Acoustic surveillance for respiratory diseases: a prospective analysis of cough trends using artificial intelligence.

Juan Gabaldon-Figueira (✉ jgabaldonfi@unav.es)  
Clinica Universidad de Navarra

Eric Keen  
Hyfe Inc

Gerard Giménez  
Hyfe Inc

Virginia Orrillo  
Universidad de Navarra

Isabel Blavia  
Universidad de Navarra

Dominique Doré  
University of Montreal Hospital Center

Nuria Amendáriz  
Servicio Navarro de Salud

Juliane Chaccour  
Clínica Universidad de Navarra

Alejandro Fernández-Montero  
Clínica Universidad de Navarra

Javier Bartolomé  
Servicio Navarro de Salud

Nita Umashankar  
San Diego State University

Peter Small  
Hyfe Inc

Simon Grandjean Lapierre  
Université de Montréal  https://orcid.org/0000-0003-3646-1573

Carlos Chaccour  
Barcelona Institute for Global Health  https://orcid.org/0000-0001-9812-050X
Abstract

Syndromic surveillance for respiratory disease is limited by an inability to monitor its protean manifestation, cough. Advances in artificial intelligence provide the ability to passively monitor cough at individual and community levels. We hypothesized that changes in the aggregate number of coughs recorded among a sample could serve as a lead indicator for population incidence of respiratory diseases, particularly that of COVID-19. We enrolled over 900 people from the city of Pamplona (Spain) between 2020 and 2021 and used artificial intelligence cough detection software to monitor their cough. We collected nine person-years of cough aggregated data. Coughs per hour surged around the time cohort subjects sought medical care. There was a weak temporal correlation between aggregated coughs and the incidence of COVID-19 in the local population. We propose that a clearer correlation with COVID-19 incidence could be achieved with better penetration and compliance with cough monitoring.

Introduction

Syndromic surveillance relies on the recognition and reporting of symptomatic patients by healthcare systems and proves challenging for identification of emerging, rapidly transmissible pathogens, particularly in low-income settings. (1)

Cough is a key symptom of most respiratory diseases, including infections of public health interest, such as COVID-19, or influenza. Approximately 57% of all COVID-19 patients will develop cough during the early stages of infection and its presence correlates with contagiousness. (2, 3) The ubiquitous presence of cough in respiratory infections, as well as its relatively low frequency in healthy individuals make it an attractive marker for syndromic surveillance. (4) Individual-level cough patterns might help identify cases, thus allowing health authorities to rapidly deploy diagnostic resources and implement confinement protocols more efficiently. Early detection of sudden changes in community-level cough frequency might indicate the emergence or recrudescence of pathogens within communities.

To date, the limited portability and automation of existing objective cough monitoring devices hinders their use at scale for population monitoring. (5) Potential incorporation of cough monitoring into existing epidemiological surveillance programs is also hampered by the lack of practical experiences with acoustic surveillance and a limited understanding of the epidemiology of cough.

An attractive alternative is using artificial intelligence systems specially trained to recognize the well-described acoustic characteristics of cough. (6, 7) Cough is typically composed of three distinct phases, each one with its own characteristic acoustic pattern. (8) These patterns can be identified by a wide range of artificial intelligence (AI) algorithms, including machine learning techniques such as convolutional neural networks. (6, 9) The widespread use of smartphones provides an opportunity to deploy these systems at the scale needed to construct a community-based surveillance network. (10, 11)

We hypothesized that a mobile phone application (Hyfe Cough Tracker, henceforth referred to as Hyfe, https://www.hyfe.ai) could detect changes in an individual's clinical condition, (as demonstrated by
changes in cough frequencies around the dates they sought medical care) and that aggregated individual
data could be used as a proxy to estimate the incidence of respiratory infections, such as COVID-19 and influenza. Although cough detection systems have been used in clinical research for years, this is the first attempt to apply such a tool on a community scale to identify epidemiologically useful insights from disease-agnostic users. (12-15)

