Validation of the Hyfe cough monitoring system: a multicenter clinical study

Carlos Chaccour, Isabel Sánchez-Olivieri, Sarah Siegel, Gina Megson, Kevin L. Winthrop, Juan Berto-Botella, Juan P. de-Torres, Lola Jover, Joe Brew, Mindaugas Galvosas, Matthew Rudd, Peter Small

Submitted to: Journal of Medical Internet Research
on: March 18, 2024

Table of Contents

Validation of the Hyfe cough monitoring system: a multicenter clinical study

Carlos Chaccour[1, 2, 3] MD, PhD; Isabel Sánchez-Olivieri[2] MD; Sarah Siegel[4] MPH, PhD; Gina Megson[4] MPH; Kevin L. Winthrop[4] MD, MPH; Juan Berto-Botella[2] MD, PhD; Juan P. de-Torres[2] MD, PhD; Lola Jover[5] MSc; Joe Brew[5] PhD; Mindaugas Galvosas[5] MD; Matthew Rudd[5, 6] PhD; Peter Small[5, 7] MD, PhD

[1]ISGlobal, Barcelona Institute for Global health Barcelona ES
[2]Clinica Universidad de Navarra Pamplona ES
[3]Centro de Investigación Biomédica en Red de Enfermedades Infecciosas Madrid ES
[4]Oregon Health Science University Portland US
[5]Hyfe Inc. Wilmington US
[6]University of the South Sewanee US
[7]University of Washington Seattle US

Corresponding Author:
Carlos Chaccour MD, PhD
ISGlobal, Barcelona Institute for Global health
Rosello 132
5a 2o
Barcelona
ES

Abstract

Background: The ability to passively and continuously monitor coughing would significantly improve cough management and research. To date there is no automated clinically validated cough monitor that can be routinely used in clinical care and research. Here we describe the validation of such an automated cough monitor.

Objective: To assess the overall performance of the Hyfe Cough Monitoring System when used by individuals with problematic cough, under common living conditions

Methods: This multicenter observational study compared the results of the Hyfe CoughMonitor wrist-worn device with manually counted coughs in subjects with a variety of etiologies as they went about their usual daily activities. We collected 24 hours of continuous sounds from subjects while they simultaneously wore a CoughMonitor and an audio recorder. Coughs were labelled by multiple trained annotators who listened to the continuous audio recordings using validated methodology. The time stamps of these human-detected coughs were compared to those of the CoughMonitor to determine the system's overall performance using event-to-event and hourly rate correlation analyses.

Results: Over the 546 hours monitored, 4454 cough events were recorded; The overall sensitivity was 90.4% (95% CI of 88.3% to 92.2%). The overall false positive rate was 1.03 false positives per hour (95% CI of 0.84 to 1.24). The overall correlation between manual and CoughMonitor measured hourly coughing was high (Pearson correlation coefficient of 0.99 with OLS slope 0.94 and OLS intercept 0.68).

Conclusions: The present analysis of cough events demonstrated that the Hyfe CoughMonitor accurately reflects them with a high sensitivity and a low false positive rate. Future studies should confirm its potential role in the management of patients with cough in clinical practice. Clinical Trial: Clinicaltrials.gov: NCT05723159

(JMIR Preprints 18/03/2024:58545)
DOI: https://doi.org/10.2196/preprints.58545

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?
✓ **Please make my preprint PDF available to anyone at any time (recommended).**

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✔ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Validation of the Hyfe cough monitoring system: a multicenter clinical study

Carlos Chaccour[1,2,3]* carlos.chaccour@isglobal.org

Isabel Sánchez-Olivieri[2] isa.sanchez5995@gmail.com

Sarah Siegel[4] siegels@ohsu.edu

Gina Megson[4] megson@ohsu.edu

Kevin L. Winthrop[4] winthrop@ohsu.edu

Juan Berto Botella[2] jberto@unav.es

Juan P. de-Torres[2,5] jpdetorres@unav.es

Lola Jover[6] lola.j@hyfe.ai

Joe Brew[6] joe@hyfe.ai

Mindaugas Galvosas[6] mindaugas.g@hyfe.ai

Matthew Rudd[6,7] matthew.r@hyfe.ai

Peter Small[5,8] peter@hyfe.ai

* Corresponding author

[1]ISGlobal, Barcelona Institute for Global Health, Barcelona, Spain

[2]Clínica Universidad de Navarra, Pamplona, Spain

[3]Centro de Investigación Biomédica en Red de Enfermedades Infecciosas, Madrid, Spain

[4] Oregon Health Science University, Portland, OR, USA

[5]IDisNa, Pamplona, Spain

[6]Hyfe, Wilmington, DE, USA

[7]University of the South, Sewanee, TN, USA

[8]University of Washington, WA, USA

Running title: Validation of the Hyfe CoughMonitor

Abstract

Background: The ability to passively and continuously monitor coughing would significantly improve cough management and research. To date there is no automated clinically validated cough monitor that can be routinely used in clinical care and research. Here we describe the validation of such an automated cough monitor.

Methods: This multicenter observational study compared the results of the Hyfe CoughMonitor wrist-worn device with manually counted coughs in subjects with a variety of etiologies as they went about their usual daily activities. We collected 24 hours of continuous sounds from subjects while

they simultaneously wore a CoughMonitor and an audio recorder. Coughs were labelled by multiple trained annotators who listened to the continuous audio recordings using validated methodology. The time stamps of these human-detected coughs were compared to those of the CoughMonitor to determine the system's overall performance using event-to-event and hourly rate correlation analyses.

Results: Over the 546 hours monitored, 4454 cough events were recorded; The overall sensitivity was 90.4% (95% CI of 88.3% to 92.2%). The overall false positive rate was 1.03 false positives per hour (95% CI of 0.84 to 1.24). The overall correlation between manual and CoughMonitor measured hourly coughing was high (Pearson correlation coefficient of 0.99 with OLS slope 0.94 and OLS intercept 0.68).

Conclusion: The present analysis of cough events demonstrated that the Hyfe CoughMonitor accurately reflects them with a high sensitivity and a low false positive rate. Future studies should confirm its potential role in the management of patients with cough in clinical practice.

Introduction

Cough is one of the most common reasons for which patients seek medical care, it is associated with a broad range of medical conditions and greatly contributes to healthcare expenditure all over the world (1, 2). For some diseases, such as COPD and COVID-19, cough rates also help to predict adverse outcomes (3, 4). However, in an era when symptom quantification drives refinement in diagnosis and precision in therapy (5), cough is currently not measured as part of clinical practice. This is not because of lack of interest, as efforts to quantify cough date back to the 1950's (6, 7).

