1

Novel Speech and Voice Biomarkers for the Remote Detection of Cardiovascular Disease

Cardiovascular Grand Rounds
May 19 2025

Jas Sara MBChB PhD

2

Disclosures

- No disclosures

3

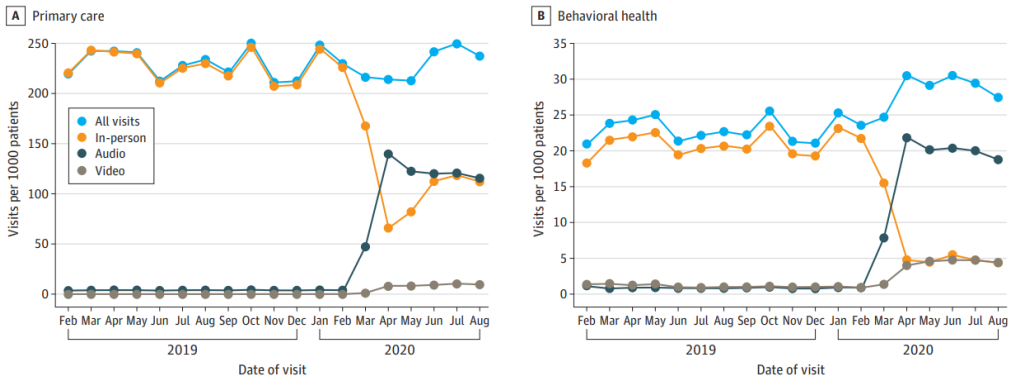
Learning Objectives

- Discuss potential mechanisms that may explain the association between voice biomarkers and disease
- Describe the process of developing a voice biomarker
- Describe evidence for the association between voice biomarkers and various diseases
- Discuss important future considerations in this area of research

4

Background

- Use of telemedicine and digital biomarkers has increased over recent years
- Shift from in-person care to non-face-to-face healthcare visits

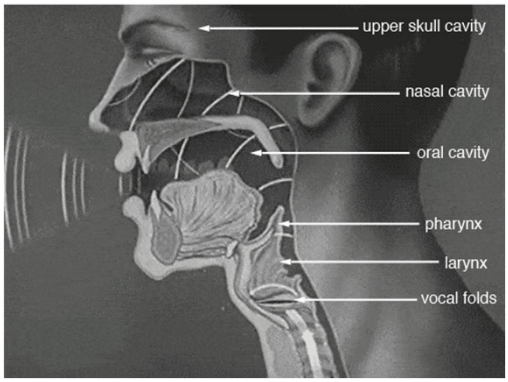


Uscher-Pines, JAMA 2021

5

Voice and Speech Analysis

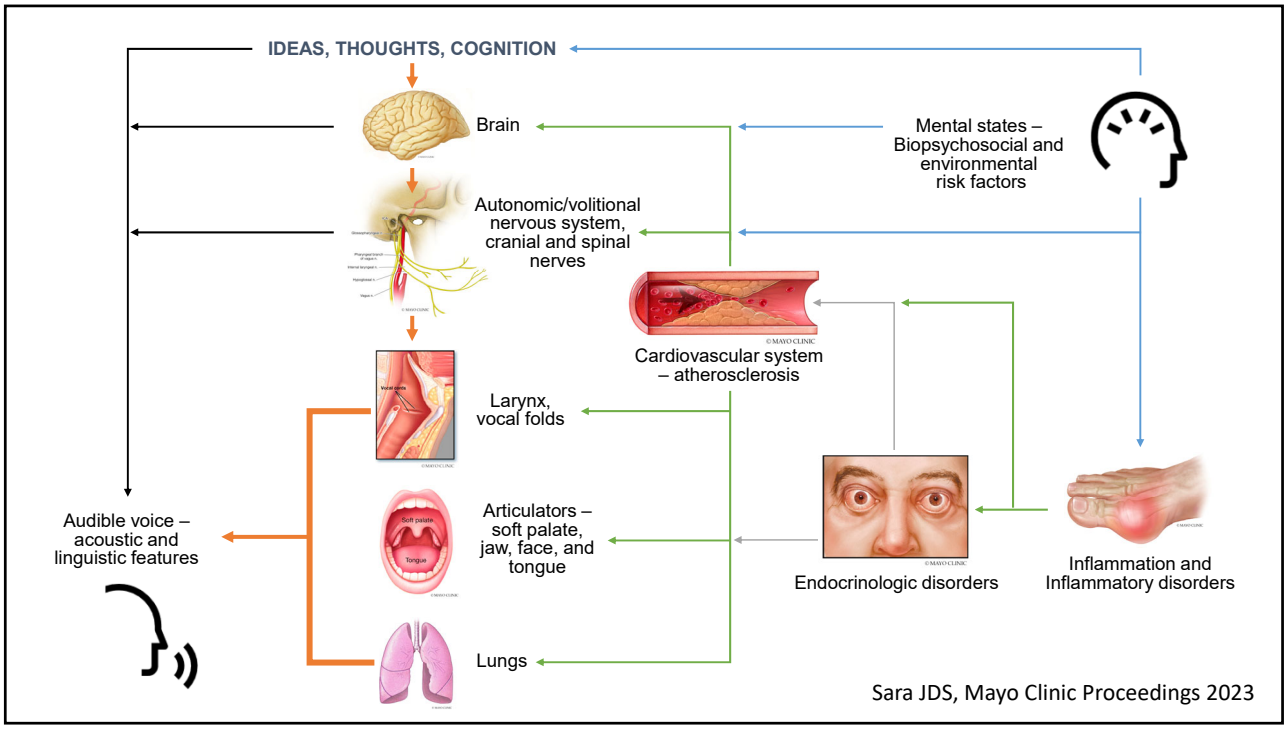
- Emerging non-invasive digital technology
- Coupled with machine-based learning



