe and severe asthma. *Sleep Breath* 2012;16:1129–37.
7. Marsden PA, Satia I, Ibrahim B, *et al*. Objective Cough Frequency, Airway Inflammation, and Disease Control in Asthma. *Chest* 2016;149:1460–6.
8. Accurate and privacy preserving cough sensing using a low-cost microphone. *Proceedings of the 13th international conference on Ubiquitous computing*: ACM, 2011.
9. Toss'n'turn: smartphone as sleep and sleep quality detector. *Proceedings of the sigchi conference on human factors in computing systems*: ACM, 2014.
10. Morice AH, Fontana GA, Belvisi MG, *et al*. European Respiratory Society (ERS). ERS guidelines on the assessment of cough. *Eur Respir J* 2007;29:1256–76.
11. Birring SS, Prudon B, Carr AJ, *et al*. Development of a symptom specific health status measure for patients with chronic cough: Leicester Cough Questionnaire (LCQ). *Thorax* 2003;58:339–43.
12. Spinou A, Birring SS. An update on measurement and monitoring of cough: what are the important study endpoints? *J Thorac Dis* 2014;6:S728–34.
13. Buysse DJ, Reynolds CF, Monk TH, *et al*. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res* 1989;28:193–213.
14. Nathan RA, Sorkness CA, Kosinski M, *et al*. Development of the asthma control test: a survey for assessing asthma control. *J Allergy Clin Immunol* 2004;113:59–65.
15. Barley EA, Quirk FH, Jones PW. Asthma health status measurement in clinical practice: validity of a new short and simple instrument. *Respir Med* 1998;92:1207–14.
16. Hulley SB, Cummings SR, Browner WS, *et al*. Designing clinical research: Lippincott Williams & Wilkins. 2013.
17. Bonett DG. Sample size requirements for estimating intraclass correlations with desired precision. *Stat Med* 2002;21:1331–5.
18. I. nternational Conference on Design Science Research in Information Systems*Design and evaluation of a mobile chat app for the open source behavioral health intervention platform mobilecoach*: Springer, 2017.
19. Filler A, Kowatsch T, Haug S, *et al*. *MobileCoach: A Novel Open Source Platform for the Design of Evidence-based, Scalable and Low-Cost Behavioral Health Interventions - Overview and Preliminary Evaluation in the Public Health Context. 14th annual Wireless Telecommunications Symposium (WTS)*. USA, New York: IEEE, 2015.
20. Walch OJ, Cochran A, Forger DB. A global quantification of "normal" sleep schedules using smartphone data. *Sci Adv* 2016;2:e1501705.
21. Kowatsch T, Nißen M, Rüegger D, *et al*. The Impact of interpersonal closeness cues in text-based healthcare chatbots on attachment bond and the desire to continue interacting: an experimental design. 2018.
22. Bickmore T, Gruber A, Picard R. Establishing the computer-patient working alliance in automated health behavior change interventions. *Patient Educ Couns* 2005;59:21–30.
23. Bickmore TW, Puskar K, Schlenk EA, *et al*. Maintaining reality: Relational agents for antipsychotic medication adherence. *Interact Comput* 2010;22:276–88.
24. Richard H, Thaler CRS. *Nudge: Improving decisions about health, wealth, and happiness*: Springer, 2008.
25. Hastie T, Tibshirani R, Friedman J. *The elements of statistical learning data mining, inference, and prediction*. 2ndCorr. 7th printing 2013 edn. New York: Springer, 2009.
26. Desai D, Brightling C. Cough due to asthma, cough-variant asthma and non-asthmatic eosinophilic bronchitis. *Otolaryngol Clin North Am* 2010;43:123–30.
27. Field A, Miles J, Field Z. *Discovering Statistics Using R*. 1st edn. London: Sage, 2012.
28. Swiss National Science Foundation. Data Management Plan (DMP) - Guidelines for researchers. http://www.snf.ch/en/theSNSF/research-policies/open_research_data/Pages/data-management-plan-dmp-guidelines-for-researchers.aspx
29. Janson C, Gislason T, Boman G, *et al*. Sleep disturbances in patients with asthma. *Respir Med* 1990;84:37–42.
30. Campos FL, de Bruin PFC, Pinto TF, *et al*. Depressive symptoms, quality of sleep, and disease control in women with asthma. *Sleep Breath* 2017;21:361–7.
