Figures

TABLE OF CONTENT

Figure 1........................................................................................................................................p. 2
Figure 2........................................................................................................................................p. 3
Figure 3........................................................................................................................................p. 4

LIST OF INVESTIGATORS

Joe Brew MPH, Astrid M Knoblauch PhD, Lai Yu Tsang, Andry Andriamiadanarivo, Niaina Rakotosamimanana PhD, Peter M Small MD, Simon Grandjean Lapierre MD
Figure 1 – Temporal Patterns of Cough

Temporal distribution of coughs

[Bar charts showing temporal patterns of coughs for 18 persons, labeled 1 to 18. The x-axis represents time of day (0:00 to 24:00), and the y-axis represents the number of coughs (0 to 150). The charts display varying patterns of cough occurrence throughout the day.]
Figure 2 – Spatial Patterns of Cough
Figure 3 - Cough Exposure Risk Heatmap
TITLE
Mobile Phone Digital Cough Monitoring for Respiratory Disease Surveillance

AUTHORS
Joe Brew MPH¹, Astrid M Knoblauch PhD²,³, Lai Yu Tsang⁴, Andry Andriamiadanarivo⁵, Niaina Rakotosamimanana PhD³, Peter M Small MD⁴, Simon Grandjean Lapierre MD⁶

AFFILIATIONS
¹Databrew, LLC. 715 NE 2nd Street, Gainesville, FL, 32601 USA.
²Swiss Tropical and Public Health Institute, P.O. Box, 4002 Basel, Switzerland.
³Institut Pasteur de Madagascar, Ambatofotsikely, Antananarivo 101, Madagascar
⁴Stony Brook University, Global Health Institute, N-203 Social and Behavioral Sciences Building, Stony Brook, New York, 11794 USA.
⁵ValBio Research Center, Ranomafana, BP 33, Madagascar.
⁶Centre de Recherche du Centre Hospitalier de l’Université de Montréal, 900 Saint-Denis, Montréal, Québec, Canada.

CORRESPONDING AUTHOR
Simon Grandjean Lapierre
Simon.grandjean-lapierre.chum@ssss.gouv.qc.ca / +1.514.890.8000 #20935

WORDS NUMBER
702
ABSTRACT

Background
Cough, which is often ignored, is an increasingly important symptom in the context of Covid-19.

Methods
We used a smartphone app to collect putative cough sounds in remote Madagascar villages. We developed an artificial intelligence algorithm to verify which were actual coughs. We plotted cough data in space and time to describe its epidemiology.

Results
Cough was ubiquitous and common (median of 17.6 per hour), but the temporal and spatial distribution of coughing varied considerably between subjects.

Conclusions
Using mobile phones to detect and monitor coughing could be an important tool for clinical medicine and public health. However, pragmatic challenges and policy issues need urgent attention for this to become a practical approach.
BACKGROUND

Cough is a common symptom that often goes unnoticed or ignored \(^1\). However, because it is a presenting symptom in up to 75\% of Covid-19 infections, increased attention and better cough monitoring approaches are important in controlling the pandemic \(^2\).

Digital and mobile phone based technologies are capable of detecting cough as well as determining the etiology and severity of pulmonary disease \(^3,4\). In the context of Covid-19, this information could be used to promote early self-isolation and health-seeking, thereby preventing transmission especially for undocumented cases which are significantly contributing to the burden of disease \(^5\). Anonymized geocoded trends in cough frequency could alert public health practitioners to the early detection of geographic hotspots and help inform containment policies. Furthermore, real time awareness of a patient’s cough frequency could help health care providers remotely monitor disease progression in their patients thus limiting the need for high-risk person-to-person interactions \(^6\). Smartphone based detection technology is potentially simple and robust enough to be deployed at scale globally \(^7\).

METHODS

In July 2018, we launched an iOS mobile phone application (ResApp Health™) to explore the relationship between cough, tuberculosis and indoor air pollution in villages of remote Madagascar. The phones recorded putative cough sounds and their geographic coordinates from 18 consenting adult volunteers for 2 days each (6 tuberculosis patients and 12 controls). We subsequently listened to the 31,704 sound files that were recorded and classified each as ‘cough’ (23\%), ‘not cough’ (75\%) or ‘undetermined’ (2\%). With the resulting dataset [see supplementary
appendix 1], we trained a convolutional neural network to generate an automated cough classifier. The classifier performed well on an independent test set of 1,832 putative coughs with a 60/40 no-cough/cough split, correctly categorizing 92.1% of sounds, with a sensitivity of 91.3% and specificity of 92.6% (Area Under the Curve of 96.9%) [see supplementary appendix 2].

RESULTS
We found cough to be ubiquitous and frequent with villagers coughing a median of 17.6 (SD: ± 11.2) times each hour. We mapped coughs over time and found significant (as measured through cross-wise Kolmogorov-Smirnov tests) person-to-person variability in the circadian pattern of cough (Figure 1). While merely anecdotal, it is interesting that a villager with tuberculosis (subject 6) had a relatively constant cough, while others without tuberculosis (subjects 2, 4 and 18) have more intermittent coughs. The relative concentration vs. dispersion of coughs over space and time is consistent with illness versus cooking-related indoor air pollution provoked cough episodes. At a population level, the combination of spatial and temporal variations in cough could be a tool for disease surveillance and early outbreak detection. Figure 3 [animated in supplementary appendix 3] shows a cough heatmap that integrates both time and space, thus depicting areas at higher risk of exposure to cough and presumably, airborne disease transmission.