Results

Characteristics of the cohort

A total of 930 participants were enrolled. However, only 616 used the application for at least one hour, and were therefore included in further analysis. Although this is the largest cough cohort continually monitored to date, it represents just 1.7% of the 35,000 people estimated to have been reached in the recruitment campaign. Participants were aged 14-76 years, (Median: 21, IQR: 20-25), mostly female (64.9%) and iOS users (56%). In total, 178 (28.9%) participants registered more than 100 hours of monitoring, and 21 (3.4%) registered at least 240 hours, equivalent to ten days of continuous monitoring. The proportion of Android users increased in groups who used Hyfe for longer periods of time (44% of users with one hour, 71% of users with 100 hours or more, and 85.7% of those with 240 hours or more, $X^2 = 52.3$, df= 2, $p < 0.001$). These groups were also older (median age for those with 100 hours or more: 25, IQR: 21-50, median age for those with 240 hours or more: 50, IQR: 39-56), as presented in table 1.

Cough frequency is higher around the time of medical consultation

Two-hundred and seventy-two participants attended at least one medical consultation during the study period, 33 of whom had at least 24 hours of monitoring both during and outside the 10-day consultation period (days -5 to +4, day 0= consultation date). For these 33 patients, hourly cough rates were higher during the consultation period, than during the rest of the monitoring history, with a difference of 0.77 coughs/hour ($p$ value < 0.00001), equivalent to approximately 19 extra coughs per day. This effect was driven by lower cough rates during the post-consultation history (after day +4, difference: 1.08 coughs/hour, $p < 0.00001$). When exclusively compared to the pre-consultation history (before day -5), cough rates were not significantly different ($p= 0.855$). (Figure 1, table 2) Similar results were observed when comparing subdivisions of the consultation period to the rest of the monitoring history (supplementary material 1).

Syndromic surveillance and COVID-19 incidence

Over 79,000 aggregated hours of monitoring, equivalent to 3,316 person-days (or 9.08 person-years) and 62,401 coughs were registered between November 6th, 2020, and August 18th, 2021 (n= 616 participants). Peaks of coughs were registered in February, April-May, and August 2021 (Figure 2, panel B).

In total 14,051 cases of COVID-19 were diagnosed in the study area. During the study, three waves of COVID-19 occurred: in January and February, late March, and April, and between July and August 2021.
(Figure 2, panel A). Only 54 cases were diagnosed among study participants.

A descriptive analysis of cough frequency in the cohort and COVID-19 incidence in the population shows that those parameters are correlated, notably for cough and disease incidence peaks during the first and last waves (Figure 2, panel A and B). The ARIMA model analysis confirmed this association which reached maximum strength (ACF= 0.43) when lagging the cough time series by 17 days compared to the COVID-19 incidence. This indicates that cough increased, in average, 17 days after peaks in COVID-19 cases.

The number of users registering coughs changed throughout the study period such that changes in cough frequency measured were based on a diminishing number of participants (Figure 2, panel D). Specifically, during the February COVID-19 wave, approximately 40 participants were contributing a combined total of over 555 hours of monitoring time per day. At the time of the third peak there were only about 12 participants contributing 111 hours of monitoring time per day. Upon closer inspection, it was noted that almost half of all coughs registered during this wave came from a single chronic cougher who did not have COVID-19. Further investigation revealed that during this period her cough increased to approximately 400 times per day following the discontinuation of her antitussive medication. The smaller peak observed in February also coincides with this participant’s recruitment into the study.

While the peak in coughs during the summer remains evident after excluding this participant from the ARIMA model (Figure 2, panel C), the strength of the correlation markedly reduced (ACF= 0.28), indicating that given the low number of participants, trends were influenced by one person with severe chronic cough.

Usage and perception of the application

Participants with at least 1 hour of monitoring data (n=616), used the application for an average of 336 minutes per day (SD: +/- 188 minutes), equivalent to approximately 5.6 hours. Average daily usage was discretely increased in participants who received more email reminders ($\beta$=5 minutes per reminder, $p < 0.001$), and those older ($\beta$= 4.5 minutes per year, $p < 0.001$). On the contrary, using iOS rather than Android was associated with significantly reduced daily usage time ($\beta$= -103 minutes, $p < 0.001$), matching the increasing proportion of Android users in groups that monitored for longer periods of time (table 1). Full results for different predictors can be found in supplementary file 2.