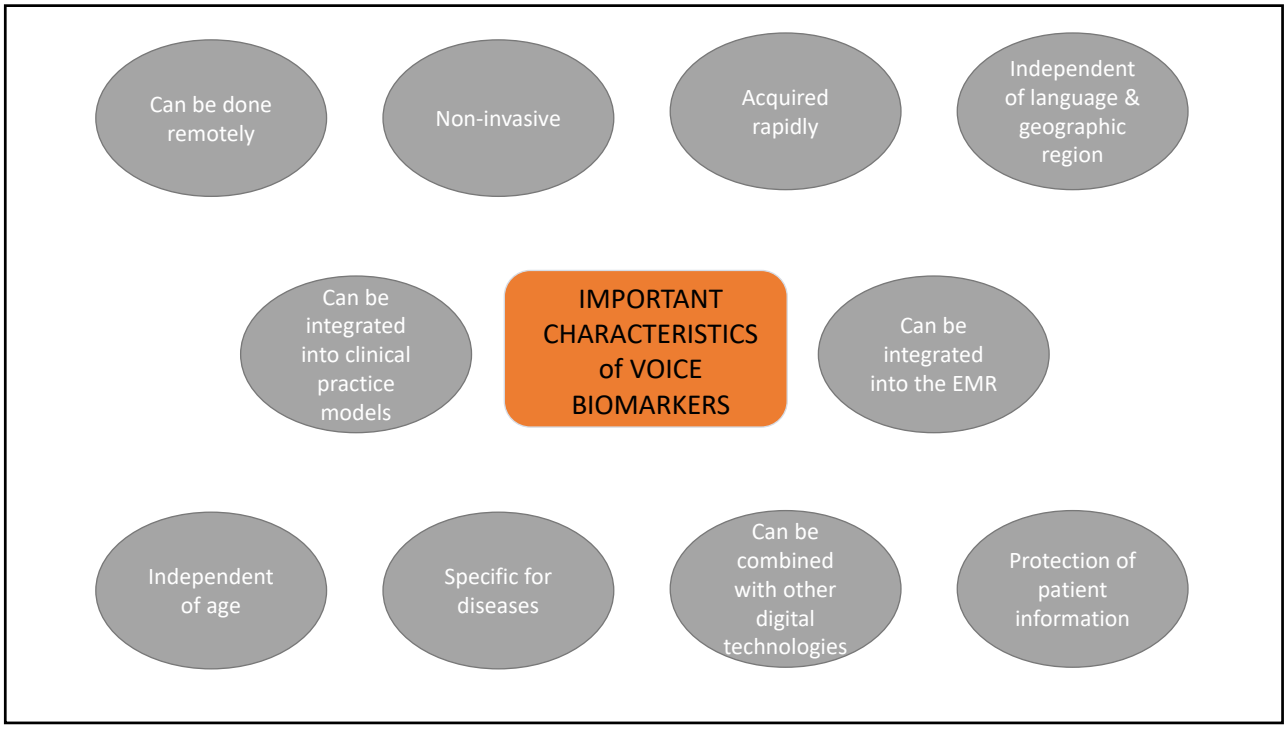
<u>Why Might this Be Useful</u>

- Coordination of cognitive and motor processes
- Simple and engaging for patients
- Relatively inexpensive and convenient
- Remote testing
- Repeated and frequent testing

