PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Predicting asthma attacks using connected mobile devices and machine learning; the AAMOS-00 observational study protocol
AUTHORS	Tsang, Kevin; Pinnock, Hilary; Wilson, Andrew; Salvi, Dario; Shah, Syed Ahmar

VERSION 1 – REVIEW

REVIEWER	Aslanidis, Theodoros St.Paul General Hospital of Thessaloniki, Intensive Care Unit
REVIEW RETURNED	20-May-2022

GENERAL COMMENTS	Well structured, clearly explained manuscript, easy to follow in order to understand every concept and phase of the study protocol. The included images provide an usefull visualisation of the elements of the study. I feel that the paper could be publihsed in its present form.

REVIEWER	Salman, Omar H Al Iraqia University, Network Department
REVIEW RETURNED	05-Jun-2022

GENERAL COMMENTS	The topic is interesting. The structure of the article is well presented. However, I would like to ask the authors to do: 1) There is a large amount of research on this topic. So, the authors need to update their references with the relevant articles to stand on the most significant variables/ attributes that shuold be cosidered in thier approach. 2) The authors claimed that they will evaluate the performance of more than 4 machine learning algorithms. However, when the authors do analyses of the relevant study, they would find recommendations from other researchers about the optimal methods. This will reduce their simulation and testing time. 3) The expectation and the outcome of the machine learning method need to be described as initial plans. in the other words, what does the result of machine learning mean to the authors? this question needs to be answered as a flowchart/ figure. 4) The paper has no results, discussion, and statistical outcome as it is a protocol study.

REVIEWER	Gonem, Sherif Nottingham University Hospitals NHS Trust
REVIEW RETURNED	20-Jul-2022

GENERAL COMMENTS	This study protocol aims to develop and assess the feasibility of an asthma attack predictor system, with data inputs from a number of smart devices.

	Major comments: 1) As a feasibility study the sample size of 30 participants is adequate, but it is unlikely that training and validation of a machine learning algorithm could be adequately carried out with this number. The authors estimate that they will obtain 3470 daily records and that 12 exacerbation events may occur during the study. Their argument for accepting these low numbers is that they will use the iForest algorithm, an anomaly detection system that in theory does not require any events at all. However, this argument is not convincing, as with only 12 events they will have no way of reliably testing the validity of any machine learning algorithm they develop. To put this in context, our recently published study using machine learning applied to daily peak flow monitoring and symptom scores included 728,535 daily records in 2010 patients, with 576 severe exacerbation events (DOI: 10.1080/02770903.2020.1802746). Therefore I would recommend that the feasibility outcomes are treated as the primary outcomes of the study, and that any machine learning / algorithm development is purely exploratory. 2) The authors correctly point out that the attrition rate for the use of asthma apps in general is quite high. In the introduction they distinguish between active and passive monitoring, with the implication being that patients are more likely to adhere in the long-term with passive monitoring. However, the proposed system includes a number of elements that require daily active engagement from the participants, including daily symptom questionnaires and peak flow measurements. Therefore the attrition rate using this system is likely to be similarly high, even given the 1 month run-in period. The authors cite work from their own group (reference 24), showing that only 17% of people continued to use an asthma self-management app after 1 month; yet they have chosen to use a different study that was not specific to asthma (reference 22) on which to base their predicted attrition rate - please provide justification for this. In general, what is the justification for believing that the attrition rate for this system will be any better than previously tested systems? Even if the investigators succeed in recruiting a highly motivated cohort to participate in this study, how likely is it that these findings will be generalisable to a broader population of people with asthma? Minor comments: 3) It would be worth mentioning the large EU-funded myAirCoach project in the introduction, as this had similar aims to the present project. What is the proposed study going to add to this previous work? 4) Line 278: Please clarify why "the longitudinal data collected in this study are likely to be more complete than the Asthma Mobile Health Study". 5) Lines 286-9: The second sentence does not follow from the first. Also, as previously mentioned, there is unlikely to be enough data to meaningfully split into training and test sets. 6) Please expand on future Patient and Public Involvement plans. 7) Please expand on the data sharing policy.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Well structured, clearly explained manuscript, easy to follow in order to understand every concept and phase of the study protocol. The included images provide a useful visualisation of the elements of the study. I feel that the paper could be published in its present form.
RESPONSE: Thank you for your comments and for taking the time to review our paper.