The average satisfaction score given to the monitoring system was 3.6, n=217, SD: +/-1.02). Scores were also higher among Android users (3.82, 95%CI: 3.64 - 4.00), compared to iOS users (3.38, 95%CI 3.18-3.58, $p= 0.001$). The most common reason to withdraw were technical problems while running the application (95, 43.8%). Most of the participants who complained of technical issues were iOS users (60 out of 95).

Nine participants agreed to take part in FGDs or guided interviews. Participants were aged 21 to 65 years (median= 48 years). Two were male (22.2%) and seven female (77.8%). Seven of these participants
belonged to the high usage group (registered 100 hours or more), while two belonged to the low usage group. In the high participation group, cough was perceived as important only if it affected daily routines, either because it was associated with a known respiratory disease, or with certain lifestyle characteristics, such as smoking. The main motivator behind constant usage was interest in helping the study team, with little interest given to its perceived health benefits. However, two participants with a history of chronic respiratory disease indicated that seeing changes in longitudinal cough trends, and their link to certain behaviors, were important motivators. Notifications were not well received in the low participation group. Repeated bugs in the iOS version were noted in both groups, confirming results from the quantitative analysis. Summary tables with common answers provided by participants can be found in supplementary material 3.

Discussion

This is the first population-based syndromic surveillance study using passive digital cough monitoring at scale. Through a concerted community outreach campaign, we enrolled and digitally monitored 616 inhabitants of a community in Northern Spain for 10 months. Over the course of the study, we monitored over 9 years of person-time and detected 62,000 cough sounds. We showed that cough monitoring can detect changes in cough frequencies at individual and community levels.

Our observation that cough frequency is higher around the time that individuals seek medical care suggests that passively detected cough patterns are clinically relevant. These changes were driven by reductions after the 5 days following consultations. The lack of a significant difference with the pre-consultation history is likely a result of the fact that many of these participants were recruited during COVID-19 testing sessions at the University of Navarra's campus. Therefore, when recruited, many participants already presented cough and other respiratory symptoms. For these participants, the pre-consultation monitoring history was short, or non-existing, and likely included days during which respiratory symptomatology was already present, not reflecting their baseline cough frequency.

We also demonstrated that longitudinal changes in aggregate cough data from this cohort were temporally associated with the incidence of COVID-19 in the community. However, the hypothesized causal nature of this association is challenged by the fact that the cough frequency peaked, on average, 17 days after COVID-19 incidence. Although the duration of COVID-19 is highly variable, this period is longer than the 11 days within which most symptoms of mild infection resolve. (16)

Further confounding in this association came from a large proportion of coughs originated by a single individual with severe chronic cough. The lack of a stronger temporal association is likely caused by a low population coverage (1.7% of reached individuals), and the low incidence of COVID-19 in the cohort. It may also be due to a low percentage of patients experiencing cough, confounding by other infectious and non-infectious causes of cough within the cohort, or exposure to other environmental tussive stimuli.

Nonetheless, the fact that the association remains visible in more than one COVID-19 wave, and even after removing the chronic cougher, might indicate that some subjects in the cohort were infected in the
last part of a community wave and remained undiagnosed. The relevance of this signal should be further evaluated in studies that achieve a better penetration in the population.

Uptake and retention of the cough monitoring system represent different challenges. Reaching an adequate uptake to produce truly representative data seems particularly complicated, considering the low uptake observed in this study, as well as those reported for similar contact tracing software in the past. (17) The general lack of interest observed among participants and the qualitative results suggesting that participants with chronic respiratory disease are more likely than the general population to use these systems provide both an opportunity and a bias. Larger multicentric projects, ideally supported by public health authorities, might help increase the number of participants in future studies. Our observations provide an understanding of barriers needed to be overcome to address retention problems. The fact that technical problems were the leading cause for discontinuation is encouraging, as these are solvable engineering challenges, many of which have already been implemented. The human factors of maintaining interest and ensuring privacy may be harder to address.

