Vocal Biomarkers for Respiratory Symptom Risk Detection and Monitoring: COVID-19 and Beyond

The COVID-19 pandemic has highlighted the critical need for real-time monitoring of potential disease transmission and symptoms at population scales – a challenge that has not been met, despite enormous investments in test strategies and infrastructure around the world. Software-only solutions compatible with existing smartphones and mobile networks offer potential to rapidly reach population scale with relatively little implementation cost. But these solutions have much broader potential: frequent, accessible, non-invasive "health checks" could help people monitor their health and wellness and allow early detection of changes that can lead to pro-active management. Sonde Health has built a vocal biomarker platform that runs on smartphones and uses voice analysis to probe health in areas such as respiratory, brain, and cardiovascular. The COVID-19 pandemic offers a case example of how vocal biomarkers, the extraction of health information from voice, can help build more robust health management approaches for both individuals as well as organizations.

Vocal biomarkers are objective acoustic features in voice that correlate with changes in physical or mental health. Using a microphone to record voice and then analyzing the structure of that voice with software-based solutions is similar to obtaining a blood sample and running a lab test to determine the presence or concentration of molecules in the bloodstream. The difference is that vocal biomarker-based information can be processed on a smartphone with a very brief (5-30 sec) voice sample, analyzed in real-time, with results returned immediately on the user's phone. The vocal biomarker field has advanced significantly in recent years, offering the potential for rapid development and deployment of tools for remote monitoring of potential disease onset and symptom progression of COVID-19 on the basis of voice acoustics.

Audio feature extraction plays an important role in analyzing and characterizing audio content. Filter banks, cepstral analysis, and local discriminant bases technique are some generic mathematical approaches to create arrays of discrete representations, or feature sets, that are the raw material for a range of applications in speech and audio processing. For specific application areas ranging from automatic speech recognition to scene recognition, specific representations are often developed with a goal to produce stable representations that are relatively stable for different examples of the same speech sound despite differences in the speaker and environmental characteristics. Sonde has created software that generates custom feature set from an audio recording of speech that combines characteristics of multiple previous generic approaches like those described above, but with a goal of stably representing acoustic changes that occur when the underlying physiology of a speaker is changed due to changes in health.

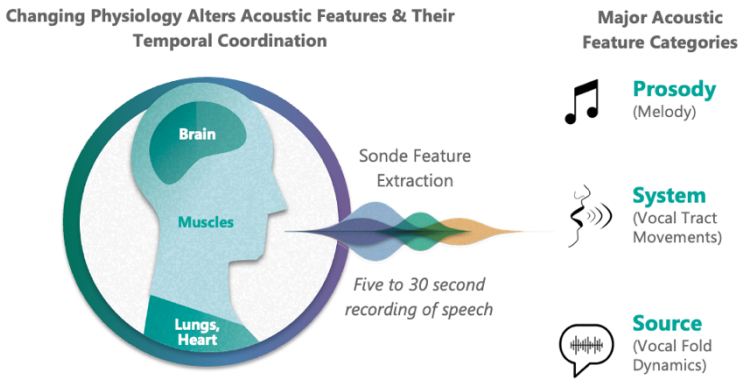

Figure 1. Sonde Feature Extraction and Categories. Sonde's feature extraction software code receives a raw audio recording of varying duration and generates up to thousands of features that represent discrete audio characteristics that capture elements of the rhythm and melody of speech (prosodic features), vocal tract movements (system features), and vocal fold dynamics (source features)

To detect respiratory symptom risk changes that can occur in patients with COVID-19 using Sonde's technology, a user holds an "ahhhh" vowel sound for six seconds while the audio is recorded by a smartphone or tablet. That recording is encrypted and transmitted to a secure cloud server and input into Sonde's feature extraction software module. After that recording is translated into an array of extracted acoustic feature representations, a small subset of these that have been determined to correlate strongly with the presence or absence of respiratory disease diagnoses are then passed as inputs to the trained machine learning model. The process and data used to select these features and train the associated model is summarized now in more detail.