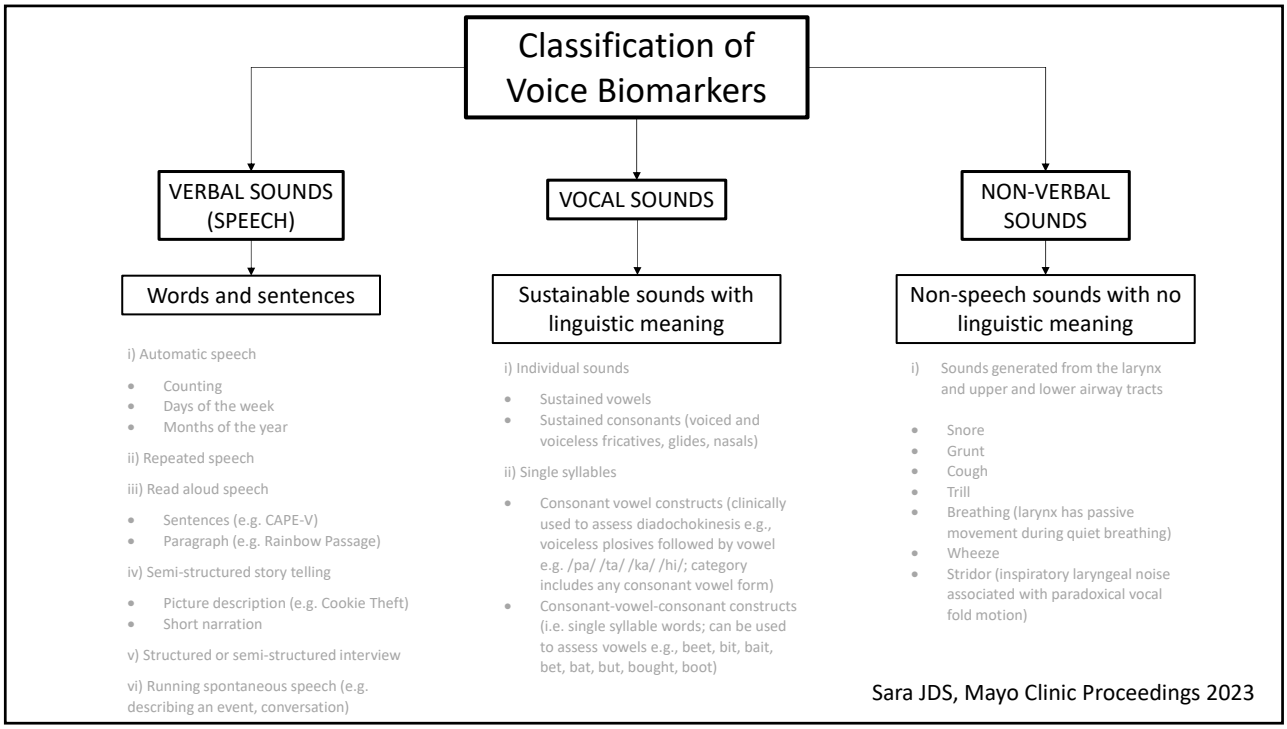
6


IDEAS, THOUGHTS, COGNITION
Brain
Mental states – Biopsychosocial and environmental risk factors
Autonomic/volitional nervous system, cranial and spinal nerves
Cardiovascular system – atherosclerosis
Larynx, vocal folds
Audible voice – acoustic and linguistic features
Articulators – soft palate, jaw, face, and tongue
Endocrinologic disorders
Inflammation and Inflammatory disorders
Lungs
Sara JDS, Mayo Clinic Proceedings 2023

7


Can be done remotely
Non-invasive
Acquired rapidly
Independent of language & geographic region
Can be integrated into clinical practice models
IMPORTANT CHARACTERISTICS of VOICE BIOMARKERS
Can be integrated into the EMR
Independent of age
Specific for diseases
Can be combined with other digital technologies
Protection of patient information