Taken together these results indicate that acoustic surveillance systems, such as Hyfe, are technically capable of detecting changes in cough frequency associated with the onset and evolution of respiratory disease in individuals and populations. However, the use of these changes as a proxy for the incidence of infections in a community will require greater uptake and more constant use than we were able to obtain. This study provides some of the parameters needed to determine a minimally representative sample size using this tool for syndromic surveillance. With further validation, these tools are likely to be particularly valuable for the longitudinal monitoring of patients diagnosed with respiratory disease, or for the evaluation of participants in studies where changes in cough frequency are an outcome of interest.

AI-powered systems such as Hyfe can detect longitudinal changes in the cough frequency associated with the progression of respiratory disease. However, their practical utility in the context of epidemiological surveillance largely depends on a high uptake and a regular use. This could be achieved by enhancing user experience and with explicit collaboration of public health agencies and similar partners.

Online Methods

Study design and population

The primary objective of this study was to assess the value of digital cough surveillance as a proxy for community incidence of COVID-19 and other respiratory diseases. Secondary objectives were to (i) determine if changes in cough frequency were associated with the moment of medical consultation, and (ii) quantitatively and qualitatively assess the barriers and facilitators to participation in smartphone-based acoustic surveillance programs.

This was a prospective observational study. Participants were recruited between November 2020 and June 2021 in the campus of the University of Navarra in Pamplona (Navarra, Spain), as well as in the
neighboring communities of Zizur Mayor and the Cendea de Cizur, all located within a 5 km range. Participants included students and staff of the University of Navarra, as well as other residents of the area. Recruitment strategies included direct solicitation, community meetings and advertisements through municipal authorities, and the university’s communication platform and social networks. Through these activities, we expected to reach up to 30,000 people, roughly 14% of the study area’s total population.

A full protocol describing sample size estimations and enrollment strategies was previously published [NCT04762693]. (11) Interested individuals were invited to information sessions on the nature of the study and the Hyfe application. To be eligible, participants needed to be 1) 13 years or older, 2) own an Android or iOS smartphone able to run Hyfe, 3) be willing to install and use Hyfe as instructed, 4) accept and comply with Hyfe’s privacy policy and terms of use, 5) grant access to their medical records during their participation in the study, 6) visit the University of Navarra regularly, either as staff or students, or be a current resident of Navarra. All participants provided informed consent.

### Automated Cough Detection

Participants were enrolled and asked to monitor their cough using Hyfe, a freely available, automated cough detection application, downloaded on their personal phones. Hyfe runs in the background of phones, monitoring ambient sounds without continuously recording. It uses a convolutional neural network (CNN) to detect and analyze explosive sounds of 0.5 seconds corresponding to putative coughs. A cough prediction score is assigned to each sound by the machine learning model. If this score lies above a predetermined threshold (in this study 0.7 out of a maximum score of 1.0), the sound is classified as a cough, stored in the participant’s smartphone, and relayed to a cloud-based central dataset. Preliminary data indicates that Hyfe has a sensitivity of 96.34% and a specificity of 96.54% differentiating coughs from other detected sounds. (11) Further validation to determine its actual performance when undetected explosive sounds are accounted for is underway [NCT05042063].

Participants were able to turn Hyfe on and off at will but were instructed to keep it active for at least 6 hours per day, while they slept, so that at minimum, night-time coughing could be monitored, while minimizing interference with the normal use of the phone during the day. All participants were instructed to monitor their cough for a 30-day period, with the possibility of prolonging their participation if desired. Daily push notifications and periodic emails were sent to participants to maintain retention and provide study preliminary data.