Sonde's COVID-19 monitoring solution was developed using voice and diagnostic data from over 3,000 respiratory disease patients collected by Sonde Health as part of a larger study involving over 50,000 outpatient visits occurring at more than 20 hospitals in India over the past several years. These data include approximately 1,700 individuals with asthma, over 600 with COPD, over 800 with persistent cough and a smaller number of individuals being treated for respiratory infections, tuberculosis, and interstitial lung disease. Sonde's research provided compelling evidence that multiple objective acoustic features extracted from as little as 6 seconds of speech correlate strongly with respiratory disease diagnosis and could help identify individuals with asthma and other respiratory diseases. Although no COVID-19 data was available early on in the pandemic, there were clues that these acoustic features were present in multiple respiratory conditions, suggesting they could also be applied to COVID-19.

Asthma-responsive acoustic feature distributions shown for asthma patients vs. healthy individuals are shown in Figure 2A. Combining these and other similar features allows a prediction of diagnosis status, because their values tend to be shifted up or down relative to those in healthy individuals. Analysis on a subset of asthma patients with Asthma Control Test (ACT) scores segregated by relative level of asthma control also indicates that distributions of these acoustic features correlate objectively with varying levels of respiratory symptom severity and may support development of tools for monitoring symptom progression (Figure 2B). Similar shifts in the distribution along levels of respiratory impairment are observed for patients with primary diagnoses of cough and COPD (Figure 2C).

Several acoustic features were combined into a vocal biomarker, the Respiratory Symptom Risk score, that was tuned to correspond to the likelihood of the presence of a lower respiratory tract condition, with approximately equal sensitivity and specificity. This machine learning model was trained on 80% of the asthma patient data and over 1,000 healthy individuals. Model performance was validated on a hold-out dataset of approximately 600 patients and 600 healthy individuals. The same model was evaluated on over 600 individuals with COPD, more than 800 with persistent cough, and approximately 100 with interstitial lung disease. As Table 1 shows, the model performed consistently across these conditions with approximately 70% sensitivity and 70% specificity. The Odds Ratio, a measure that combines performance on positive and negative predictions, indicates that this Respiratory Symptom Risk score increases the odds of correct identification by 3-9x relative to the pre-test probability of an individual having the disease. Such performance is typical of screening tests, which can be usefully deployed to channel patients to the right type of follow-up consultations or care.