Improvements in acoustic signal processing and machine learning techniques have fostered renewed attention to contactless fully automated cough monitoring (8-12). Objective, prolonged cough monitoring can provide a valuable data stream for the diagnosis, prognosis, assessment of treatment response and even syndromic surveillance of respiratory diseases as well as for the development of novel therapeutics. However, there are currently no validated, unobtrusive, and fully-automated

cough monitoring systems that are commonly used to monitor cough as patients go about their normal activities (13).

Validation of such systems relies upon having accurate ground truth cough data as gold standard. There is consensus that manual cough counting from audio recordings, albeit laborious, can achieve good interobserver agreement (13). When manual labeling cough acoustic data, the cough second (a second containing at least one explosive phase of a cough) is a unit of annotation that correlates well with true cough rates and reduces the ambiguity associated with the use of individual explosive phases or cough epochs (14, 15).

The primary objective of the study was to assess the overall performance of the Hyfe CoughMonitor (Hyfe Inc., 2022), when used by individuals with problematic cough, in comparison to the gold standard of manual cough annotation. Additionally, we also compared the CoughMonitor's performance during the daytime versus nighttime, between individuals and as a function of cough rates.

Methods

The Hyfe CoughMonitor

The CoughMonitor App runs on an Android Smartwatch (Shenszhen Domino Times Smart 4G Watch, Model DM20). In brief, this app uses the watch's microphone to continuously capture and encrypt ambient sounds in a manner that cannot be replayed and is deleted after processing, thus ensuring that sound recordings are secure and transient. Firstly, a "peak detection" algorithm detects and records a 0.5 second snippet when an explosive "cough-like" sound is detected. Secondly, this snippet is classified as a cough or non-cough by a second "cough recognition" algorithm. All processing happens in-device. While charging, timestamps and cough durations are transmitted via Wi-Fi to the Hyfe cloud, where the time of each cough is converted to cough seconds. In this study, all uploading was handled by study personnel at the end of the monitoring period of each participant. The same version of the cough detection software was used throughout this study (version 1.0.0).

Continuous sounds were recorded using the same Android smartwatch (Shenszhen Domino Times Smart 4G Watch, Model DM20), but running custom software that continuously recorded all ambient sounds.

Study design

This is a multicenter observational study of individuals with problematic cough due to a variety of cough related conditions designed to assess the overall performance of the Hyfe CoughMonitor System in comparison to manually counted cough events.

Enrollment and eligibility

Individuals of both sexes who had problematic coughs consulting at two clinical sites between March 17 and Nov 7, 2023: 1) Oregon Health & Science University (OHSU) in the US, and 2) University Clinic of Navarra in Spain A third group of participants was enrolled remotely in a decentralized manner in the US through targeted outreach. Inclusion and exclusion criteria are shown in **Table 1**.

Sample size

Hourly cough counts are necessarily non-negative integers and follow negative binomial distributions closely (16), neither these counts nor any simple transformations thereof are normally distributed, precluding application of standard formulas for the SEs of the Pearson correlation or linear regression coefficients. A simulation based on data collected in previous studies to estimate the sampling distribution of hourly cough rates and showed that a minimum of 18 participants contributing 20 hours of monitoring each, yielding 360 paired person-hours for analysis, would result in a standard error under 0.1 in average correlations and slopes (see online data supplement for the study protocol). This target was later expanded to 23 participants in order to have over 50% of the sample recruited in the US for regulatory purposes.

Data Acquisition

After obtaining informed consent, research subjects were instructed to wear two devices: (i) the Hyfe CoughMonitor on one wrist and (ii) a second identical watch running custom software that

continuously recorded all ambient sounds on the other wrist.

At bedtime, subjects were instructed to charge both devices at bedside. The exact start and stop time of monitoring was recorded by study personnel. Subjects were instructed to monitor for 24 hours and write down the times they went to bed, awoke and any times they had to leave the watch aside such as while showering. Participants who recorded for less than 20 hours were excluded from analysis.

Cough Annotation

To obtain the gold standard, the exact start time of each cough was manually annotated on the continuous recordings as previously described (3, 15, 17). In brief, continuous audio recordings were broken into 60 second segments and presented to two trained annotators using a proprietary audio playback, visualization and data entry program. Each 60 second segment was listened to by the annotators independently and blinded to one another´s results. Every cough, throat clear, sneeze or cough-like sounds was labeled as a segment, from its beginning to its end, noting if the sound was distant or unclear. Sounds for which the two labelers disagreed on the presence of a cough, the timing of a cough's start by greater than 100 milliseconds, or the indication that the cough sound was distant or unclear were adjudicated by a third expert trained annotator. The adjudicator was aware of the discrepant annotations and listened to each 60s audio segment containing discrepancies in full. The interobserver variability of this system has been quantified as negligible with an inter-labeler Pearson´s correlation of 0.96 (15).

Event Classification

The timestamps of automatically (CoughMonitor) and manually (human annotated) detected cough events were described and compared statistically. Each timestamp was converted to cough seconds, defined as a second during which at least one individual cough occurs (14). As previously shown, cough seconds can be interchangeable units with explosive phases as units of cough (14) and a more consistent annotation metric (15). We used the following event definitions:

- True Positive: a cough detected by the CoughMonitor matching human annotators within a 0.5

second margin.

- False Positive: a "cough" detected by the CoughMonitor but either (a) not labeled by humans or (b) labeled by humans as a different sound i.e., throat clear, sneeze or (c) not matching coughs labeled by humans within 0.5 seconds.
- False Negative: a cough labeled by humans that is not detected by the CoughMonitor within a 0.5 second margin.
- True Negative: a non-cough sound that is not detected by the CoughMonitor. Note however that given the continuous nature of sound, true negatives are not practically feasible to measure and thus we use false positive rate per hour to convey specificity.

Statistical Analysis

Results were analyzed on the basis of cough seconds in two complementary ways: (1) an event-to-event comparison of each cough second yielding the CoughMonitor's sensitivity, false positive rate and positive predictive value, and (2) a correlation analysis of hourly cough rates comparing human annotators and the monitor presented as a Bland-Altman plot and, in the online data supplement, as linear plots.

All statistical analyses were performed in R (R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.)

