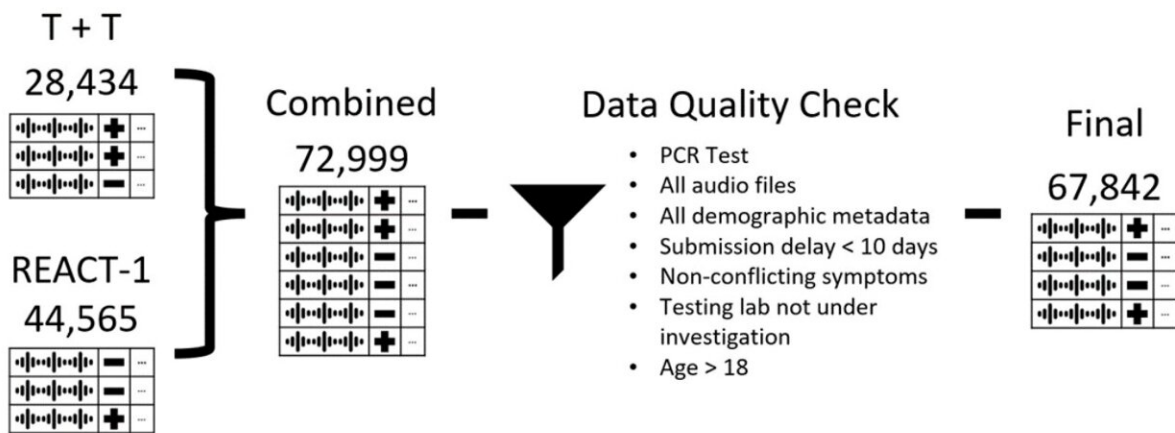


Audio-based AI unreliable for predicting COVID-19 infection, study finds

March 2 2023



(a) Schematic detailing the two recruitment sources for the study and the filtration steps applied to yield the final dataset

		Random	Standard	Matched	Long	Long Matched
train	+	20000	9379	2599		
	-	37665	16518	2599		
	Σ	57665	25897	5198		
test	+	3514	3820	907	10315	2098
	-	6663	7301	907	20509	2098
	Σ	10177	11121	1814	30824	4196

(b) Final dataset splits

Details of final collected dataset in participant numbers. A description of each split can be found in Methods. Credit: *arXiv* (2022). DOI:

10.48550/arxiv.2212.08570

Artificial Intelligence (AI) is unable to accurately predict whether someone has COVID-19 based on the sound of their cough, according to a study involving statisticians from the Department of Mathematics at King's.

By analyzing the accuracy of Machine Learning (ML) algorithms to detect COVID-19 infections, researchers found that this technology did not improve on a model that only used symptoms reported by individuals themselves along with their [demographic data](#), such as age and gender.

Commissioned by the UK Health Security Agency, as part of the government's pandemic response, researchers conducted an independent review of how ML algorithms performed as a COVID-19 screening tool to determine whether AI classifiers could be used as a potential alternative to lateral flow tests, which could be less expensive, less environmentally wasteful and more accurate.

Led by the Alan Turing Institute and Royal Statistical Society, the project involved a team including Professor Steven Gilmour, Dr. Davide Pigoli, Dr. Vasiliki Koutra and Kieran Baker from the Department of Mathematics at King's, together with researchers from the University of Oxford, Imperial College London and University College London.

The team collected and analyzed a dataset of audio recordings from 67,842 individuals recruited in part via the NHS Test-and-Trace program and the REACT-1 study, who had also taken a PCR test. Participants were asked to record themselves coughing, breathing and talking. Results of their PCR tests revealed over 23,000 of them had tested positive for COVID-19.

Researchers then trained a ML model on these audio recordings, comparing them with people's COVID-19 test results. At first, in an unadjusted analysis of the data, the AI classifiers appeared to predict COVID-19 infection with high accuracy—consistent with the findings of previous studies, including research by Massachusetts Institute of Technology, which reported up to 98.5 percent accuracy from AI classifiers predicting whether someone had COVID-19 based on audio recordings.

However with continued analysis the results from this latest study revealed something different.

"When we grouped participants of the same ages, genders and symptoms into pairs, with only one of them having COVID-19, and evaluated these models on the matched data, the AI models failed to perform well in terms of accuracy," said Kieran Baker, Ph.D. student at King's College London and research assistant at the Alan Turing Institute.

"It appeared that the accuracy was due to an effect in statistics called confounding. Almost all people in our dataset with COVID-19 have some symptoms, and so the model is learning that if you have symptoms in the audio, this is a proxy for COVID-19 infection, and no respiratory symptoms means COVID-free. It thus overdiagnosed the number of COVID-19 cases. This confounding is caused by a phenomenon called recruitment bias, because Test and Trace only recruited people with symptoms—so the sample wasn't representative of the population at large," said Baker

These findings provide evidence that audio-based AI classifiers in practical settings were not able to improve over simple predictive scores based on users reporting their symptoms.

While the study did not return a positive outcome in terms of offering a

novel solution to screening and diagnosing diseases such as COVID-19, the researchers were able to introduce new methods for characterizing complex, high-dimensional bias, and make best-practice recommendations for handling recruitment bias. The findings also offered fresh insight for assessing audio-based classifiers by their utility in relevant practical settings.

Professor Stephen Gilmour, Head of the Department of Mathematics said, "This research is timely in highlighting the need for caution when it comes to constructing machine learning evaluation procedures, aimed at yielding representative performance metrics. The important lessons from this case study on the effects of confounding extend across many applications in AI—where biases are often hard to spot and difficult to control for."

Baker said, "It is still possible that technologies that use AI audio classifiers like this could work in the future. Recent publications show progress in detecting sleep apnea and [chronic obstructive pulmonary disease](#) (COPD) from [audio recordings](#)."

"But in order to be certain that this is really working as we hope, it is vital that these models, alongside the data, go through robust model development and testing."

The paper is published on the *arXiv* preprint server.

More information: Harry Coppock et al, Audio-based AI classifiers show no evidence of improved COVID-19 screening over simple symptoms checkers, *arXiv* (2022). [DOI: 10.48550/arxiv.2212.08570](https://doi.org/10.48550/arxiv.2212.08570)

Provided by King's College London

Citation: Audio-based AI unreliable for predicting COVID-19 infection, study finds (2023, March 2) retrieved 6 July 2024 from <https://medicalxpress.com/news/2023-03-audio-based-ai-unreliable-covid-infection.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.