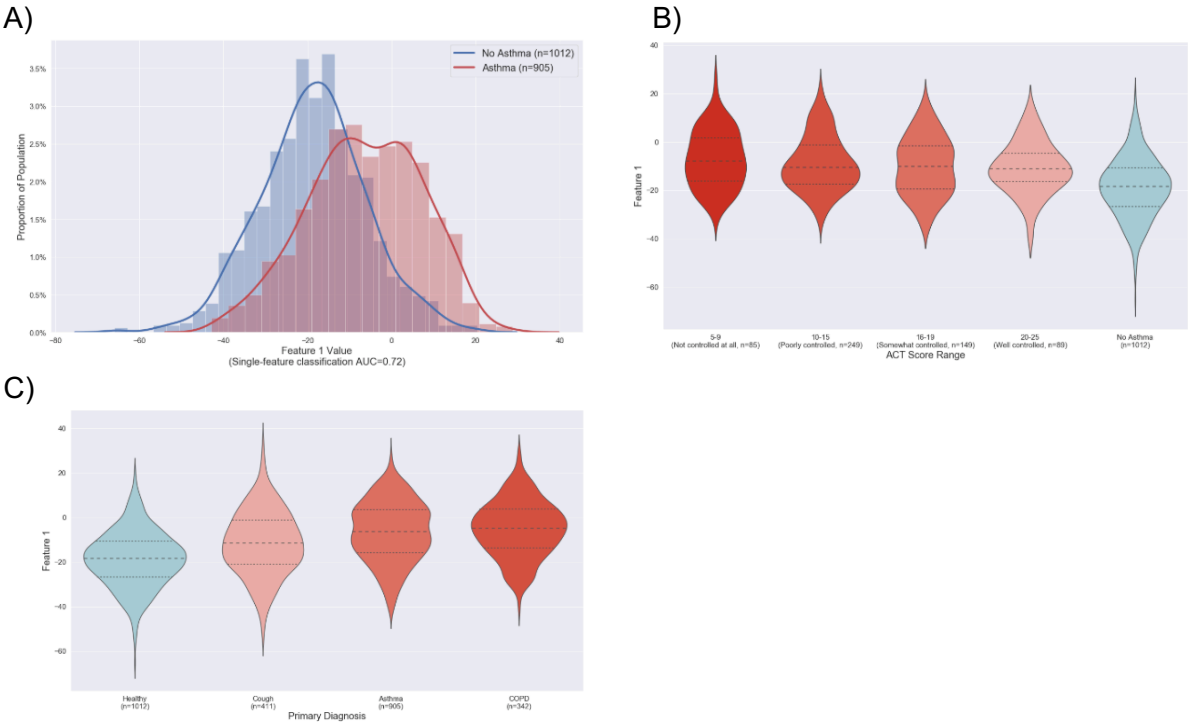

Figure 2. A: The value of a single acoustic feature extracted from 6-second recordings of the utterance "Ahhhh", like in the word "father". Note the significant shift (p<1e-82) in distribution for asthma patients (red, n=905) vs. healthy volunteers (blue, n=1012). B: Asthma patients were divided by the level of asthma control based on their ACT score (available for 572 patients). Violin plots show a shift in distribution toward higher values as relative asthma control levels decrease (p<0.01). C: COPD patients present a similar Feature 1 distribution as asthma patients, whereas patients with a primary diagnosis of cough and no underlying chronic respiratory conditions present an intermediate distribution (p<1e-8), indicating a potential shift as respiratory conditions worsen.

Table 1. Performance measures of the Respiratory Symptom Risk score applied across a range of respiratory conditions show a consistent ability to discriminate between individuals that are healthy vs. those with a respiratory condition. Statistical significance of this performance as measured using conventional statistical approaches indicates $p<10^{-4}$ in all disease types.

Disease	# patients	Sensitivity	Specificity	Odds Ratio
Asthma	576	66%	69%	4.32
COPD	625	77%	73%	9.05
Persistent cough	814	55%	72%	3.14
Interstitial Lung Disease	98	65%	68%	3.95

Based on these promising demonstrations of the Respiratory Symptom Risk score that suggest generalizable quantitative correlations to individuals diagnosed with a range of diseases that share common symptoms including difficulty breathing, tightness in the chest, and persistent cough with COVID-19, we began a series of large-scale deployments of the technology in the US and India to assist with risk-based screening approaches. Data from these deployments are validating the sensitivity and specificity of the vocal biomarker for automated detection of respiratory symptom risk, illustrated in Fig. 3.

The Respiratory Symptom Risk biomarker forms the basis of Sonde One, a return-to-work daily screening app that Sonde launched in August 2020 to help employers keep employees safer as workplaces were gradually re-opening. Combined with a CDC symptom and exposure risk questionnaire and self-reported temperature check, Sonde One provides a <1 minute screening approach that workers can do daily at home before going to work. At one US employer, investigation of ~50,000 screening sessions show respiratory symptom incidence of ~1%, approximately half of which are also identified with the vocal biomarker – rising to approximately 3 out of 4 when multiple symptoms occur together (ROC in Fig. 3).