### Review of clinical and epidemiological data

Study personnel reviewed medical records of enrolled participants every eight weeks in the private Clínica Universidad de Navarra and the regional public health system (Osasunbidea), looking for consultations associated with respiratory symptoms. Those facilities provide local healthcare services to the study population and offer general medicine and pneumology consultations, as well as COVID-19 PCR-based testing. During each round of reviews the national ID numbers of participants were searched in a
centralized dataset, and any registered consultation associated with unspecified respiratory symptoms (including COVID-19 screening tests) or a confirmed diagnosis of respiratory disease (COVID-19, influenza, respiratory syncytial virus, pneumonia, asthma, bronchitis, pharyngitis, chronic cough, chronic obstructive pulmonary disease, gastro-esophageal reflux disease, or other nonspecific respiratory tract infections) was recorded. Daily incidence of COVID-19 in the study area were obtained from public sources (18) and used to construct local epidemic curves.

Assessment of barriers and facilitators of usage

Upon withdrawal, participants were instructed to rate their appreciation of the digital cough monitoring application on a 0-5 scale (0 being very unsatisfactory, and 5 being very satisfactory), as well as to indicate the reason why they decided to withdraw from the syndromic surveillance study. Study participants were divided into two participation groups, high (≥100 hours of monitoring) and low participation users (<100 hours of monitoring). Demographic, medical, and mobile phone technology-related predictors of participation levels were identified.

Participants were invited to participate in virtual focus group discussions (FGDs) to evaluate the importance given to cough and their experience using the digital cough monitoring application.

Data analysis

Cough frequency and healthcare seeking behaviors

We defined the medical consultation period as a period of 10 days centered on the date of care seeking (days -5 to +4, with day 0 being the date of consultation). All data outside of the consultation period were defined as the user’s cough frequency history and was further divided into a pre- (before day -5) and post-consultation history (after day +4). Participants who attended at least one medical consultation during the enrollment period and for which at least 24 hours of cough monitoring was achieved within and outside the consultation period were included in the analysis. Changes in cough frequency from history to peri-consultation periods as well as between the pre- and post-consultation periods were analyzed.

Comparison tests were carried out using a randomization routine, which protects results against bias from individual-level effects (e.g., large differences in user activity and/or cough rates) while preserving the uncertainties inherent to low sample sizes. Cough rates during the consultation period and the participant’s whole history were calculated for each user, then their differences (consultation - history) were determined, yielding a delta cough rate for each user. The mean of these deltas is treated as the average effect size in the sampled population.

To determine the significance of this observed effect size, it was compared to a distribution of effect sizes that would be expected under a null model of no difference between the two levels. This was determined using 1,000 iterations of a randomization routine in which the user records were shuffled (specifically, the field indicating days since consultation), and the average of the users’ cough rate difference (consultation - history) was re-calculated. This routine produced a null distribution of
simulated effect sizes. The proportion of null values greater than the observed value is treated as a p-value.

**Acoustic surveillance and COVID-19 incidence**

Participants for which at least one hour of cough monitoring was achieved were included in this analysis. Cough was aggregated in time at cohort level to create a cough frequency curve. An epidemic curve including all cases of COVID-19 diagnosed in the study area was superposed to cough data. An autoregressive moving average (ARIMA) analysis was carried out to compare confirmed cases of COVID-19 with cough frequency in the cohort, measured as coughs per person-hour, and excluding participants with less than an hour of data on any specific day. The strength of the association between both variables was expressed with the auto-correlation function (ACF). This parameter ranges from -1 to +1, indicating the direction of the correlation with values closer to 1 representing a stronger association. This analysis was only carried out for COVID-19, due to the low circulation of other respiratory pathogens during the study period.

**Usage and perception of the acoustic syndromic surveillance system**

Predictors of regular use were evaluated using a linear regression model to compare differences in the average daily monitoring period by age, gender, phone operating system used (Android or iOS), number of medical consultations during the study period, and number of email reminders sent to each participant.

Mean appreciation scores from participants who completed the withdrawal questionnaire were disaggregated by mobile phone operating systems and compared with a two-tailed unpaired t-test. Barriers and facilitators for uptake and use of Hyfe were also qualitatively assessed in FGDs.