Reviewer: 2
The topic is interesting. The structure of the article is well presented. However, I would like to ask the authors to do:
1) There is a large amount of research on this topic. So, the authors need to update their references with the relevant articles to stand on the most significant variables/ attributes that should be considered in their approach.
RESPONSE: Thank you for your suggestion. We have added some additional relevant articles to support the selection of monitoring devices in our study in lines 79-81 and reproduced below:
"There has been an increasing number of mHealth studies to predict asthma attacks and develop passive monitoring to support asthma self-management,[11] including the use of smart peak flow meters,[12] night-time activity tracking,[13,14] smart inhalers,[15] smartphone administrated questionnaires,[6,16,17] and weather data.[18,19]"

2) The authors claimed that they will evaluate the performance of more than 4 machine learning algorithms. However, when the authors do analyses of the relevant study, they would find recommendations from other researchers about the optimal methods. This will reduce their simulation and testing time.
RESPONSE: We agree that there have been previous studies about asthma attack prediction using home monitoring mHealth data, however, there is no consensus on the optimal algorithm. See DOI: 10.2147/JAA.S285742. Thus, we have chosen to take a broad approach to test the performance of different classes of classification algorithms. We have revised our manuscript in lines 316 – 318. The text now reads:
"There is no consensus on the optimal algorithm for classification as previous studies are not comparable.[11] Therefore, we have taken a broad approach to use five state-of-the-art algorithm classes including Bayesian networks, decision trees, iForest, logistic regression, and support vector machines.[44,45]"

3) The expectation and the outcome of the machine learning method need to be described as initial plans. in the other words, what does the result of machine learning mean to the authors? this question needs to be answered as a flowchart/ figure.
RESPONSE: In our previous analysis, our asthma attack classifiers built on self-reported questionnaire data achieved AUC > 0.87 when differentiating between stable and unstable periods (DOI: 10.1109/EMBC44109.2020.9175679). We anticipate that the model performance will improve with the addition of objective data. This study will also provide an opportunity to test the prediction performance using different data subsets, such as self-reported data only, self-reported and all objective data (passive and active monitoring data), and passive monitoring only. We have now added a new figure (Figure 8), and revised the manuscript (lines 312 – 315) accordingly:
"Using different subsets of the data and features, we will test the performance of predictions made using different modes of monitoring, such as self-reported data alone, self-reported and objective data (active and passive monitoring), and passive monitoring data only (see Figure 8). Our previous analysis using only self-reported data achieved AUC > 0.87 and we expect the performance to increase with the addition of objective data.[6]"

4) The paper has no results, discussion, and statistical outcome as it is a protocol study.
RESPONSE: This is correct. Thank you for taking the time to review our paper.

Reviewer: 3
This study protocol aims to develop and assess the feasibility of an asthma attack predictor system, with data inputs from a number of smart devices.

Major comments:
1) As a feasibility study the sample size of 30 participants is adequate, but it is unlikely that training and validation of a machine learning algorithm could be adequately carried out with this number. The authors estimate that they will obtain 3470 daily records and that 12 exacerbation events may occur during the study. Their argument for accepting these low numbers is that they will use the iForest algorithm, an anomaly detection system that in theory does not require any events at all. However, this argument is not convincing, as with only 12 events they will have no way of reliably testing the validity of any machine learning algorithm they develop. To put this in context, our recently published study using machine learning applied to daily peak flow monitoring and symptom scores included 728,535 daily records in 2010 patients, with 576 severe exacerbation events (DOI: 10.1080/02770903.2020.1802746). Therefore I would recommend that the feasibility outcomes are treated as the primary outcomes of the study, and that any machine learning / algorithm development is purely exploratory.
RESPONSE: Thank you for your comment. We have revised the manuscript in line with the recommendation and we now treat the feasibility outcomes as primary and the machine learning asthma attack prediction as exploratory.
Lines 104 – 107 have been revised and now reads:
“The overarching aim of this study is to develop and assess the feasibility and acceptability of an asthma self-management system using existing smart devices, collect novel monitoring data and leverage machine learning to explore the feasibility of an asthma attack prediction algorithm based on passive monitoring.”
Lines 164 – 187 have been revised and now reads:
“Primary endpoints
The primary endpoints of this study are adherence to monitoring, which are defined by the collection of data using different devices. For each task, we will measure the percentage of total days completed.
Secondary endpoints
The secondary endpoints of this study are asthma attacks. Primarily, we use severe asthma attacks as defined in the American Thoracic Society (ATS)/European Respiratory Society (ERS) Task Force 2009 statement.[34] The definition is as follows:
• Severe asthma attacks are events that require urgent action on the part of the patient and physician to prevent a serious outcome. Such attacks are defined as a deterioration requiring use of systemic corticosteroids (or an increase from a stable maintenance dose).[34]
Severe asthma attacks will be identified using the use of systemic corticosteroids question on the weekly self-reported questionnaire. Courses of corticosteroids separated by 1 week or more will be treated as separate severe attacks.
We will also explore the use of moderate asthma attacks, as defined in the ATS/ERS Task Force 2009 statement.[34] The definition is as follows:
…”
The subtitle in line 296 has been revised and now reads:
“Exploring asthma attack prediction (using monitoring data)”
Lines 324 – 326 have been revised and now reads:
“Furthermore, if there are sufficient data, retrospective analysis will simulate the effects of limiting the use of active monitoring data, to simulate patients only taking active measurements (e.g., peak flow