Cough Event Based Analysis

To calculate the CoughMonitor`s sensitivity, false positive rate and positive predictive value as measures of its success at detecting individual cough seconds we compared (i) ground truth annotations (converted to cough seconds) with (ii) cough seconds as per the CoughMonitor. When a cough second was detected by the CoughMonitor within 0.5 seconds of a human labeled cough, the event was considered to be matched and therefore a "true positive". Following a match, both the CoughMonitor cough second and human labeled cough second are removed from eligibility for

further matching (i.e., if a human-labeled cough fell within 0.5 seconds of the CoughMonitor, the human-labeled cough could only match with one CoughMonitor cough, meaning that only one event is considered a true positive; the remaining, unmatched CoughMonitor event would be considered a false positive despite it being within 0.5 seconds of a human labeled cough. A “false positive” was considered to be a CoughMonitor detection which was not matched to a human labeled cough. By comparison with ground truth annotations, each CoughMonitor timestamp is therefore either a true positive or a false positive. Sensitivity and the false positive rate were then calculated according to the usual formulas,

$$\text{CoughMonitor Sensitivity} = \frac{\textit{Total number of true positives}}{\textit{Total number of cough seconds}},$$

where the denominator is determined by ground truth, and

$$\text{False positive rate} = \frac{\textit{Total number of false positives}}{\textit{Total number of hours of monitoring}}.$$

Cough Rate Based Analysis

To evaluate the CoughMonitor’s overall performance, its hourly cough second counts were compared, person-hour by person-hour, with the gold standard hourly cough second counts obtained from trained human annotators. The agreement between these paired person-hour counts and the ideal model $y = x + \epsilon$, where ϵ denotes an integer-valued random error term with mean 0 and constant variance, was assessed with both a linear analysis and a Bland-Altman analysis.

To account for within-subject correlations, the clustered bootstrap was used to calculate all of these confidence intervals, with each cluster consisting of one individual’s paired person-hour counts.

Secondary Analyses

We then conducted three pre-specified secondary analyses comparing the CoughMonitor’s performance (1) between day and nighttime, using a two-sample t-test (2) between each individual, using descriptive metrics and (3) as a function of cough rate by means of one-way analyses of the

variance.

There was no missing data in this study.

Ethics Approval and Registration

The study was approved by the OHSU Institutional Review Board (number 24749), the Ethics committee for medical Research in Navarra (Comité de Ética de la Investigación con medicamentos de Navarra) (number PI_2022/101), and by WCG IRB for the distributed trial (number 20232553). The study was prospectively registered in Clinicaltrials.gov (NCT05723159)

Results

Enrollment

A total of 28 participants with problematic coughs due to a variety of etiologies were recruited. A battery misconfiguration resulted in the device turning off in the first 3 subjects who were enrolled and this misconfiguration was corrected after the 3rd participant, the detection algorithm was not changed. No subsequent technical interruptions were observed. Two participants inadvertently prematurely terminated monitoring prior to 20 hours and were thus excluded from analysis (**Figure 1**). There were no adverse events related to the use of CoughMonitor.

The final analysis included data from 23 subjects (**Table 2**). Almost two thirds were female (15, 65%), the mean age was 52 (range 24 to 72 years), and 12 (52%) were recruited in the US (5 from Oregon Health Science University and 7 from the dispersed clinical trial). The mean monitoring time was 23.8 hours (range 22.9 to 25.64) and the mean hourly cough rate annotated by humans was 8.1 (range 1.5-32). All diagnoses are provided in **Table 2** (see **Table E1** in the online data supplement for all participant-related data).

Continuous Recordings

A total of 546 hours of continuous audio/monitoring time containing 4,454 cough seconds was captured from all participants. The mean number of cough seconds by participant was 200 (range 36 to 821). The mean hourly cough rate was 8.15 (range 1.5 to 32).

It was apparent to annotators listening to the continuous recordings that most monitoring occurred in subjects' homes and workplaces. However, monitoring also occurred in a wide variety of acoustic environments such as grocery stores, restaurants, in vehicles, and even while listening to loud techno music.

Cough Event Based Performance Results

The overall sensitivity was 90.4% (95% CI of 88.3% to 92.2%). The overall false positive rate was 1.03 false positives per hour (95% CI of 0.84 to 1.24). **Figure 2** shows the cumulative number of cough seconds detected by the CoughMonitor and the human annotators. The sex and age of the participants had no impact on performance (**Figure E1** and **E2** in the online data supplement).

The sensitivity rate for individual subjects ranged from 78.1% to 96.5% with false positivity rates ranging from 0.38 to 2.23 cough seconds per hour (**Figure 3, Figure 4**).

Cough Rate Based Performance Results

CoughMonitor performance based on hourly cough rate is shown in a standard Bland-Altman plot in **Figure 5**, providing a visual assessment of the agreement between the hourly cough rates as measured by the CoughMonitor and by the human annotators. We calculated the upper and lower 95% limits of agreement (LOAs), and also computed the 95% confidence interval for the upper and lower COAs using cluster bootstrap resampling techniques. As shown by the dashed horizontal lines, the overall bias (mean difference) is 0.23 (95% CI -0.039 to 0.51), the lower limit of agreement (LOA) is -3.7 (95% CI -5.2 to -3), and the upper limit of agreement is 4.8 (95% CI 4 to 6).

Linear Cough Rate Performance Results

A linear validation analysis assumes that the paired person-hour counts exhibit a linear relationship. **Figure 6** below confirms this assumption visually, and the overall Pearson correlation coefficient of 0.99 (95% CI 0.962 to 0.996) quantifies the strength of this linear relationship. While having a Pearson correlation coefficient close to +1 is a necessary criterion for satisfactory performance, this is insufficient alone, as it only implies that the paired counts cluster tightly about some line; the

coefficients of the regression line (the dashed black line in **Figure 6**) measure how close this line is to the ideal line y=x (the solid black line in **Figure 6**). The slope and the intercept of the OLS line of best fit in **Figure 6** are 0.94 (95% CI 0.91 to 0.97) and 0.74 (95% CI 0.50 to 0.99), respectively, indicating good agreement and hence satisfactory performance.

Because the results reported by the Hyfe CoughMonitor are expressed as hourly coughing rates, we calculated the correlation based on all 477 of the hours monitored. As shown in **Figure 3**, the overall linear correlation was high, with a Pearson correlation coefficient of 0.99, an OLS slope of 0.94, and an OLS intercept of 0.74. This same information is presented in a Bland-Altman plot in **Figure 4** and **Figure E3** in the online supplement.

Secondary analyses

Daytime versus nighttime performance

There were no statistically significant differences in the performance of CoughMonitor in between day and night periods. (**Table E2** and **Figure E4** in the online data supplement).

Performance by individuals

The mean difference between the true cough second rate and that estimated by CoughMonitor was -0.25 cough seconds/hour (IQR -0.6) i.e. CoughMonitor overestimated rates in about one quarter of cough per hour. All individual ground truth and CoughMonitor-calculated cough seconds are presented in the **Table E1** in the online data supplement.

Performance as a function of cough rate

The daily coughing rates for the 23 participants ranged from 1.5 cough seconds per hour to 32 cough seconds per hour. Participants with cough rates below 4 cough seconds per hour were categorized as low (n = 7), those with cough rates between 4 and 8 (n = 8) as middle and those with rates above 8 as “high” (n = 8). (**Table E3** in the online data supplement). The scatterplots of paired hourly cough second counts and the boxplots of sensitivity and false positive rates in **Figure E5** in the online data supplement show that the CoughMonitor performs similarly across these groups. One-way analyses

of variance show no differences in sensitivity (p = 0.63) or in false positive rates (p = 0.76) between these three groups.