Data was organized and analyzed using Excel 365 (Microsoft Corporation, Redmond, WA), R Studio version 1.3 (RStudio Team, 2020. RStudio: Integrated Development for R. RStudio, PBC, Boston, MA), and SPSS Statistics version 28 (IBM Corp, Armonk, NY)

**Declarations**

**Authors contributions**

Conceptualization: SGL, JCGF, CCh

Methodology: SGL, JCGF, CCh, EK, GG, NU

Software: EK, GG

Formal analysis: JCGF, EK, GG, NU

Investigation: JCGF, CCh, VO, IB, NA, JCh, AFM, JB
Funding

This study was funded by the Patrick J McGovern Foundation (grant name: 'Early diagnosis of COVID-19 by utilizing Artificial Intelligence and Acoustic Monitoring'). CCh received salary support from Unitaid (grant name: BOHEMIA). SGL received salary support from the Fonds de Recherche en Santé Québec.

ISGlobal acknowledges support from the Spanish Ministry of Science and Innovation through the 'Centro de Excelencia Severo Ochoa 2019–2023' Programme (grant number: CEX2018-000806-S), and support from the Generalitat de Catalunya through the CERCA programme.

Data availability statement

Datasets including anonymized records with the date of medical consultation of individual participants, responses to the satisfaction questionnaire, and individual cough and monitoring data are provided as supplementary files to this article.

Code availability statement

The code used to run the analysis described above is provided as supplementary files to this article.

Conflicts of interest:

EK, GG, and PS are employees of Hyfe, Inc. CCh has received consultancy fees and owns equity in Hyfe Inc. All other authors declare no conflict of interest.

Ethics statement

This project has been reviewed and approved by the Ethics Committee for Medical Research of the Chartered Community of Navarra (Pamplona, Spain), under the reference number: PI2020/107 and the Comité d’éthique à la recherche du Centre Hospitalier de l’Université de Montréal (Montréal, Canada), under the reference number: 2021-9247

References
1. Chowdhury R, Luhar S, Khan N, Choudhury SR, Matin I, Franco OH. Long-term strategies to control COVID-19 in low and middle-income countries: an options overview of community-based, non-pharmacological interventions. European Journal of Epidemiology. 2020 2020/08/01;35(8):743-8.

2. Grant MC, Geoghegan L, Arbyn M, Mohammed Z, McGuinness L, Clarke EL, et al. The prevalence of symptoms in 24,410 adults infected by the novel coronavirus (SARS-CoV-2; COVID-19): A systematic review and meta-analysis of 148 studies from 9 countries. PLoS One. 2020;15(6):e0234765-e. PubMed PMID: 32574165. eng.

3. Agrawal A, Bhardwaj R. Probability of COVID-19 infection by cough of a normal person and a super-spreader. Physics of Fluids. 2021 2021/03/01;33(3):031704.

4. Grandjean Lapierre S. Making Cough Count in Tuberculosis Care: Under review in Communications Medicine; 2021.

5. Cho PSP, Birring SS, Fletcher HV, Turner RD. Methods of Cough Assessment. The Journal of Allergy and Clinical Immunology: In Practice. 2019 2019/07/01/;7(6):1715-23.

6. Pramono RXA, Imtiaz SA, Rodriguez-Villegas E. A Cough-Based Algorithm for Automatic Diagnosis of Pertussis. PLoS One. 2016;11(9):e0162128.

7. Lee KK, Matos S, Ward K, Rafferty GF, Moxham J, Evans DH, et al. Sound: a non-invasive measure of cough intensity. BMJ Open Respiratory Research. 2017;4(1):e000178.

8. Korpás J, Sadlonová J, Vrabec M. Analysis of the cough sound: an overview. Pulmonary pharmacology. 1996 Oct-Dec;9(5-6):261-8. PubMed PMID: 9232662. Epub 1996/10/01. eng.

9. Liu J, You M, Wang Z, Li G, Xu X, Qiu Z, editors. Cough detection using deep neural networks. 2014 IEEE International Conference on Bioinformatics and Biomedicine (BIBM); 2014 2-5 Nov. 2014.