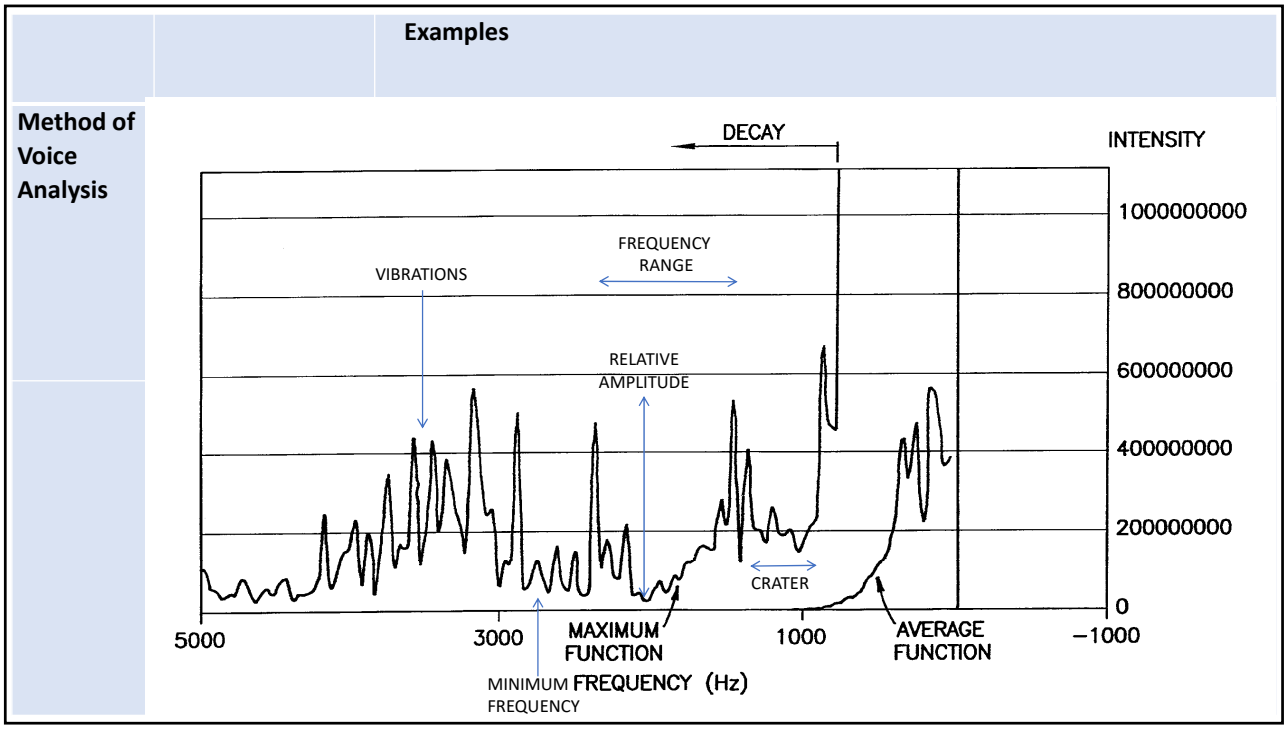
8



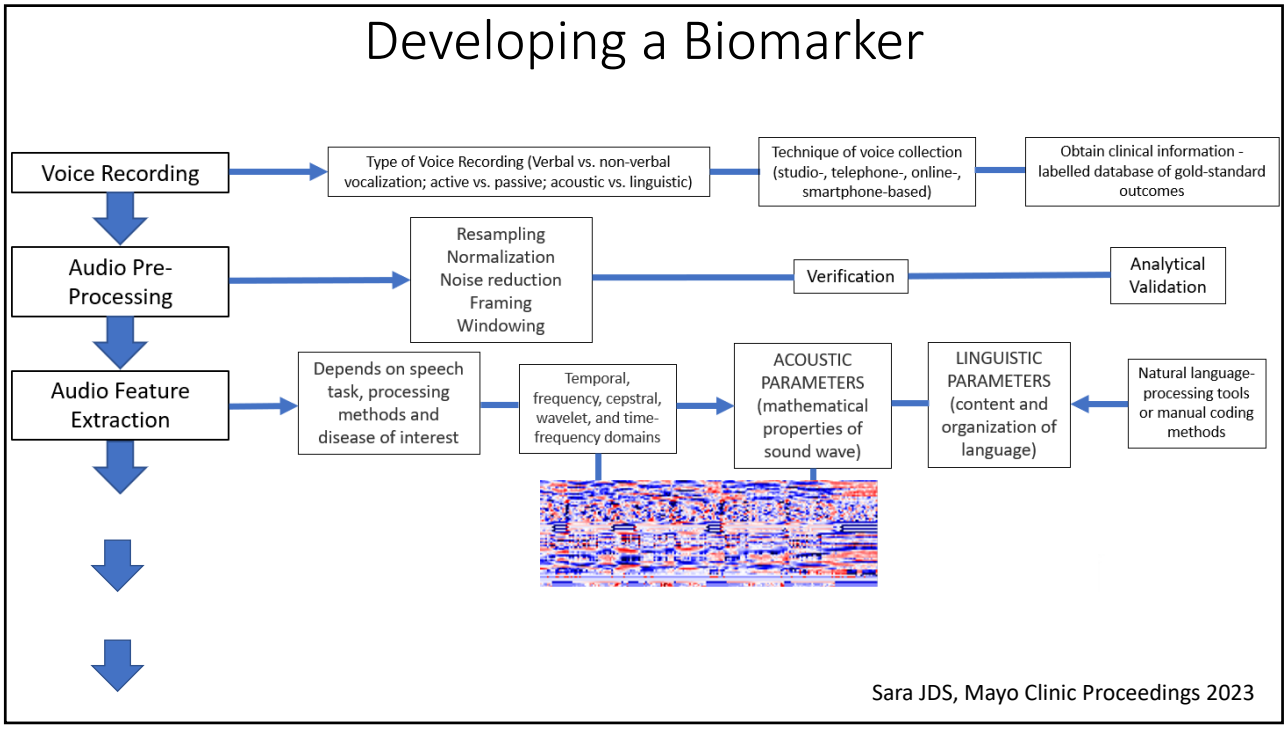
9

I	Source of Vocalization	S – speech V – vocal sounds N – non-verbal sounds D – all types
II	Method of Voice Assessment	A – active recording P – passive recording D – both types
III	Method of Voice Analysis	A – acoustic L – linguistic D – both types
IV	Location of Recording	I – in-person R – remote

