



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*

Multimodal Remote Patient Monitoring Solutions for Neurologic **Conditions**



Clinical importance of remote monitoring of respiratory function

- Weakness of respiratory muscles in neurological conditions like \bigcirc Parkinson's Disease or Amyotrophic Lateral Sclerosis may result in dysarthria¹
- Respiratory function is key to efficient speech production and an objective Ο measure for disease diagnosis and management
- Current clinical standard is a spirometry test²: patients exhale forcefully Ο into a device that measures the flow of exhaled air
- Telemedicine has been gaining traction; current COVID-19 pandemic³ \bigcirc highlights the need to make clinical tests available to patients at home

NEMSI bridges this gap!

- **NEurological and Mental health Screening Instrument** (NEMSI)⁴
 - cloud-based multimodal dialogue system that conducts \bigcirc automated screening interviews by engaging with conversational AI over a device of the user's choice (smartphone, tablet, laptop) from the comfort of their home
 - deployed in an automatically scalable cloud environment Ο allowing it to serve an arbitrary number of end users at a very small cost per interaction
 - natively equipped with real-time speech and video analytics Ο

Unmet need for remote spirometry to allow on-demand remote monitoring \bigcirc

of patients' respiratory function

modules that extract a variety of features of direct relevance to clinicians, thus allowing for measurement of multiple subsystems (motoric, phonatory) in conjunction with lung function

Flowchart for a remote spirometry call

After asking for

details like age,

sex, height and

weight, the AI

agent instructs

users to exhale

forcefully into the

microphone of

their device

Users call a secure web link to start the interview Users are guided to set up their webcam and microphone and greeted by an Al agent (Figure 1)

our call? Modality.ci Patients Calls Sign-out	II? modelity.ai Patients Calls ver it to reveal a definition.	
our call? Patients Calls Sign-out	II? modality.ai Patients Calls Sign-out vikram.ramanarayanan@moda	ession
Patients Calls Sign-out	ver it to reveal a definition.	blems during your call?
Calls Sign-out	ver it to reveal a definition.	
Calls Sign-out	ver it to reveal a definition.	
Sign-out	ver it to reveal a definition.	
	ver it to reveal a definition. vikram.ramanarayanan@modalit	

modality.ai	М	odality:De	emos							
ents	Pa	atients \mathcal{C}								
	A	dd owing rows 1 to 19 of	19							
- OU t ramanaravanan@modalitv.ai	- >	Patient Id ↑↓	Access code ↑↓	Patient type ↑↓	Provide access ↑↓	Session Date (UTC) ↑↓	ALSFRS-R score ↑↓	Speaking rate (words/minute including pauses) ↑↓	Articulation (words/minu excluding pauses) ↑↓	rate ite
	>	AMTWorkers	j462w0	Crowdsourced testers (xx2x)	Re-invite 🖸	11/27/2019 18:44 >	44/44 🗸	211.99 🗸	231.68	4
	>	AndrewC	w9egv3	Modality tester (xx0x)	Invite 🖸	02/08/2020 17:52				
	>	Carlo	4f03au	ePHI (xx3x)	Re-invite 🖸	01/06/2020 20:49	No survey responses (1)	No speech metrics (1)		
	>	Cullen	t3cx06	Modality tester (xx0x)	Re-invite 🗹	12/16/2019 22:46 >	48/48	222.02	297.17	
	>	DavidM	sr1cyz	Internal tester (xx1x)	Re-invite 🖸	01/24/2020 23:50 (48/48	214.39	232.08	
	>	Doug	cc0db8	Modality tester (xx0x)	Re-invite 🖸	11/21/2019 23:33 (Þ)	No survey responses (4)	150.56	170.32	1
	>	DSO	xt034d	Modality tester (xx0x)	Re-invite 🖸	02/03/2020 23:20	No survey responses (4)	No speech metrics (2)		
	>	Harris	o6gjxq	Internal tester (xx1x)	Re-invite 🖸	01/21/2020 23:10 >	48/48	145.63	197.33	

Analytics modules extract metrics automatically

- Amplitude envelope in units of pressure (pascals) extracted from the audio signal after calculating the Hilbert envelope and adding it to the raw signal.
- \circ Pressure at lips or p_{lips} (t) is assumed to be equivalent to pressure at mic (by standardisation of mic-to-mouth distance as 2) inches or ~5cm) and flow rate at lips or $u_{lips}(t)$ is estimated as done in previous studies⁵, $u_{lips}(t) \sim 2\pi r_{lips}^2 / 2p_{lips}(t)$ where r_{lips} is the radius of the lip aperture. Exhaled volume of air is estimated by integrating flow with respect to time.
- Measures like Forced Vital Capacity (FVC) or the total exhaled volume of air and Forced Expiratory Volume in 1 second (FEV1)



	Speech
	Normal speech processes
	Detectable speech disturbance
Call status: Active. Start Conversation Hangup	Intelligible with repeating
When your call is complete, we will press the Hangup button for you and go to the survey.	Speech combined with nonvocal communication
	Loss of useful speech
Figure 1: Example call	Next If an underlined word isn't familiar, move your mouse over it to reveal a definition.

Figure 2: Questionnaires

and FEV1/FVC ratio extracted.

 Measures displayed on a user-friendly dashboard accessible by clinicians and researchers (Figure 3).

Analysis of metrics

Al agent helps

users fill out

relevant

questionnaires

(Figure 2)

- The work presented here is currently still in early stages of development and has only undergone internal testing.
- We attempted to standardise various aspects of data collection like mic type and mic-to-mouth distance.
- However, as in clinical spirometry, poor user adherence to instructions and resulting user errors are a major hurdle.
- Slow exhalation (as opposed to a forced burst) may result in an overestimation of FVC values and underestimation of FEV₁ values (see Figure 4A) whereas incomplete exhalation results in an underestimation of FEV₁ and FVC values (see Figure 4B).
- Not holding the microphone at the recommended distance causes the audio signal to be too quiet resulting in incorrect FEV₁ and FVC values but an accurate FEV_1/FVC ratio (see Figure 4C).



Figure 4: Volume vs Time plots

 In the absence of user errors, all estimated metrics fall within the range of predicted values for the user's age, sex and height (see Figure 4D) as specified in reference charts⁶.

(A) Slow exhalation, (B) Incomplete exhalation, (C) Microphone held far away, (D) No user errors Note: The axes are not uniform across panels

Next steps

- Our cloud-based multimodal dialogue system provides an integrated scalable solution to remote diagnosis and monitoring of respiratory function in patients with respiratory muscle weakness and dysarthric speech.
- However, there are several limitations that need to be overcome before the solution can be implemented.
- We look forward to working with the research community on the following points:
 - 1. Ideas on ensuring user adherence by means of user training and

intelligent feedback

- 2. Developing a reliable method to calibrate microphones
- 3. Collection of preliminary data from patient populations and healthy cohorts to validate and verify our methods.
- Remote spirometry is not a replacement for clinical spirometry but provides a

valuable telehealth and telemedicine tool.

References

¹ Darley, F. L. et al. (1975). Motor speech disorders. Saunders.

² 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.

³ Hollander, J. E., & Carr, B. G. (2020). Virtually perfect? Telemedicine for COVID-19. New England Journal of Medicine, 382(18), 1679-1681.

⁴ Suendermann-Oeft, D., et al. (2019, July) Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents (pp. 245-247).

⁵ Larson, E. C., et al. (2012, September). Proceedings of the 2012 ACM conference on ubiquitous computing (pp. 280-289).

⁶ 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