10. Kvapilova L, Boza V, Dubec P, Majernik M, Bogar J, Jamison J, et al. Continuous Sound Collection Using Smartphones and Machine Learning to Measure Cough. Digit Biomark. 2019;3(3):166-75. PubMed PMID: 32095775. eng.

11. Gabaldon-Figueira JC, Brew J, Doré DH, Umashankar N, Chaccour J, Orrillo V, et al. Digital acoustic surveillance for early detection of respiratory disease outbreaks in Spain: a protocol for an observational study. BMJ Open. 2021;11(7):e051278.

12. Larson S, Comina G, Gilman RH, Tracey BH, Bravard M, López JW. Validation of an Automated Cough Detection Algorithm for Tracking Recovery of Pulmonary Tuberculosis Patients. PLoS One. 2012;7(10):e46229.

13. Ryan NM, Birring SS, Gibson PG. Gabapentin for refractory chronic cough: a randomised, double-blind, placebo-controlled trial. The Lancet. 2012 2012/11/03/;380(9853):1583-9.
14. Smith JA, Kitt MM, Morice AH, Birring SS, McGarvey LP, Sher MR, et al. Gefapixant, a P2X3 receptor antagonist, for the treatment of refractory or unexplained chronic cough: a randomised, double-blind, controlled, parallel-group, phase 2b trial. The Lancet Respiratory Medicine. 2020 2020/08/01/;8(8):775-85.

15. Hsu JY, Stone RA, Logan-Sinclair RB, Worsdell M, Busst CM, Chung KF. Coughing frequency in patients with persistent cough: assessment using a 24 hour ambulatory recorder. European Respiratory Journal. 1994;7(7):1246.

16. Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, et al. The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. Annals of Internal Medicine. 2020 2020/05/05;172(9):577-82.

17. Toussaert S. Upping uptake of COVID contact tracing apps. Nature Human Behaviour. 2021 2021/02/01;5(2):183-4.

18. Positivos covid-19 por PCR distribuidos por municipio: Gobierno de Navarra; 2021 [cited 2021 September 1st]. Available from: https://gobiernoabierto.navarra.es/es/open-data/datos/positivos-covid-19-por-pcr-distribuidos-por-municipio.

Tables

**Table 1: Cohort characteristics and cough monitoring periods**

|                         | Recruited  | One hour | 100 hours | 240 hours |
|-------------------------|------------|----------|-----------|-----------|
| **Median age (IQR)**    | 21 (20-24) | 21 (20-25)| 25 (21-50)| 50 (39-56)|
| **Operating System**    |            |          |           |           |
| Android Users (%)       | 366 (39.4) | 269 (43.7)| 127 (71.3)| 18 (85.7) |
| iOS users (%)           | 559 (60.1) | 345 (56) | 51 (28.7) | 3 (14.3)  |
| Not specified (%)       | 5 (0.5)    | 2 (0.3)  | 0         | 0         |
| **Total**               | 930        | 616      | 178       | 21        |

**Table 2: Changes in cough rates during the consultation period and participant monitoring history**
### Figures

**Figure 1**

**Difference between cough rates in the consultation period compared to the participants’ monitoring history** Cough frequency during the consultation period is compared to the entire monitoring history (n=33), and the parsed pre- (n=23) and post-consultation history (n=29). Shaded areas represent the
distribution of effect sizes predicted under a null model of no difference. The black line represents the actual observed difference between the consultation period and compared periods. Cough frequency during the consultation period significantly increased when compared to the entire history and the post-consultation history ($p < 0.00001$ in both cases), but not when compared to the pre-consultation history ($p=0.855$).
Cough and usage trends compared to COVID-19 incidence. Incidence of COVID-19 in the entire study area (A) compared to the evolution of cough trends in the monitored cohort (B), after the exclusion of the participant with chronic cough (C) and compared to the number of active users (D). The continuous line represents the 7-day rolling average.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementaryfile1.pdf
- supplementaryfile2.xlsx
- supplementaryfile3.xlsx
- consultationsforpublic.zip
- predictorsdataset.csv
- satisfactionquestionnaire.csv
- ARIMApub.rar