and symptom diary) when prompted."

2) The authors correctly point out that the attrition rate for the use of asthma apps in general is quite high. In the introduction they distinguish between active and passive monitoring, with the implication being that patients are more likely to adhere in the long-term with passive monitoring. However, the proposed system includes a number of elements that require daily active engagement from the participants, including daily symptom questionnaires and peak flow measurements. Therefore the attrition rate using this system is likely to be similarly high, even given the 1 month run-in period. The authors cite work from their own group (reference 24), showing that only 17% of people continued to use an asthma self-management app after 1 month; yet they have chosen to use a different study that was not specific to asthma (reference 22) on which to base their predicted attrition rate - please provide justification for this. In general, what is the justification for believing that the attrition rate for this system will be any better than previously tested systems? Even if the investigators succeed in recruiting a highly motivated cohort to participate in this study, how likely is it that these findings will be generalisable to a broader population of people with asthma?
RESPONSE: We have revised our estimate of the attrition rate in phase 2 after considering multiple attrition rates from different study designs in lines 133-135. In particular, financial incentives and nature of the daily tasks will influence the attrition rate. Our study has a £5 financial incentive per month and a further incentive to complete phase 1, which is that participants of phase 2 can keep the three smart devices beyond study completion.
The text now reads:
"Based on 30 participants with daily monitoring, 3,098 recordings (= 30 people × 30 days × (85% retention at baseline, 50% retention at 6 months) × 85% adherence) for each of the daily measurements can be expected in phase 2."
Justification of these assumptions is as follows:
• 85% baseline retention
o Hui: After first 30 with large drop in retention (would be observed in phase 1 of the AAMOS-00 study), day 30 to 60 had 61% retention.
o Pathiravasan: Baseline 89% retention
o Xu: For the 28-day study, there was 100% retention.
o Overall, the average of the three studies is 83%, we round up to 85% for our estimates as there are more incentives to continue the AAMOS-00 study than Hui's and Pathiravasan's studies.
• 50% retention at 6 months
o Hui: On day 0 there were 111 users, 28 users on day 30, 17 users on day 60, and 7 users on day 90. Accounting for a 30-day run-in period, this equates to a 25% retention after 2 months. Extrapolating the attrition rate to 6 months gives 2% (25%3) retention after 6 months.
o Pathiravasan: 52% retention after 6 months
o Senturia: 94% retention to the study after 6 months. After 9 months, there was an 89% retention to peak flow monitoring.
o Compernolle: Month 1 had 78% (213/274), month 3 had 72% (198/274) retention. Therefore, after 1-month run-in period, the retention from month 1 to month 3 is 93%. Extrapolated to 86% retention after 6 months.
o Overall, the average of the four studies is 59%, rounded down to 50% for a conservative estimate.
• 85% adherence
o Xu: For every day of the 28-day study, there was between 90% and 100% adherence without any clear drop off. The overall adherence was 97%.
o Senturia: Month 9 had 42% adherence. Interpolated to be 81% after 3 months and 61% after 6 months, assuming 100% on day one and linear decline.
o Overall, the average adherence of the two studies is 89%, rounded down to 85% as Xu's study was much shorter than our AAMOS-00 study.

