Remote monitoring of respiratory function using a cloud-based multimodal dialogue system

Hardik Kothare[1,2] & Vikram Ramanarayanan[1,2]

[1] **Modality.AI, Inc.,** [2] **University of California, San Francisco**

Introduction:

Speech production depends on continuous expiration of air from the lungs. Each respiratory cycle during speech involves an exchange of larger volumes of air in the lungs as compared to quiet breathing[1]. Weakness of respiratory muscles due to neurological conditions like Parkinson's Disease (PD) or Amyotrophic Lateral Sclerosis (ALS) may result in dysarthria; particularly, it may affect the overall loudness of speech[2]. Lung function is thus key to efficient production of speech and is used as an objective measure for disease diagnosis and management by physicians and speech-language pathologists.

The clinical standard for measuring lung function is a spirometry test[3] and it involves the patient exhaling forcefully into a device which measures the flow of exhaled air. With telemedicine gaining traction in recent years, especially during the current COVID-19 pandemic[4], there is an increased need to make clinical tests available to patients at home. Remote spirometry would allow physicians to monitor lung function in patients longitudinally without having to schedule a visit to the clinic. Previous work[5] has demonstrated the feasibility of collecting spirometry data using a microphone. We explore the possibility of using a multimodal dialogue system for remote monitoring of patients' lung function.

Methods:

We developed a method to compute spirometric measures from automated screening interviews conducted via our cloud-based multimodal dialogue system, NEMSI (NEurological and Mental health Screening Instrument)[6]. NEMSI instructs the user to call in using the device of their choice and extracts speech and video metrics in real time. Users are asked to inhale deeply and then blow into the microphone of their device as forcefully as possible. This exhalation is recorded in the form of an audio signal by the microphone. The amplitude envelope in units of pressure (Pascals) of this signal is then extracted and is assumed to be proportional to pressure at lips (Pascals) assuming standard dissipation losses. This pressure value is used to calculate flow rate at lips and an estimate of exhaled volume is calculated by integrating flow with respect to time. Measures like the total exhaled volume or Forced Vital Capacity (FVC), the volume of air exhaled in the first second or Forced Expiratory Volume in 1 second (FEV_1) and Peak Expiratory Flow (PEF) are then extracted and displayed on a user-friendly dashboard accessible by clinicians / researchers.

We are currently working on verifying these measures against the gold standard for clinical spirometry, which requires measurements using a spirometer device. However, we already find promising initial results in this direction. Specifically, the FVC values that we estimate are within ecologically valid limits that are specified in reference charts[7].

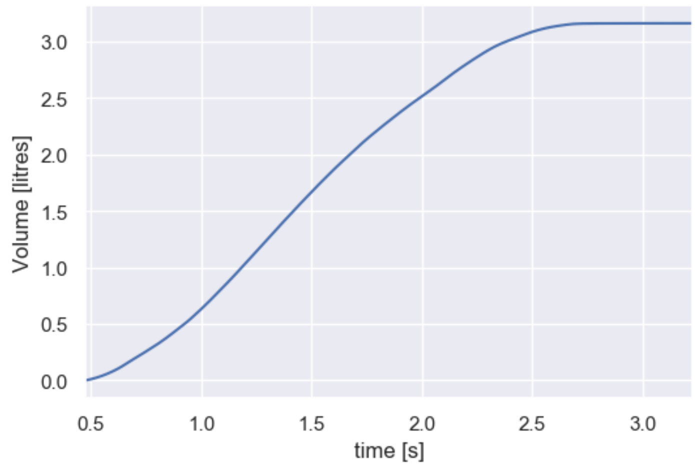

Figure 1: Example Volume vs Time plot obtained using NEMSI where FVC = 3.16 litres and FEV_1 = 1.62 litres. Note that the time axis starts where the audio signal is detected and ends where the audio signal ends.

Discussion and Conclusion:

Our cloud-based multimodal conversational AI system provides a fully-integrated solution to remote diagnosis and monitoring of patient speech metrics, including spirometry. Audio signals can be used to extract clinically-relevant measures. As in clinical spirometry, there will likely be user errors in this method of collecting data, but robust measures can be obtained with user training and intelligent feedback. We aim to validate and verify our methods in collaboration with speech-language pathologists and pulmonologists. In future work, we will collect spirometric data from healthy cohorts as well as patient populations to further confirm the efficacy of our proposed solution. Successful estimation of spirometric data using our remote patient monitoring solution will enable doctors to not only detect and diagnose respiratory ailments but also track lung capacity in patients with respiratory muscle weakness and dysarthric speech. Although this solution is not a replacement for clinical spirometry, it may prove to be a valuable tool in the field of telehealth and telemedicine.

References:

[1] Gick, B., et al. (2012). *Articulatory phonetics*. John Wiley & Sons.
[2] Darley, F. L. et al. (1975). *Motor speech disorders*. Saunders.
[3] Townsend, M. C., & Occupational and Environmental Lung Disorders Committee. (2011). Spirometry in the occupational health setting--2011 update. *Journal of occupational and environmental medicine*, 53(5), 569.
[4] Hollander, J. E., & Carr, B. G. (2020). Virtually perfect? Telemedicine for COVID-19. *New England Journal of Medicine*, *382*(18), 1679-1681.
[5] Larson, E. C., et al. (2012, September). *Proceedings of the 2012 ACM conference on ubiquitous computing* (pp. 280-289).
[6] Suendermann-Oeft, D., et al. (2019, July) *Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents* (pp. 245-247).
[7] Quanjer, P. H., et al. (2012). Multi-ethnic reference values for spirometry for the 3–95-yr age range: the global lung function 2012 equations. *European Respiratory Journal*: 1324-1343