31. Fitzpatrick MF, Martin K, Fossey E, *et al*. Snoring, asthma and sleep disturbance in Britain: a community-based survey. *Eur Respir J* 1993;6:531–5.
32. Garden M, O'Callaghan M, Suresh S, *et al*. Asthma and sleep disturbance in adolescents and young adults: A cohort study. *J Paediatr Child Health* 2016;52:1019–25.
33. Assessment of audio features for automatic cough detection. *Signal Processing Conference, 2011 19th European*: IEEE, 2011.
34. McGuinness K, Holt K, Dockry R, *et al*. P159 Validation of the VitaloJAK™ 24 Hour Ambulatory Cough Monitor. *Thorax* 2012;67:A131–31.
35. Stores G, Ellis AJ, Wiggs L, *et al*. Sleep and psychological disturbance in nocturnal asthma. *Arch Dis Child* 1998;78:413–9.
36. Pavord ID, Beasley R, Agusti A, *et al*. After asthma: redefining airways diseases. *Lancet* 2018;391:350–400.
37. Shadish WR, Cook TD, Campbell DT. *Experimental and quasi-experimental designs for generalized causal inference*. Houghton: Mifflin and Company, 2002.
38. Barata F, Kowatsch T, Tinschert P, *et al*. *Personal MobileCoach: tailoring behavioral interventions to the needs of individual participants*. Heidelberg, Germany: Proceedings of the 2016 ACM UbiCompACM, 2016:1089–94.
39. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med* 2005;37:360–3.
40. Lee KK, Matos S, Ward K, *et al*. Sound: a non-invasive measure of cough intensity. *BMJ Open Respir Res* 2017;4:e000178.
41. Gender differences in voluntary cough sound spectra demonstrated by an inverse power law analysis. *Engineering in medicine and biology, 2002 24th annual conference and the annual fall meeting of the biomedical engineering society EMBS/BMES Conference, 2002*: Proceedings of the Second JointIEEE, 2002.
42. Berger V. Selection bias and covariate imbalances in randomized clinical trials: John Wiley & Sons. 2007.
43. Hanson VL. Influencing technology adoption by older adults. *Interact Comput* 2010;22:502–9.
44. Suruki RY, Daugherty JB, Boudiaf N, *et al*. The frequency of asthma exacerbations and healthcare utilization in patients with asthma from the UK and USA. *BMC Pulm Med* 2017;17:74.
45. Frey U, Brodbeck T, Majumdar A, *et al*. Risk of severe asthma episodes predicted from fluctuation analysis of airway function. *Nature* 2005;438:667–.
46. Rothwell PM. Treating individuals 2. Subgroup analysis in randomised controlled trials: importance, indications, and interpretation. *Lancet* 2005;365:176–86.
47. NDSS. AccelPrint: Imperfections of Accelerometers Make Smartphones Trackable. 2014.
48. Kardous CA, Shaw PB. Evaluation of smartphone sound measurement applications. *J Acoust Soc Am* 2014;135:EL186–EL192.
49. Understanding Human-Smartphone Concerns. *A study of battery life*. Berlin, Heidelberg.: Springer Berlin Heidelberg, 2011.
50. Kenney M, Pon B. Structuring the smartphone industry: is the mobile internet OS platform the key? *Journal of Industry, Competition and Trade* 2011;11:239–61.
51. Felt AP, Egelman S, Wagner D. *I've got 99 problems, but vibration ain't one: a survey of smartphone users' concerns. Proceedings of the second ACM workshop on Security and privacy in smartphones and mobile devices*: Raleigh, North Carolina, USA: ACM, 2012:33-–44.

BMJ Open: first published as 10.1136/bmjopen-2018-026323 on 7 January 2019. Downloaded from http://bmjopen.bmj.com/ on 20 January 2019 by guest. Protected by copyright.