Referenced studies:
Study Task frequency Study duration Task nature Financial incentive
Hui [DOI:10.14236/jhi.v25i4.1056] Daily 3 months Questionnaire No
Pathiravasan [DOI:10.2196/24773] 3 months 12 months Questionnaire No
Xu [DOI:10.1371/journal.pone.0199838] Daily 1 month Questionnaire and smartwatch Yes
Senturia [DOI:10.1016/S0197-2456(98)00032-4] Twice daily 3 2-week periods across 9 months Questionnaire and peak flow Yes
Compernolle [DOI:10.2196/jmir.3402] Continuous 3 months Smartwatch No

The aim of this study is to collect a novel dataset and explore the feasibility of monitoring and the possibility of asthma attack prediction on this data. We recognise that the analysis will need to simulate the adherence levels of the wider population when considering generalisability. We have expanded on this limitation in lines 342-347 which now reads:.
“This study is limited to patients with severe asthma at risk of acute attacks, and the findings may not generalise to the wider asthma population. Furthermore, only participants using a pressurised metered dose relief inhaler that fits our smart device are able to join the study, which is around 80% of the UK’s asthma population.[46] If there are sufficient data, we will explore the generalisability of the results through simulating different adherence to monitoring.”

Minor comments:
3) It would be worth mentioning the large EU-funded myAirCoach project in the introduction, as this had similar aims to the present project. What is the proposed study going to add to this previous work?
RESPONSE: The myAirCoach study is indeed similar to our project. However, the frequency of monitoring in the myAirCoach study is daily monitoring for one month followed by weekly monitoring with an additional two weeks of daily monitoring, whereas our study will provide six months of continuous daily monitoring. Furthermore, our recruitment will allow patients across the whole UK to participate and is not limited to practices in London and Manchester, who are more likely to be based in urban areas. To our knowledge, the myAirCoach dataset is not yet publicly accessible and has not been used to test machine learning-based algorithms for asthma attack prediction.
We have added reference to the myAirCoach project in lines 90-94, which now reads:
“A related study is the myAirCoach study,[22] which investigated asthma home-monitoring using connected devices. However, their participants conducted daily monitoring for the first month with an additional randomly allocated two weeks, compared to seven months in total proposed in this study. To our knowledge, the dataset from myAirCoach is not publicly available and it has not yet been used to test any machine learning-based algorithms for asthma attack prediction.[23,24]”

4) Line 278: Please clarify why "the longitudinal data collected in this study are likely to be more complete than the Asthma Mobile Health Study".
RESPONSE: We have clarified that the AMHS had a less specific participant selection criteria, compared to our study population who are more likely to be adherent to monitoring and thus provide more complete data. The text (page 11) now reads:
“The methods of linear fit and bin-algorithms will be used to collate and produce summary variables over irregular time-series and to handle missing data.[6] After processing the data, machine learning classifiers will be trained to predict asthma attacks. Evaluation of these classifiers will allow comparison with the benchmarks set using daily questionnaires alone.[6] However, due to the selection of participants with higher adherence to monitoring in phase 2 of this AAMOS-00 study, the longitudinal data collected in this study are likely to be more complete than the data collected by the wide range of participants in the Asthma Mobile Health Study.[43]”

5) Lines 286-9: The second sentence does not follow from the first. Also, as previously mentioned,

there is unlikely to be enough data to meaningfully split into training and test sets.
RESPONSE: We have revised the sentence which now reads:
“Furthermore, if there are sufficient data, retrospective analysis will simulate the effects of limiting the use of active monitoring data, to simulate patients only taking active measurements (e.g., peak flow and symptom diary) when prompted.”

6) Please expand on future Patient and Public Involvement plans.
RESPONSE: We have added more detail about our planned work with PPI members. The additional text on page 12 now reads:
“The findings about feasibility and acceptability will be interpreted with input from PPI members. Furthermore, we plan to continue working with PPI members to develop a system that supports asthma self-management.”

7) Please expand on the data sharing policy.
RESPONSE: We have added more information about the Edinburgh DataShare service and a reference to the website in lines 372-373. The text now reads:
“At the end of the study, the anonymised research data will be stored at Edinburgh DataShare (a digital repository of research data produced at the University of Edinburgh) in perpetuity. Researchers will be able to access and download the data from the website for their own research.”
Thank you for taking the time to review our paper.

VERSION 2 – REVIEW

REVIEWER	Gonem, Sherif Nottingham University Hospitals NHS Trust
REVIEW RETURNED	05-Sep-2022

GENERAL COMMENTS	Thank you, my previous comments have been satisfactorily addressed.