CONCLUSION
We have shown that it is feasible to use mobile phones to detect and monitor coughs. Because mobile phone-based technology is rapidly scalable we propose that such an approach can facilitate screening, detection and surveillance of Covid-19. It could be integrated into telemedicine or other mobile phone-based applications for Covid-19 or other infectious diseases. This acoustic
surveillance could be a powerful public health tool in high income countries with broad network coverage or in low-income countries where traditional surveillance is difficult to deploy. However, doing so will require rapid progress in solving technical and pragmatic challenges. The generalizability of our model is limited by the fact that it was trained on data from only one setting and more AI training is needed with more diverse data sets before it can be implemented more widely. To facilitate rapid improvement and deployment, we are making our annotated cough data and cough classifier software code open access and encourage others to do the same.

The Covid-19 pandemic presents a unique opportunity for technologists, clinicians and public health authorities to rapidly collect and share data to bring cutting edge artificial intelligence-based approaches to improve health globally. The urgency of doing so, however, must not compromise the principles of privacy, equity, and freedom, and must include mechanisms that prevent the proliferation of bias and spurious solutions. Addressing those and other practical implementation challenges now will lead to increased pandemic preparedness for the future.
FIGURES

Figure 1 – Temporal Patterns of Cough

Figure 1 - Temporal patterns of cough frequency (bar height) by time of day (horizontal axis) between the participants (numbered panels). There is a significant difference in the diurnal patterns, with some fairly constant (e.g. subject 6) and others episodic (e.g. subjects 2, 4 and 18).

Figure 2 – Spatial Patterns of Cough

Figure 2 - Spatial pattern of cough showing the movement (lines and green markers) and coughs (red markers) of two Malagasy villagers. Participant 2 (left) traveled widely but his/her coughs were concentrated in space/time, a pattern suggesting indoor air pollution exposure. Participant 8 (right) remained near his/her village, and had a more uniform temporal distribution of coughs, a pattern more compatible with respiratory illness. Note: coordinates were re-centered so as to hide true locations.

Figure 3 - Cough Exposure Risk Heatmap

Figure 3 - A cough heatmap that integrates both time and space, thus depicting areas at higher (red) and lower (blue) risk of exposure to cough, and presumably, disease prevalence and likelihood of airborne disease transmission.
FINANCIAL ASSOCIATIONS / CONFLICT OF INTEREST

This work was supported in part by the Stop TB Partnership’s TB REACH wave 5 award to NR and PMS (http://www.stoptb.org/global/awards/tbreach/wave5.asp). The Stop TB Partnership is funded by the Government of Canada and the Bill and Melinda Gates Foundation. LYT was supported by a New York Academy of Medicine David E. Rogers Student Fellowship Award. iOS devices and recording software were provided by ResApp™ Health.

ETHICAL REVIEW

This work was approved by the Stony Brook University Internal Review Board (CORIHS# 2017-4056-F) and the Madagascar’s Ministry of Health “Comité d’Éthique de la Recherche Biomédicale” in Madagascar (073-MSANP/CERBM). All participants provided written consent for cough recording.

CONFLICTS OF INTEREST

Joe Brew (Databrew) has been working with others to develop Hyfe App, an impact first effort to create phone apps to collect open source data to accelerate cough monitoring. Simon Grandjean Lapierre is an uncompensated scientific advisor to that effort. Peter Small, (currently employed at Global Good, Bellevue WA.) is supporting multiple efforts to accelerate open access acoustic data collections.
AUTHORSHIP

All cited authors have substantially contributed to the conception and design of this study along with data collection, analysis and interpretation and thus fulfill the ICJME prerequisites for authorship. Also, all individuals meeting the ICJME prerequisites for authorship are named as article authors. All authors have approved the final version to be published and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the article are appropriately investigated and resolved.
REFERENCES


SUPPLEMENTARY APPENDIX

TABLE OF CONTENT
Supplementary Appendix 1 ................................................................. p. 2
Supplementary Appendix 2 ................................................................. p. 2
Supplementary Appendix 3 ................................................................. p. 3

LIST OF INVESTIGATORS
Joe Brew MPH, Astrid M Knoblauch PhD, Lai Yu Tsang, Andry Andriamianarivo, Niaina Rakotosamimanana PhD, Peter M Small MD, Simon Grandjean Lapierre MD
Supplementary Appendix 1

Supplementary Appendix 1 contains the complete data set of putative coughs recorded. Each putative cough sound file is matched with a denominative ID (linkable to audio file) and human classification (cough vs. not cough).

Labels: [http://droneots.com/labels](http://droneots.com/labels)

Sounds: [http://droneots.com/sounds](http://droneots.com/sounds)

Password:

Supplementary Appendix 2

Supplementary Appendix 2 contains the software code for our cough classification, as well as full model evaluation results: [https://github.com/joebrew/coughtracker](https://github.com/joebrew/coughtracker)
Supplementary Appendix 3

An annotated cough heatmap that depicts areas at higher (red) and lower (blue) risk of exposure to cough over time in a defined region.