Precision Assessment of COVID-19 Phenotypes Using Large-Scale Clinic Visit Audio Recordings: Harnessing the Power of the Patient Voice

Paul J Barr, James Ryan, Nicholas C Jacobson

Submitted to: Journal of Medical Internet Research
on: June 03, 2020

Table of Contents

Precision Assessment of COVID-19 Phenotypes Using Large-Scale Clinic Visit Audio Recordings: Harnessing the Power of the Patient Voice

Paul J Barr[1, 2] PhD, MScPH; James Ryan[3] DO; Nicholas C Jacobson[1, 4, 5] PhD

[1]The Center for Technology and Behavioral Health Geisel School of Medicine at Dartmouth College Lebanon US
[2]The Dartmouth Institute for Health Policy & Clinical Practice Geisel School of Medicine at Dartmouth College Lebanon US
[3]Ryan Family Practice Ludington US
[4]Biomedical Data Science Geisel School of Medicine at Dartmouth College Lebanon US
[5]Department of Psychiatry Geisel School of Medicine at Dartmouth and Dartmouth Hitchcock Health Lebanon US

Corresponding Author:
Paul J Barr PhD, MScPH
The Dartmouth Institute for Health Policy & Clinical Practice
Geisel School of Medicine at Dartmouth College
46 Centerra Parkway
Lebanon
US

Abstract

The novel coronavirus (SARS-CoV-2) and its related disease, COVID-19, are exponentially increasing across the world, yet there is still uncertainty about the clinical phenotype. Natural Language Processing (NLP) and machine learning may hold one key to quickly identify individuals at high risk for COVID-19 and understand key symptoms in its clinical manifestation and presentation. In healthcare, such data often come the medical record, yet when overburdened, clinicians may focus on documenting widely reported symptoms that appear to confirm the diagnosis of COVID-19, at the expense of infrequently reported symptoms. A comprehensive record of the clinic visit is required—an audio recording may be the answer. If done at scale, a combination of data from the EHR and recordings of clinic visits can be used to power NLP and machine learning models, quickly creating a clinical phenotype of COVID-19. We propose the creation of a pipeline from the audio/video recording of clinic visits to the clinical symptomatology model and prediction of COVID-19 infection. With vast amounts of data available, we believe a prediction model can be quickly developed that could promote the accurate screening of individuals at risk of COVID-19 and identify patient characteristics predicting a greater risk of a more severe infection. If clinical encounters are recorded and our NLP is adequately refined, then benchtop-virology will be better informed and risk of spread reduced. While recordings of clinic visits are not the panacea to this pandemic, they are a low cost option with many potential benefits that have only just begun to be explored.

(JMIR Preprints 03/06/2020:20545)
DOI: https://doi.org/10.2196/preprints.20545

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

Title: Precision Assessment of COVID-19 Phenotypes Using Large-Scale Clinic Visit Audio Recordings: Harnessing the Power of the Patient Voice

Authors: Paul J. Barr PhD MScPH,[1,2] James Ryan DO,[3] Nicholas C. Jacobson PhD[2,4,5]

Affiliations:

1. The Dartmouth Institute for Health Policy & Clinical Practice, Geisel School of Medicine at Dartmouth College, Lebanon NH, USA

2. The Center for Technology and Behavioral Health, Geisel School of Medicine at Dartmouth College, Lebanon NH, USA

3. Ryan Family Practice, Ludington MI, USA

4. Biomedical Data Science, Geisel School of Medicine at Dartmouth College, Lebanon NH, USA

5. Department of Psychiatry, Geisel School of Medicine at Dartmouth and Dartmouth Hitchcock Health, Lebanon NH, USA

Correspondence:
Paul J. Barr, PhD, MSc PH
The Center for Technology and Behavioral Health
46 Centerra Parkway, Lebanon, New Hampshire, 03766
Phone: +1-603-653-0863
Email: paul.j.barr@dartmouth.edu

Conflict of interest: None to report

Acknowledgments
We would like to acknowledge the review of this article by Prof. Lisa Marsch and editing by Dr. Jesse Schoonmaker.

Challenges in identifying COVID-19 Clinical Phenotypes.

The novel coronavirus (SARS-CoV-2) and its related disease, COVID-19, are exponentially

increasing across the world, yet there is still uncertainty about the clinical phenotype. A clinical phenotype is the observable characteristics (i.e., symptoms) of a disease in a given individual. A meta-analysis of COVID-19 related symptom presentations found that the most frequent clinical symptoms are fever, cough, fatigue, and dyspnea.[1] However, the meta-analysis found I^2 statistics of 84.9 – 96.4% which suggests that there is considerable heterogeneity between studies, and this is likely suggestive of the extreme heterogeneity of these symptoms at the individual patient level. Other less frequent symptoms of COVID-19 have been reported including anosmia, dysgeusia, headache, sore throat, rhinorrhea, diarrhea, nausea and myalgias.[2] Yet, it is unclear as to their significance, prevalence, and importance.