Discussion

We evaluated the Hyfe CoughMonitor in over 546 hours of continuous audio containing 4,454 coughs from 23 participants with a broad range of cough-causing diagnoses. This resulted in a sensitivity of 90.4% and an overall false positive rate was 1.03 false positives per hour. This translates into a margin of error of 2.2%. Thus, the monitor is unique in its ability to continuously monitor cough frequency in a manner that is unobtrusive, fully automated, and privacy preserving as users go about their usual activities of daily living. These attributes promise to improve patients' understanding of their cough and its triggers, as well as to improve the diagnosis and management of disease. Furthermore, continuous cough monitoring addresses some of the fundamental statistical inadequacies of short-term cough monitoring that have limited clinical trials of new antitussive drugs as recently raised by the US Food and Drug Administration (18, 19).

Using cough seconds as unit of analysis and applying a previously described annotation protocol, has proven accurate to establish the ground truth, showing an intra-labeler disagreement of less than two per hour monitored; (Pearson’s correlation 0.98) and inter-labeler agreement with a Pearson’s correlation of 0.96 (15). In addition to showing high reproducibility, this analysis establishes the error rate of even the most rigorous human annotation with implications for their use in trials and as comparators when validating automated monitors.

Sensitivity is an essential metric of event-level performance, but detecting or failing to detect a single cough is not clinically meaningful. Hourly cough rates, on the other hand, are highly informative to patients, providers, and investigators, the Bland-Altman plots reported here show the robustness of hourly rates determined using CoughMonitor.

There are several limitations to this study. First, the results were obtained with a single make and model of Android watch which could limit its generalizability to other devices. However, in the

recent literature resulting from a sharp increase in the general interest in cough recognizing algorithms, comparable results have been obtained using a variety of devices that employ comparable microphones and chip sets (20, 21). Second, in this study we have made only limited efforts to distinguish between coughs from the wearer and others in close proximity. Thus, the device should not be used in environments with a high burden of non-user coughs until such an analysis has been conducted. Third, cough recognition will be influenced by the acoustic environment in which the CoughMonitor is worn, as a large volume of peaks may artificially inflate the detected cough second rate, even with a low false positivity rate. Although subjects were instructed to avoid unusually loud environments, they were actually worn in a wide variety of settings such as while using public transport and listening to loud techno music. Nonetheless, users must be informed that results may be less accurate in settings with extremely loud background noise.

Conclusion

The Hyfe CoughMonitor System has an overall sensitivity for detecting a cough event above 90% and a false positivity rate of about one per hour. These results were observed in men and women with a variety of diagnoses as individuals went about their usual activities of daily living. Given the system's high accuracy, usability, and scalability, it has the potential to greatly improve clinical care, drug development and regulatory efforts.

Acknowledgements and statements

Acknowledgements

The authors would like to thank the participants and study team for their efforts to make this research happen.

Author contributions

Conceptualisation: CCh, JB, PS

Data curation: CCh, ISO, SS, LJ

Formal analysis: MR, JB

Investigation: CCh, ISO, SS, GM, KW, JPDT, JB, PS

Methodology: CCh, MR, JB, PS

Supervision: CCh, SS, PS

Writing - original draft: CCh, PS

Writing - review & editing: all authors contributed, reviewed and approved the last draft.

Funding information

This study was funded by Hyfe Inc. The sponsor participated in study design, collection and analysis; it had no role in the decision to submit this manuscript for publication. ISGlobal acknowledges support from the Spanish Ministry of Science and Innovation through the “Centro de Excelencia Severo Ochoa 2019-2023” Program (CEX2018-000806-S), and support from the Generalitat de Catalunya through the CERCA program.

Data integrity

CCh, ISO, MR and KW had full access to all of the data in this study and take complete responsibility for the integrity of the data and the accuracy of the data analysis.

Declaration of interests

MR, MG, LJ, MG, JB and PS are employees of Hyfe, Inc and own equity in Hyfe Inc. CCh has received consultancy fees and owns equity in Hyfe Inc. All other authors declare no conflict of interest.

Data sharing statement

Anonymized individual data and code relevant to this article will be publicly available in GitHub (https://github.com/hyfe-ai/validation/) as well as study protocol, immediately following publication, indefinitely.

References

1. Cornford CS. Why patients consult when they cough: a comparison of consulting and non-consulting patients. *Br J Gen Pract* 1998; 48: 1751-1754.

2. Irwin RS, French CT, Lewis SZ, Diekemper RL, Gold PM, Panel CEC. Overview of the management of cough: CHEST Guideline and Expert Panel Report. *Chest* 2014; 146: 885-889.

3. Altshuler E, Tannir B, Jolicoeur G, Rudd M, Saleem C, Cherabuddi K, Dore DH, Nagarsheth P, Brew J, Small PM, Glenn Morris J, Grandjean Lapierre S. Digital cough monitoring - A potential predictive acoustic biomarker of clinical outcomes in hospitalized COVID-19 patients. *J Biomed Inform* 2023; 138: 104283.

4. Crooks MG, den Brinker AC, Thackray-Nocera S, van Dinther R, Wright CE, Morice AH. Domiciliary Cough Monitoring for the Prediction of COPD Exacerbations. *Lung* 2021; 199: 131-137.

5. Gomes B, Ashley EA. Artificial Intelligence in Molecular Medicine. *N Engl J Med* 2023; 388: 2456-2465.

6. Hall JI, Lozano M, Estrada-Petrocelli L, Birring S, Turner R. The present and future of cough counting tools. *J Thorac Dis* 2020; 12: 5207-5223.

7. Loudon RG, Spohn SK. Cough frequency and infectivity in patients with pulmonary tuberculosis. *Am Rev Respir Dis* 1969; 99: 109-111.

8. Gabaldon-Figueira JC, Keen E, Gimenez G, Orrillo V, Blavia I, Dore DH, Armendariz N, Chaccour J, Fernandez-Montero A, Bartolome J, Umashankar N, Small P, Grandjean Lapierre S, Chaccour C. Acoustic surveillance of cough for detecting respiratory disease using artificial intelligence. *ERJ Open*

Res 2022; 8.

9. Gabaldon-Figueira JC, Keen E, Rudd M, Orrilo V, Blavia I, Chaccour J, Galvosas M, Small P, Grandjean Lapierre S, Chaccour C. Longitudinal passive cough monitoring and its implications for detecting changes in clinical status. *ERJ Open Res* 2022; 8.

10. Huddart S, Asege L, Jaganath D, Golla M, Dang H, Lovelina L, Derendinger B, Andama A, Christopher DJ, Nhung NV, Theron G, Denkinger CM, Nahid P, Cattamanchi A, Yu C. Continuous cough monitoring: a novel digital biomarker for TB diagnosis and treatment response monitoring. *Int J Tuberc Lung Dis* 2023; 27: 221-222.