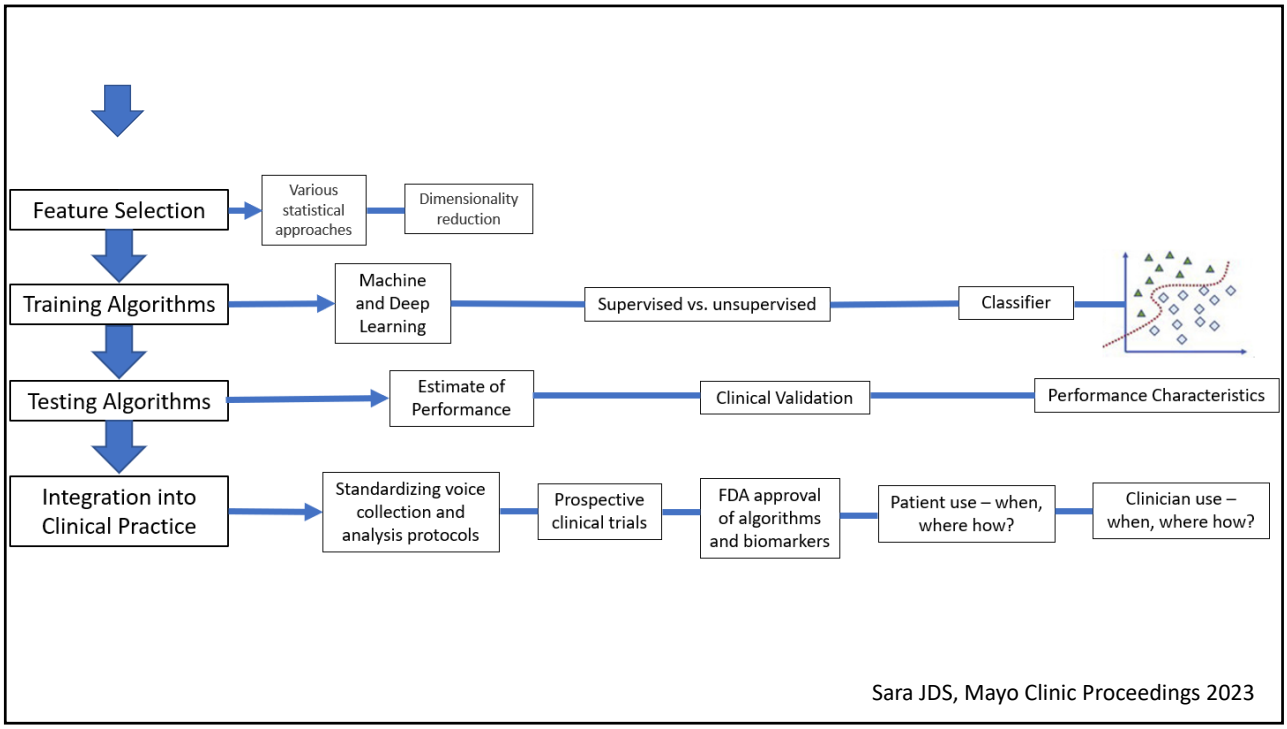
10


Examples
Method of Voice Analysis
DECAY
INTENSITY
1000000000
800000000
600000000
400000000
200000000
0
FREQUENCY RANGE
VIBRATIONS
RELATIVE AMPLITUDE
CRATER
5000
3000
MAXIMUM FUNCTION
1000
AVERAGE FUNCTION
−1000
MINIMUM FREQUENCY
FREQUENCY (Hz)

11


Developing a Biomarker
Voice Recording
Type of Voice Recording (Verbal vs. non-verbal vocalization; active vs. passive; acoustic vs. linguistic)
Technique of voice collection (studio-, telephone-, online-, smartphone-based)
Obtain clinical information - labelled database of gold-standard outcomes
Audio Pre-Processing
Resampling
Normalization
Noise reduction
Framing
Windowing
Verification
Analytical Validation
Audio Feature Extraction
Depends on speech task, processing methods and disease of interest
Temporal, frequency, cepstral, wavelet, and time-frequency domains
ACOUSTIC PARAMETERS (mathematical properties of sound wave)
LINGUISTIC PARAMETERS (content and organization of language)
Natural language-processing tools or manual coding methods
Sara JDS, Mayo Clinic Proceedings 2023

12



13

Examples

Individuals asked to perform 3 separate 30-second voice recordings:

- i) reading a prespecified text
- ii) describing a positive emotional experience
- iii) describing a negative emotional experience

- Spontaneous conversational speech?
- Reading scripted text?

Selective approach:

- Linguistic
  - Depression
  - Alzheimer's Disease
- Acoustic
  - Parkinson's Disease
  - Pneumonia

14

Specific Diseases

Neurologic Diseases

- Parkinson's disease
- Alzheimer's disease
- Mild cognitive impairment

Psychiatric Conditions

- Autistic spectrum disorder
- Post-traumatic stress disorder
- Major depression
- Bipolar disorder

Infectious Diseases

- COVID-19

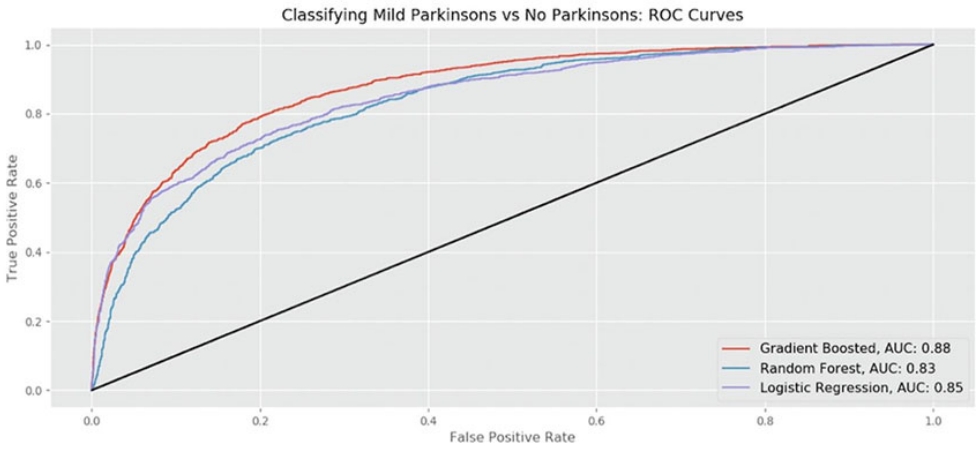
Cardiovascular Diseases

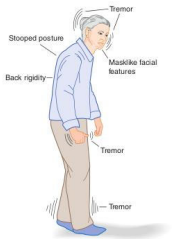
- Coronary artery disease at baseline
- Coronary artery disease events at follow-up
- Pulmonary hypertension
- Heart failure hospitalization and mortality