As such, a traditional reductionist approach to finding treatments for COVID-19 is not as simple as extrapolating the current set of knowledge toward our limited model of the novel virus. Clinical treatments often arise from the set of biochemical agents currently developed and discovering less frequent symptoms of a disease may lead to a biochemical pathway that can be treated with a pharmacologic substance that may not have been considered. Only laboratory tests can confirm the presence of COVID-19, but such tests are in short supply. This presents an unprecedented need to develop better assessment methods to identify and synthesize heterogeneity in the clinical profile of COVID-19, and other viruses across the entire healthcare system. The urgency of this need cannot be understated, as it holds a key to understand how to more accurately identify and treat COVID-19.

Using 'big data' to understand the clinical manifestations of COVID-19

Natural Language Processing (NLP) and machine learning may hold one key to quickly identifying individuals at high risk for COVID-19 and understanding key symptoms in its clinical manifestation and presentation.[3] Existing applications of NLP and machine learning in the field of diagnostics in medicine are based on a combination of structured (e.g., symptom codes, medications, lab results)

and unstructured (e.g., visit notes, radiology reports) data captured by clinicians in the electronic health record (EHR). Utilizing NLP and machine learning, data on documented signs and symptoms in the EHR are already used to identify clinical conditions (computational phenotyping).[4] Such NLP efforts are currently being applied to unstructured text data captured in the EHR from telehealth visits to develop better screening tools for COVID-19.[5] Ancillary data can improve the accuracy of computational phenotyping, such as information from disease registries. However, the performance of any model is determined by the quality of data used to create it and concerns exist about the fullness of data captured in the EHR.

The limitations of EHR reported data

This considerable degree of symptom heterogeneity reported by patients with COVID-19 can present problems for the accurate documentation of less frequently reported symptoms in the EHR. Inaccuracy in EMR documentation is not a new phenomenon; an analysis of 105 clinic encounters, 90% of clinician notes had at least one error, including 636 documentation errors, comprising 181 charted findings that did not take place and 455 findings that were not charted.[6] Data on such symptoms may not be accurately synthesized into patient records due to the pressing need to treat patients in overburdened health care settings. In this scenario, clinicians may focus on documenting widely reported symptoms that appear to support the diagnosis of COVID-19, at the expense of infrequently reported symptoms. This is not surprising as overburdened clinicians are more likely to be affected by 'cognitive biases' such as anchoring and confirmation bias.[7] Additionally, it has recently been reported that ICD-10 codes, the mainstay of EMR documentation, do not adequately capture COVID-19 related symptoms.[8] While NLP solutions can play a key role in creating clinical phenotypes of COVID-19, they are limited by the resulting limitations in EHR data. A comprehensive record of the clinic visit is required—an audio recording may be the answer.[9]

Clinical phenotypes based on audio recordings of clinic visits

A small, but growing number of health systems are routinely audio recording, and, in some cases, video recording clinic visits.[9,10] For example, human scribes are commonly employed to review recordings of clinic visits and make detailed notes, reducing documentation burden on clinicians and improving the accuracy of data entered to the EHR. A recording of the clinic visit represents a more comprehensive and accurate record of patient-reported symptoms. If done at scale, a combination of data from the EHR and recordings of clinic visits can be used to power NLP and machine learning models, quickly creating a clinical phenotype of COVID-19 and its subsequent strains. In addition to a more comprehensive record of symptoms discussed, recordings also asynchronously collect additional ancillary information such as type and frequency of cough, that may be used to improve the precision of phenotyping.

Creating NLP and machine learning models requires 1) the transcription of vast quantities of clinic visit conversations of patients being investigated for COVID-19 (with latter confirmatory lab testing for the presence of the disease) and 2) annotation of these transcripts by annotators trained to identify symptom mentions. There have been significant improvements in the performances of automated speech recognition algorithms,[10] allowing the real-time use of audio data, rather than more time consuming human transcription of the audio data. Real-time risk assessment is critical when responding to an infectious disease such as COVID-19, it allows for individuals to identify their risk level and more quickly isolate, reducing risk of spread. The annotation of the data to create models that can accurately identify symptoms, does not come without challenges—many of which have been nicely summarized by Quiroz et al[11]. It can be difficult for annotators to identify vaguely mentioned symptoms from unstructured natural language clinic visit conversations, with a negative impact on model performance. Rigorous training of annotators can help mitigate this challenge, yet such training and annotation is time-consuming and would require a large team of annotators to quickly

meet the immediate need for this analysis. In addition, model training requires human input and time. Also, creating optimal data will require data which continues to be refined, wherein suspected cases are replaced by the results of confirmatory tests, so as not to correspond to clinician views or biases.