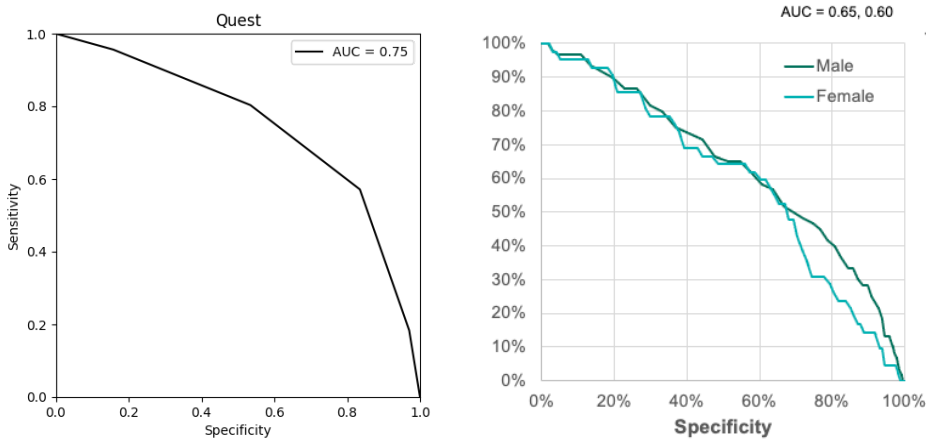

Figure 3. Left panel: Receiver Operating Curve (ROC) of the Respiratory Symptom Risk biomarker illustrating identification of asthma diagnosis vs. healthy individuals, as described above and illustrated in Fig. 1 and Table 1. Right panel: ROC from a ~50,000 voice sample dataset in a single US employer deployment showing the correspondence of the Respiratory Symptom Risk vocal biomarker with ~100 self-reported symptoms of shortness of breath, cough, and sore throat.

While not all of these cases would be due to COVID-19 infection, this does provide a simple and standardized way to avoid taking unnecessary risks by keeping these employees out of the workplace until their health status can be confirmed. Organizations using Sonde One also commented positively on the fact that this approach can be highly standardized across all company locations and departments, and that the screening activity and outcomes are available in real-time and form an auditable record that helps investigations when COVID-19 cases do occur. Employees appreciate the simplicity and find comfort in the fact that everyone present at the workplace has been cleared using a standard set of objective criteria. Fig. 4 illustrates the clear relationship between various symptom types and Respiratory Symptom Risk scores, which allows Sonde One screening to enhance risk determination using this quantitative measure of respiratory health.

Adding objective, quantitative measures of risk significantly enhances the reliability of screening that otherwise would rely purely on subjective self-reports. This is powerfully demonstrated by the discrepancy in the actually measured oral temperature being above 100.4 F, which was reported at a 25x higher rate as compared to self-reported fever status alone (yes/ no response) at this particular employer. Likewise, it should be anticipated that quantitative measures of respiratory function could enhance screening performance relative to the subjective symptoms alone.