11. Kang YR, Oh JY, Lee JH, Small PM, Chung KF, Song WJ. Long-COVID severe refractory cough: discussion of a case with 6-week longitudinal cough characterization. *Asia Pac Allergy* 2022; 12: e19.

12. Zimmer AJ, Ugarte-Gil C, Pathri R, Dewan P, Jaganath D, Cattamanchi A, Pai M, Grandjean Lapierre S. Making cough count in tuberculosis care. *Commun Med (Lond)* 2022; 2: 83.

13. Morice AH, Fontana GA, Belvisi MG, Birring SS, Chung KF, Dicpinigaitis PV, Kastelik JA, McGarvey LP, Smith JA, Tatar M, Widdicombe J, European Respiratory S. ERS guidelines on the assessment of cough. *Eur Respir J* 2007; 29: 1256-1276.

14. Kelsall A, Decalmer S, Webster D, Brown N, McGuinness K, Woodcock A, Smith J. How to quantify coughing: correlations with quality of life in chronic cough. *Eur Respir J* 2008; 32: 175-179.

15. Sanchez-Olivieri I, Rudd M, Gabaldon-Figueira JC, Carmona-Torre F, Del Pozo JL, Moorsmith R, Jover L, Galvosas M, Small P, Grandjean Lapierre S, Chaccour C. Performance evaluation of human cough annotators: optimal metrics and sex differences. *BMJ Open Respir Res* 2023; 10.

16. Rudd M, Song WJ, Small PM. The Statistics of Counting Coughs: Easy as 1, 2, 3? *Lung* 2022; 200: 531-537.

17. Galvosas MG-F, JC; Keen, EM; Orrillo, V; Blavia, I; Chaccour, J; Small, P; Giménez, G; Grandjean-Lapierre, S; Chaccour, C. Performance evaluation of the smartphone-based AI cough monitoring app

- Hyfe Cough Tracker against solicited respiratory sounds [version 1; peer review: 1 approved with reservations, 1 not approved]. *F1000Research* 2022; 11:730.

18. McGarvey LP, Birring SS, Morice AH, Dicpinigaitis PV, Pavord ID, Schelfhout J, Nguyen AM, Li Q, Tzontcheva A, Iskold B, Green SA, Rosa C, Muccino DR, Smith JA, Cough, Investigators C-. Efficacy and safety of gefapixant, a P2X(3) receptor antagonist, in refractory chronic cough and unexplained chronic cough (COUGH-1 and COUGH-2): results from two double-blind, randomised, parallel-group, placebo-controlled, phase 3 trials. *Lancet* 2022; 399: 909-923.

19. FDA. Pulmonary-Allergy Drugs Advisory Committee (PADAC) meeting on Gefapixant November 17, 2023. Full hearing available at: https://www.fda.gov/advisory-committees/advisory-committee-calendar/november-17-2023-meeting-pulmonary-allergy-drugs-advisory-committee-11172023, (Aaccessed February 2024). 2023.

20. Hegde S, Sreeram S, Alter IL, Shor C, Valdez TA, Meister KD, Rameau A. Cough Sounds in Screening and Diagnostics: A Scoping Review. *Laryngoscope* 2023.

21. Serrurier A, Neuschaefer-Rube C, Rohrig R. Past and Trends in Cough Sound Acquisition, Automatic Detection and Automatic Classification: A Comparative Review. *Sensors (Basel)* 2022; 22.

Figure Legends

Figure 1. Enrollment flow diagram of study.

Figure 2. Cumulative cough event-based performance results. Cumulative cough seconds from all participants over the course of the study with time. The ground truth human counts are shown in black and the Hyfe results are shown in red.

Figure 3. Individual performance results.

Figure 4. Cough event-based performance results. Shown are the event-based sensitivity and false positive rates for cough second detection. Each gray point is one subject, showing a range in sensitivity of 78.1% to 96.5% (y-axis) and false positivity rates from 0.38 to 2.23 false positives per hour (x-axis). The red point indicates the overall sensitivity of 90.4% and the overall FPR of 1.03 false positives per hour of the Hyfe CoughMonitor.

Figure 5. Cough rate-based performance results (Bland-Altman plot). Each point is one person-hour; its x-coordinate is the average of its manual and CoughMonitor hourly cough second counts, and its y-coordinate is their difference. The dashed horizontal lines indicate the bias (mean difference) of 0.23, the lower limit of agreement of -3.7, and the upper limit of agreement of 4.8.

Figure 6. Cough rate-based performance results (linear plot). Each point is one person-hour, the black dashed line is the OLS line of best fit, and the black solid line is the line of perfect agreement ($y=x$).

Tables

Table 1. Study inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria
1. Age 21 years old or above.	1. Inability to accept the privacy policy and terms of use of CoughMonitor due to confidentiality or other concerns.
2. Individuals who express concern about	2. Inability to avoid unusually prolonged

their active cough (problematic cough).	loud environments for the duration of the 24-hour study period.
3. Anticipate that they can collect auditory recordings and keep the devices with them continuously for 24 hours.	3. Need to conduct confidential conversations during the 24-hour monitoring period.
4. Residing in a domestic environment without unusually high and / or persistent background sound levels.	4. Individuals who have had significant change in antitussive therapy in the week preceding study.
5. Willing to wear a watch and audio recorder and keep them at bedside (within 3 ft from the mouth) during the night.	

Table 2. Demographic and cough monitoring results of study subjects.

Age, mean (range)	52 years (24-72)
Hours monitored, mean (range)	23.8 hours (22.9-25.6)
Hourly cough rate, mean (range)	8.15 (1.5-32)
Female, n (%)	15 (65%)
US-based, n (%)	12 (52%)
Diagnoses, n (%)	
Bronchitis	6 (26%)
NTM infection	5 (22%)
RCC/UCC	3 (13%)
Acute respiratory syndrome	2 (9%)
COPD	2 (9%)
COVID-19	2 (9%)
Bronchiectasis	1 (4%)
Asthma	1 (4%)
Cystic fibrosis	1 (4%)