15

Neurological Disease

Parkinson's Disease



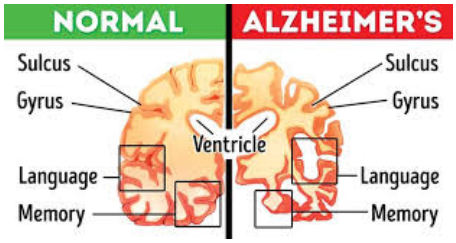


Speech **P**assive **A**coustic **R**emote (**SPAR**) biomarker

Tracy et al. J Biomed Inform 2020

16

Alzheimer's Disease

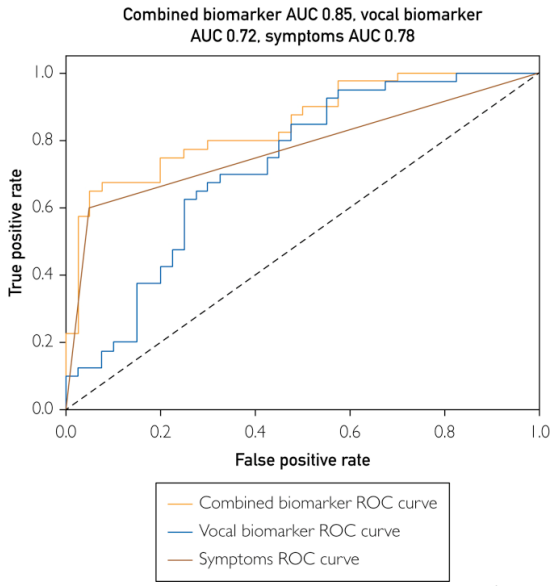


Ref.	Sample	Voice biomarker used	Principal findings
Ahmed et al. 2013	15 subjects	**S**peech **A**ctive **L**inguistic **I**n-Person (**SALI**) biomarker	Subtle changes in language evident in prodromal stages of Alzheimer's disease
Toth et al. 2018	48 with MCI 38 healthy controls	**S**peech **A**ctive **L**inguistic **I**n-Person (**SALI**) biomarker	Significant differences in most features evaluated Discrimination of groups with F1 score of 79%

17

Infectious Diseases

COVID-19



Vocal **A**ctive **A**coustic **R**emote (**VAAR**) biomarker

Maor et al. Mayo Clinic Proc Innov Qual Outcomes 2021

18

Psychiatric Diseases

Post Traumatic Stress Disorder

Ref.	Sample	Voice biomarker used	Principal findings
Marmar et al. 2019	52 cases of PTSD 77 controls	**S**peech **A**ctive **D**ual (acoustic and linguistic) **I**n-person (**SADI**) biomarker	18 speech features collectively provided an AUC of 95.4%

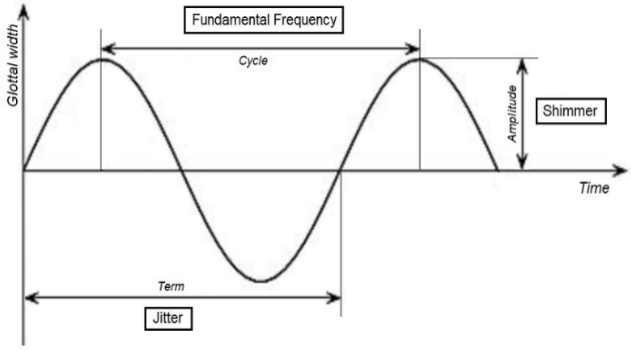
19

Ref.	Sample	Outcome	Voice biomarker used	Principal findings
Belouali et al 2021	588 audio recordings	Suicidal ideation	**S**peech **A**ctive **D**ual **I**n-person (**SADI**) biomarker	Classified suicidal ideation with 86% sensitivity, 70% specificity
Zhang et al 2020	535 audio recordings	Depression severity (using PHQ-9)	**S**peech **A**ctive **D**ual **R**emote (**SADR**) biomarker	Voice features predicted PHQ-9 scores with an AUC of 82.1%
Mundt et al 2012	105 adults with major depression	Depression severity Treatment response	**S**peech **A**ctive **A**coustic **R**emote (**SAAR**) biomarker	Significant differences in speech patterns in responders
Faurholt-Jepsen et al 2016	28 adults with bipolar disorder	Bipolar disorder	**S**peech **A**ctive **D**ual **R**emote (**SADR**) biomarker	Voice features classified manic or mixed states with an AUC of 89%

20

Cardiovascular Disease

- Non-smoking vocally untrained males
  - 6 young and healthy
  - 6 elderly and healthy
  - 6 elderly and diagnosed with “chronic atherosclerosis”
- Asked to enunciate the sound “a” at a comfortable pitch



Results

- Differences between groups in acoustic features
- Differences most evident when comparing healthy young males and elderly males with atherosclerotic disease

Orlikoff et al. Journal of Speech, Language and Hearing Research 1990

21

Ref.	Sample	Age, yrs	Male (%)	Voice biomarker used	Cardiovascular outcome	Principal findings
Maor et al 2018	101 planned coronary angiograms 37 controls	Median 61	54	81 different pre-specified voice features from each recording **S**peech **A**ctive **A**coustic **I**n-person (**SAAI**) biomarker	Coronary artery disease (CAD) at angiography	

Recording	OR	95% CI	P
Recording 1			
Feature 43	1.97	0.81-4.80	.14
Feature 71	0.63	0.32-1.22	.17
Recording 2			
Feature 43	2.79	0.90-8.62	.08
Feature 71	0.37	0.18-0.79	.009
Recording 3			
Feature 43	4.01	1.25-12.84	.02
Feature 71	0.49	0.25-0.96	.04