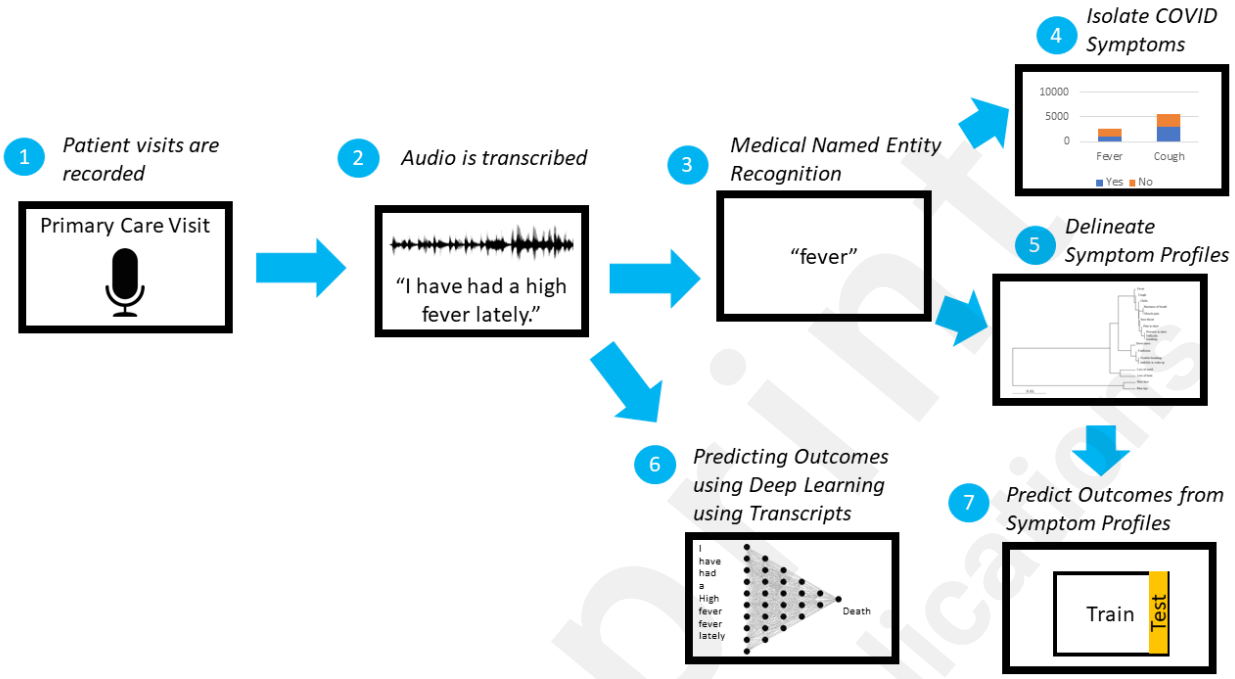

Figure 1. NLP pipeline from audio to phenotype of COVID-19

Implications of the adoption of clinic visit recording in tackling COVID-19

We propose the creation of a pipeline from the audio recording of clinic visits to the clinical symptomatology model and prediction of COVID-19 infection (See Figure 1). With vast amounts of data available, we believe a prediction model can be quickly developed that could promote the accurate screening of individuals at risk of COVID-19. Beyond the challenge of creating a clinical phenotype, an unfiltered and raw account of a patient's clinical experience of the disease allows us to answer other pressing questions, such as understanding the constellation of patient characteristics that may predict greater risk of a more severe infection. If clinical encounters are recorded and our NLP is adequately refined, then benchtop-virology will be better informed. Recordings of clinic visits also offer a historical artifact to learn from, so that we may be better prepared for the next

pandemic.