NTM: nontuberculous mycobacteria. RCC/UCC: refractory/unexplained chronic cough

Figures

Figure 1

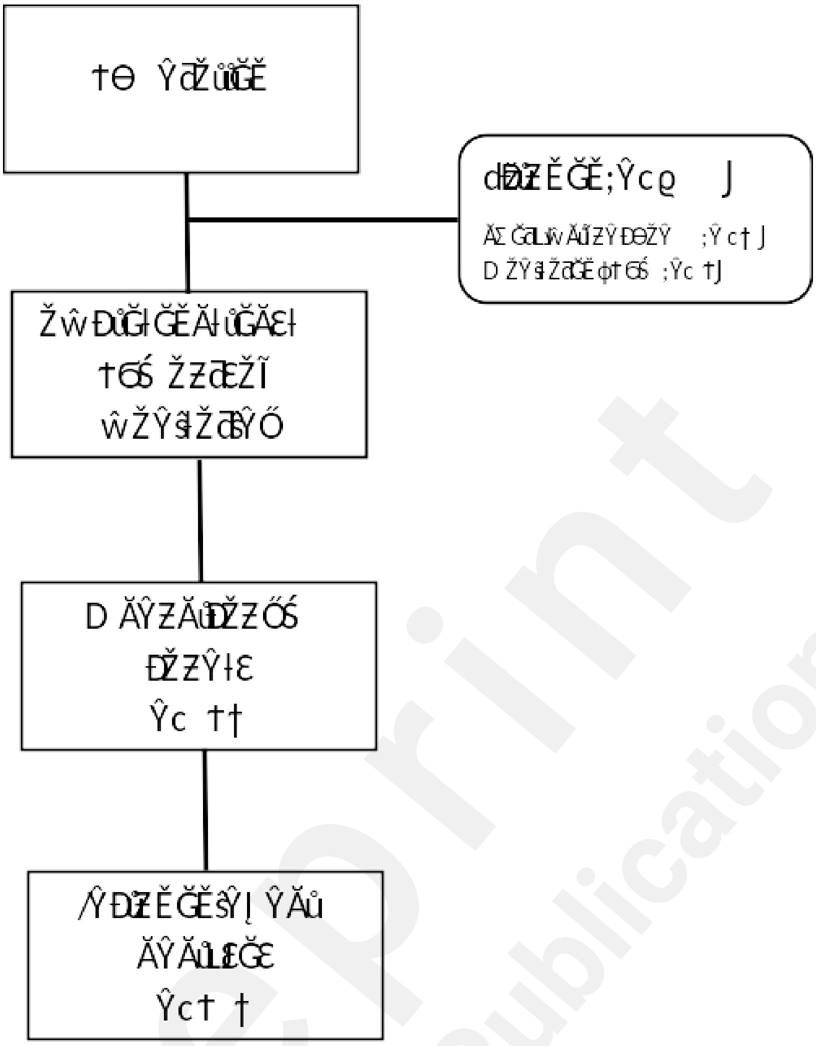


Figure 2

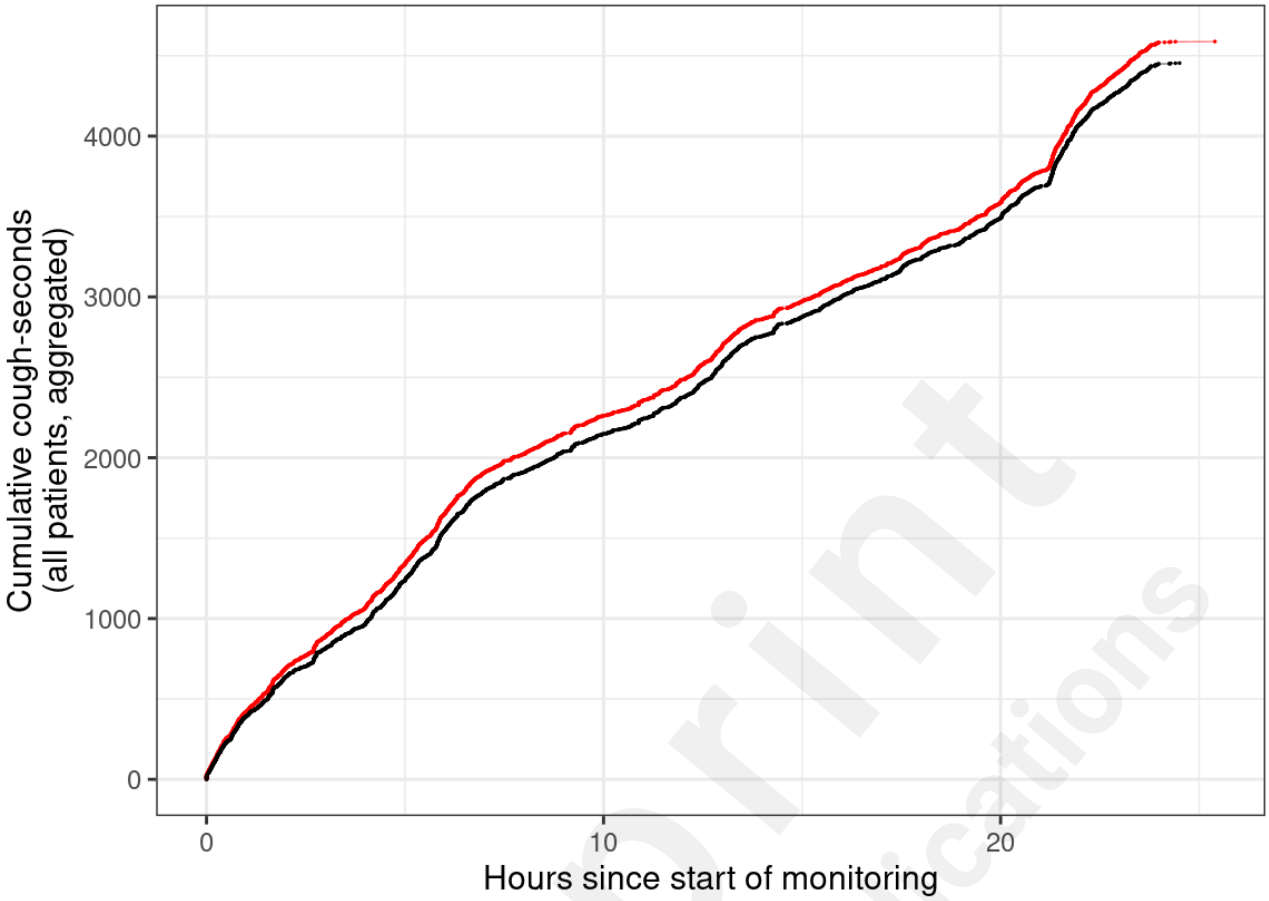


Figure 3