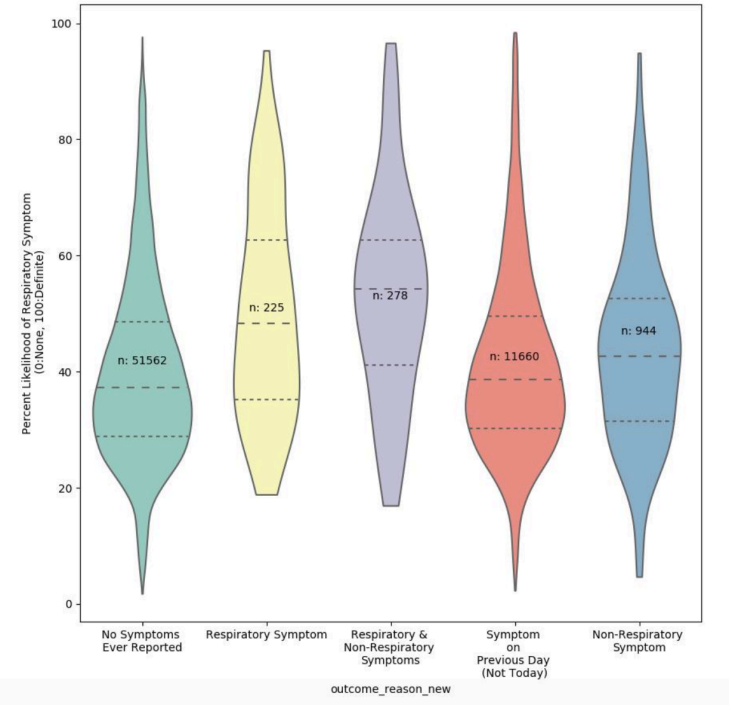

Figure 4. Respiratory Symptom Risk score distributions as measured in the absence or presence of various symptoms from a US employer deployment of Sonde One. The first distribution shows scores for users that have never reported any symptoms, while the second distribution is elevated, consistent with the fact that respiratory symptoms are also reported during these sessions. The third distribution is even more elevated, and represents scores as measured when respiratory symptoms co-occur with non-respiratory symptoms. The larger increases for these scores may reflect a more severe health condition given multiple symptoms are occurring together. The fourth distribution shows the scores return to baseline once respiratory symptoms are no longer reported, and the final and fifth distribution illustrates minimal to no response of the Respiratory Symptom Risk scores to non-respiratory symptoms.

In parallel with the launch of Sonde One, we had already been preparing a multi-site and multi-country clinical trial to validate the Respiratory Symptom Risk scores in COVID-19 patients. At three US hospitals treating COVID-19 patients and one India-based site, starting in September 2020, a total of 400 patients with COVID-19 relevant symptoms are being enrolled in a trial that collects vocal biomarkers, self-reported symptoms, and COVID-19 molecular test results, with participation of 200 healthy volunteers as a comparator group. At the interim analysis point, representing ~3/4 of the full enrollment, vocal biomarker performance is observed to be on par with that seen in asthma, COPD and the other conditions described above. This provides encouraging preliminary evidence from medically confirmed cases that daily screening with Sonde One will help to objectively identify symptoms in COVID-19. Full results will be published at the conclusion of the trial.

The Respiratory Symptom Risk vocal biomarker integrated into Sonde One has shown great potential to aid organizations in keeping workplaces safer, but applications go beyond COVID-19. The strong evidence for generalizability across multiple respiratory conditions suggests that this vocal biomarker, accessible from any smartphone, can function as a respiratory thermometer of sorts. Like the thermometer, an elevated score does not tell the individual *what* may be ailing them, but it allows objective determination of the severity or change in their condition. This smartphone-based rapid assessment of respiratory symptoms can offer a range of remote detection, monitoring and triage data streams in various categories, e.g.:

1. To aid earlier infection diagnosis in the general population, especially in high-risk individuals (e.g. health care workers or those with pre-existing risk factors due to co-morbidities), for COVID-19 but also seasonal influenza and other acute respiratory infections.

2. To monitor symptom progression in individuals with diagnosed chronic respiratory conditions to aid early onset detection for pulmonary exacerbations, e.g., in asthma, COPD, cystic fibrosis, or fluid build-up in congestive heart failure patients.

3. To aid the monitoring of recovering high-risk post-acute patients to understand potential long-term impacts of severe infection and promote earlier detection of potentially serious complications.

The vocal biomarker architecture and mobile applications that enabled the deployment of Sonde One and the results reported herein are compatible with iOS and Android smartphones that are already owned by more than 80% of adult Americans and approximately 50% across emerging economies as a whole. This uniquely enables potential to scale useful objective monitoring capability in high and low resource populations globally with no new capital equipment or reagent expenditures. As the COVID-19 pandemic has accelerated the incorporation of telemedicine and digital health into everyday care, vocal biomarkers will also play an increasingly larger role in helping people monitor their health and wellness. Sonde is excited to continue working with our partners to put the power of voice into everyone’s hands.