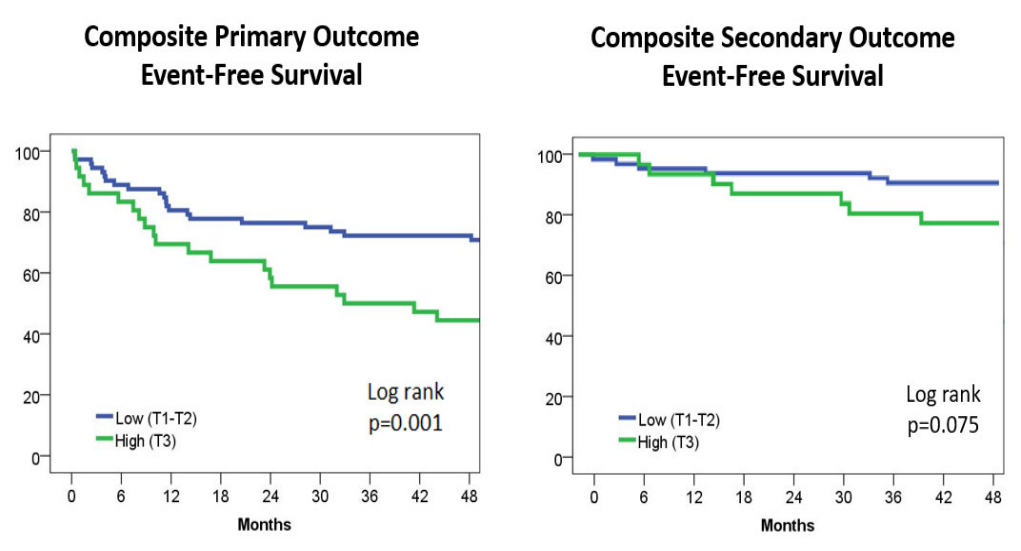
22

Ref.	Sample	Age, yrs	Male (%)	Voice biomarker used	Follow-up	Cardiovascular outcome
Sara et al 2022	108 planned coronary angiograms	Mean 59.5	55	223 acoustic features extracted to develop a voice biomarker **S**peech **P**assive **A**coustic **R**emote (**SPAR**) biomarker	Median 24 months Range 1 – 60 months	**Primary composite outcome:** - presenting to the ED with chest pain - admission to hospital due to chest pain - diagnosis of ACS **Secondary composite outcome:** - incidence of a positive stress test - presence of CAD on coronary angiography

23

Principal findings

- n=43 (39.8%) - primary outcome
- n=15 (13.8%) - secondary outcome



24

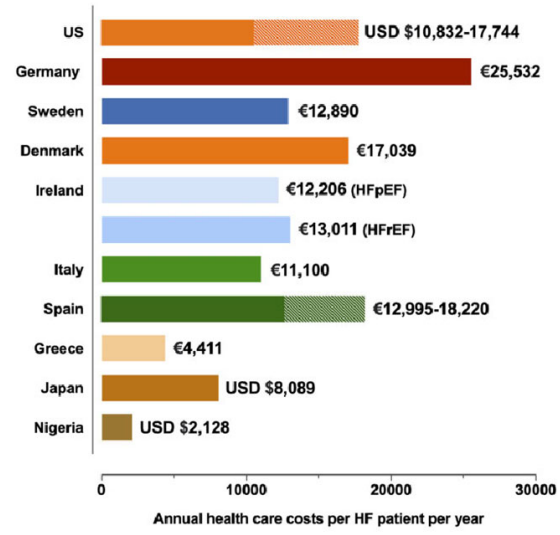
Association Between Voice Biomarker and Composite Outcomes

		Hazards Ratio (95% CI) for the Association with High Mean Voice Biomarker (T3 vs. T1 and T2)	P value
Primary Composite Outcome	Univariable	2.29 (1.26 – 4.17)	0.007*
	Multivariable†	2.61 (1.42 – 4.80)	0.002*
Secondary Composite Outcome	Univariable	2.44 (0.89 – 6.74)	0.080
	Multivariable†	3.13 (1.13 – 8.68)	0.030*

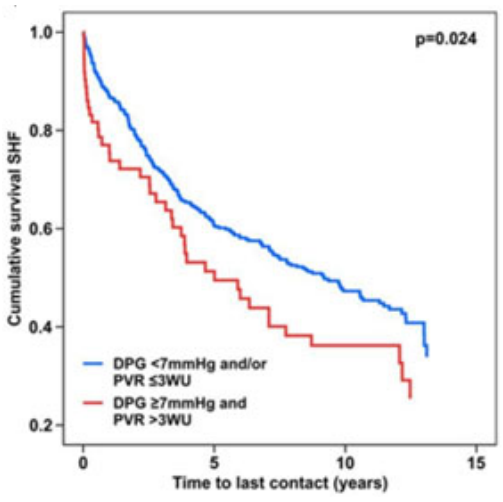
†adjusted for CAD grade at baseline angiogram