With the mass transition to telehealth visits, and guidance for conducting remote assessments of COVID-19 via telehealth in primary care are now available,[12] an opportunity now exists to capture the audio of clinic consultations at scale. An accurate model predicting higher risk of COVID-19 could be applied to telehealth visits with the added benefit of reducing the risk of exposure to clinicians, patients, and the public. The use of NLP to remotely screen for COVID-19 is already emerging, for example, audio recordings of cough sounds are being used to identify individuals with COVID-19.[13,14]

Data from beyond the clinic

While recordings of clinic visits are not the panacea to this pandemic, they are a low cost option with many potential benefits that have only just begun to be explored. Beyond audio-recordings, video recordings of telehealth visits can provide additional diagnostic information such as skin appearance.[12] At home voice-based technologies such as Alexa, Siri and Google Home can also be used, allowing further information from outside of the clinic visit to supplement predictive models.[15] For example, the Mayo Clinic has recently added a skill to Amazon Alexa called “Answers on COVID19”, providing resources on COVID-19 and also offering a series of questions to determine a person's symptoms and whether they should get tested for COVID-19.[16]

In the rush to solutions in our efforts to confront COVID-19, care must be taken to rigorously protect sensitive data, while challenges exist in accessing the corpus of patient recordings needed to create these models.[11] Using a data collection method should only be done entirely on an opt-in voluntary framework to preserve privacy and confidentiality, but this has the potential to collect data on COVID-19 symptom deterioration at a scale unavailable from all traditional means. This, as is often

the case, points towards an evolving learning health system capable of managing computable knowledge.

References

1. Yang J, Zheng Y, Gou X, et al. Prevalence of comorbidities in the novel Wuhan coronavirus (COVID-19) infection: a systematic review and meta-analysis. *Int J Infect Dis*. 2020. doi:10.1016/j.ijid.2020.03.017
2. McIntosh K, Hirsch M, Bloom A. Coronavirus disease 2019 (COVID-19). UpToDate. https://www.uptodate.com/contents/coronavirus-disease-2019-covid-19-clinical-features. Published 2020.
3. Liao KP, Cai T, Savova GK, et al. Development of phenotype algorithms using electronic medical records and incorporating natural language processing. *BMJ*. 2015. doi:10.1136/bmj.h1885
4. Richesson RL, Hammond WE, Nahm M, et al. Electronic health records based phenotyping in next-generation clinical trials: A perspective from the NIH health care systems collaboratory. *J Am Med Informatics Assoc*. 2013. doi:10.1136/amiajnl-2013-001926
5. Obeid J, Davis M, Turner M, Meystre S, Heider P, Lenert L. An AI approach to COVID-19 infection risk assessment in virtual visits: a case report. *JAMIA*. 2020;https://doi.org/10.1093/jamia/ocaa105.
6. Weiner SJ, Wang S, Kelly B, Sharma G, Schwartz A. How accurate is the medical record? A comparison of the physician's note with a concealed audio recording in unannounced standardized patient encounters. *J Am Med Informatics Assoc*. 2020;27(5):770-775. doi:10.1093/jamia/ocaa027
7. Yuen T, Derenge D, Kalman N. Cognitive bias: Its influence on clinical diagnosis. *J Fam Pract*. 2018;67(6):366-372.
8. Crabb BT, Lyons A, Bale M, et al. Comparison of International Classification of Diseases and Related Health Problems, Tenth Revision Codes With Electronic Medical Records Among Patients With Symptoms of Coronavirus Disease 2019. *JAMA Netw open*.

2020;3(8):e2017703. doi:10.1001/jamanetworkopen.2020.17703

9. Elwyn G, Barr PJ, Piper S. Digital clinical encounters. *BMJ*. 2018;361:k2061. doi:10.1136/BMJ.K2061

10. Barr PJ, Bonasia K, Verma K, et al. Audio-/Videorecording Clinic Visits for Patient's Personal Use in the United States: Cross-Sectional Survey. *J Med Internet Res*. 2018;20(9):e11308. doi:10.2196/11308

11. Quiroz JC, Laranjo L, Kocaballi AB, Berkovsky S, Rezazadegan D, Coiera E. Challenges of developing a digital scribe to reduce clinical documentation burden. *npj Digit Med*. 2019;2(1):1-6. doi:10.1038/s41746-019-0190-1

12. Greenhalgh T, Koh GCH, Car J. Covid-19: A remote assessment in primary care. *BMJ*. 2020;368(m1182). doi:10.1136/bmj.m1182

13. Tobias M. AI And Medical Diagnostics: Can A Smartphone App Detect Covid-19 From Speech Or A Cough? Forbes.

14. Imran A, Posokhova I, Quershi H, et al. AI4COVID-19: AI Enabled Preliminary Diagnosis for COVID-19 from Cough Samples via an App. *arXiv:200401275v5*. 2020.

15. Dojchinovski D, Ilievski A, Gusev M. "Interactive home healthcare system with integrated voice assistant," *2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, Opatija, Croatia, 2019, pp. 284-288, doi: 10.23919/MIPRO.2019.8756983.

16. Furst J. Mayo Clinic introduces skill for Amazon's Alexa about COVID-19. News Network, Mayo Clinic.

Supplementary Files

Figures

NLP pipeline from audio to phenotype of COVID-19.