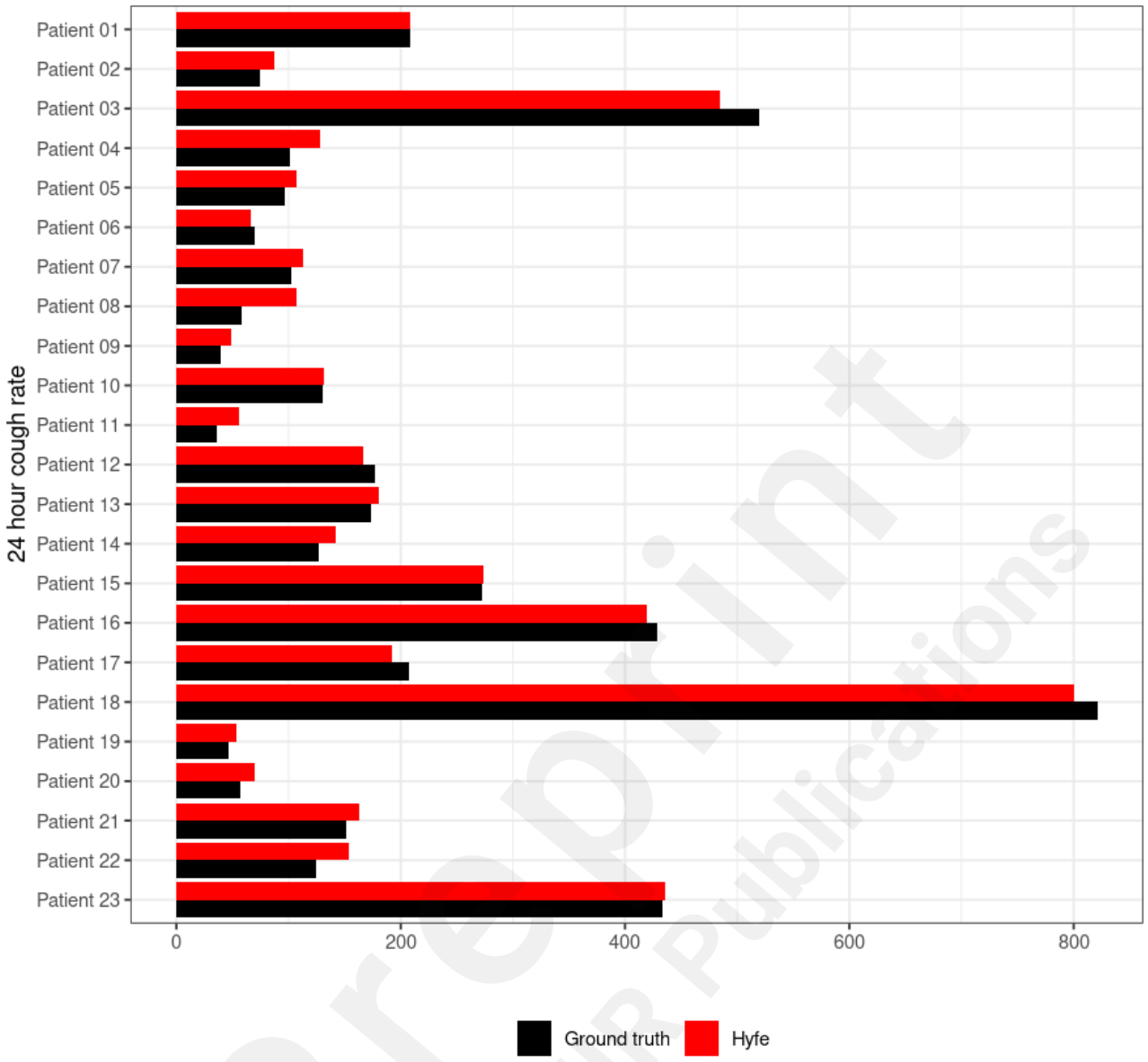


Figure 4

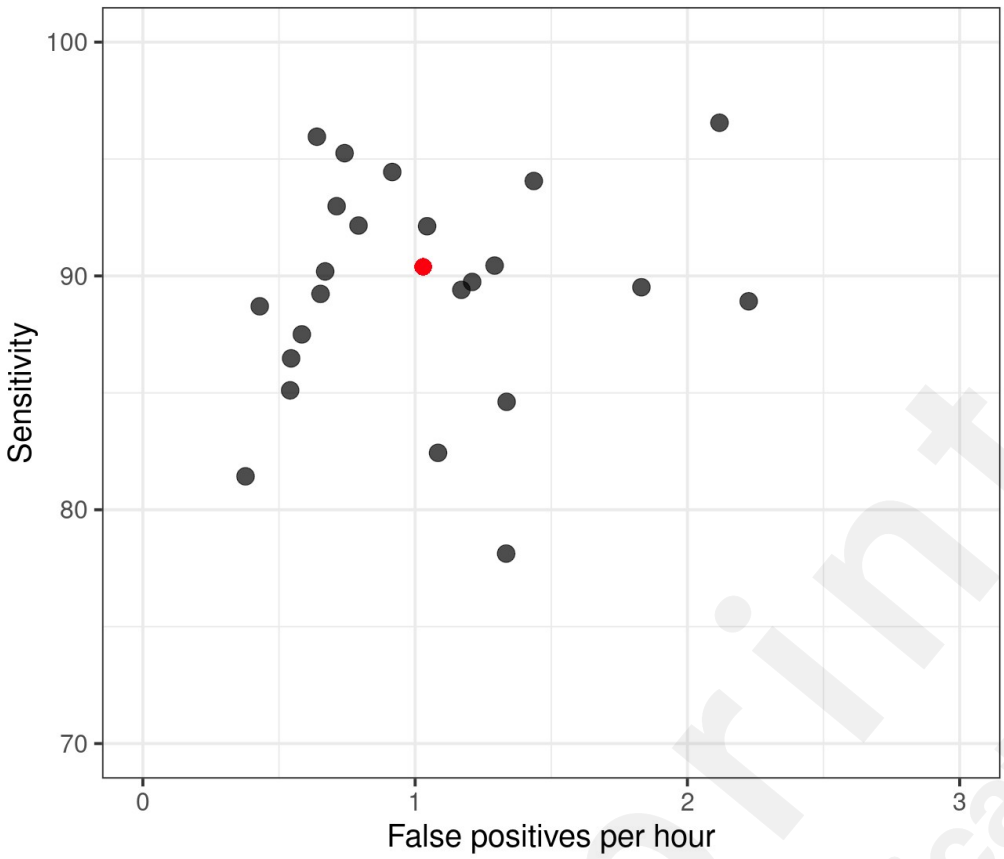


Figure 5

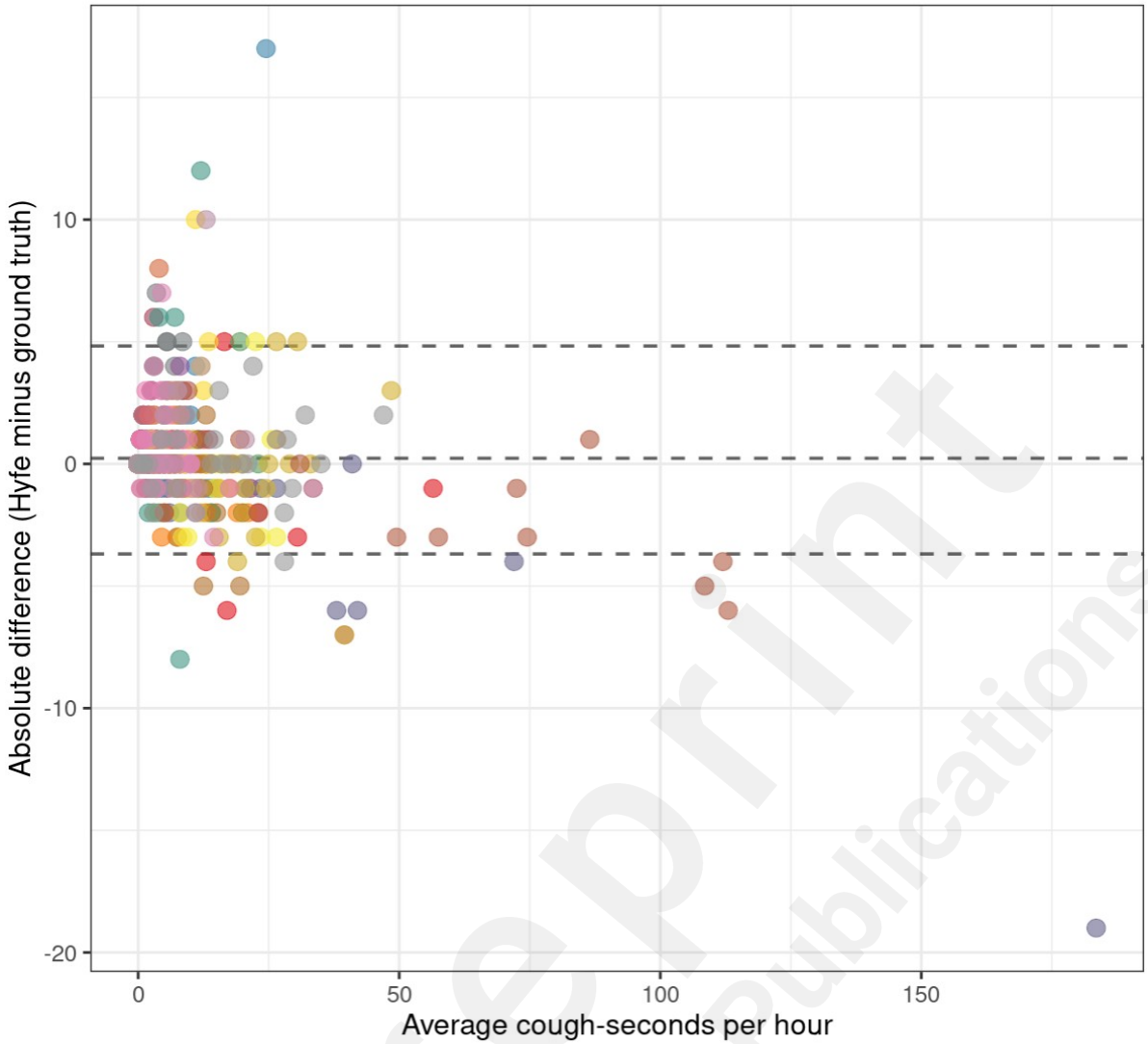
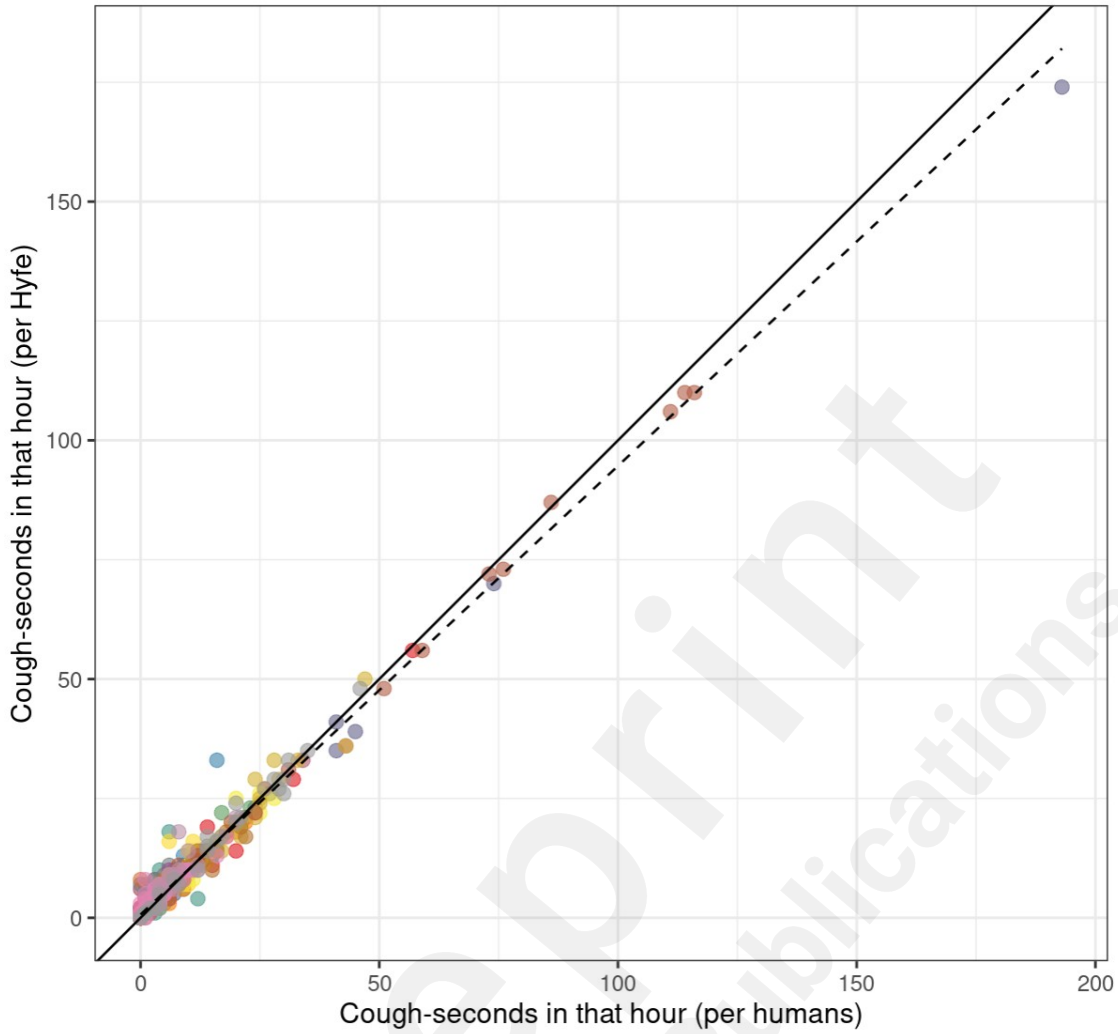


Figure 6



Supplementary Files

Multimedia Appendixes

Untitled.
URL: http://asset.jmir.pub/assets/310c1baeabb1e81197e08a2bdb083c49.docx

Untitled.
URL: http://asset.jmir.pub/assets/faa1c8aaf5b56e92a55a017b2f47d16b.docx