25

Heart Failure and Pulmonary Hypertension

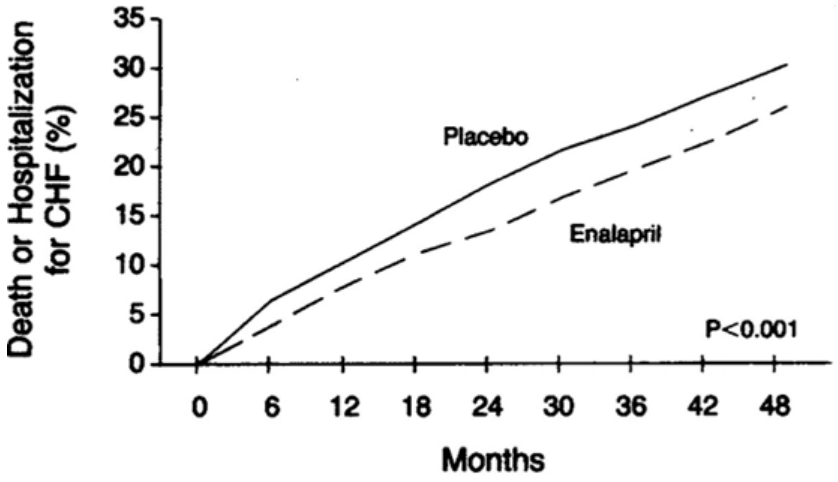


Cardiovascular Research Feb 2022



Circ Heart Fail Research Sep 2017

26

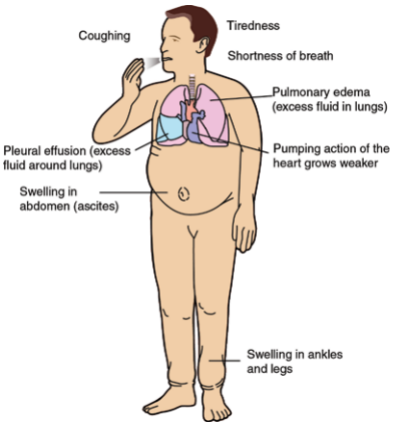


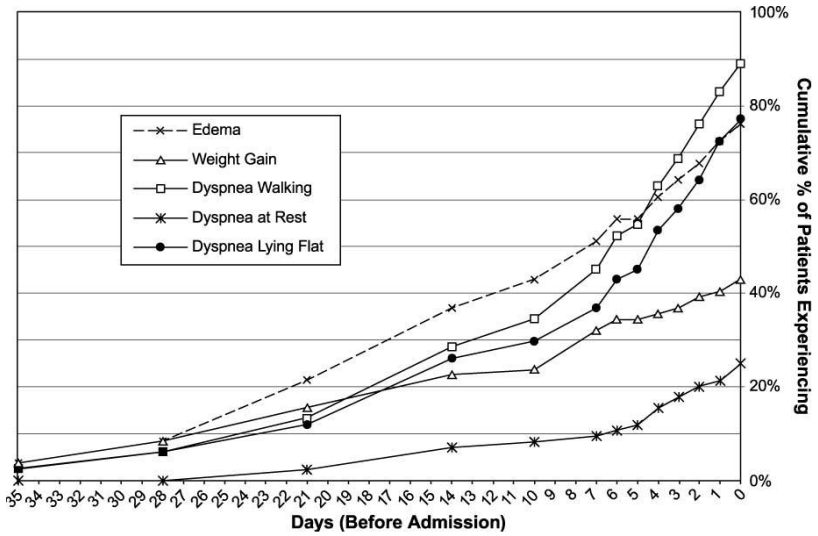
ACE-inhibitors are associated with better outcomes compared to placebo in patients with Stage B (early) heart failure

The SOLVD Investigators NEJM Sep 1992

27

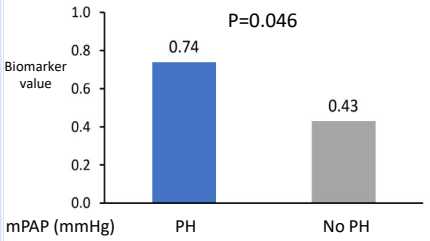
Edema is an early feature of heart failure prior to admission





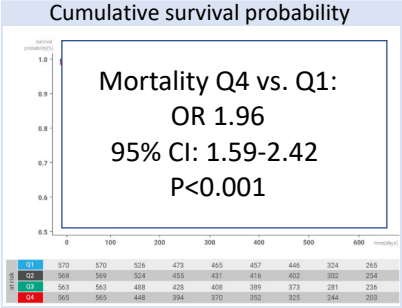
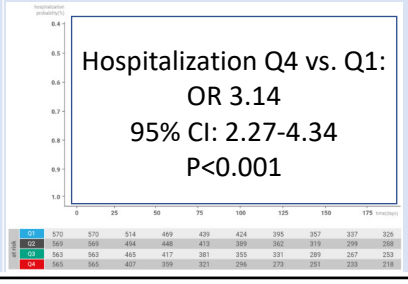
Tex Heart Inst J 2009

28

Ref.	Sample	Age, yrs	Male (%)	Voice biomarker used	Cardiovascular outcome	Principal findings
Sara et al 2020	83 patients undergoing right heart catheterization	Mean 61.6	45	Same biomarker developed for CAD study	**Primary outcome:** pulmonary hypertension (PH)	

MULTIVARIATE ANALYSES		Odds Ratio (95% CI) for Association with Voice Biomarker	P value
Pulmonary Hypertension	All	1.92 (1.00 – 3.65)	0.049*
	PCWP ≥ 15mmHg	1.89 (0.87 – 4.12)	0.008*
	PCWP < 15mmHg	2.09 (0.64 – 6.82)	0.223

29

Ref.	Sample	Age, yrs	Male (%)	Follow-up	Voice biomarker used	Cardiovascular outcome	Principal findings
Maor et al 2020	2,267 patients with clinical heart failure	Median 77	63	Median 20 months	Same biomarker developed for CAD study	**Primary outcome:** all-cause mortality **Secondary outcome:** hospitalization during follow-up	Cumulative survival probability  Cumulative hospitalization probability 

30

Reference	Sample Size	Population	Recording Method	Analytical Technique	Results
Murton et al. JASA 2017	10	Acute HF undergoing diuresis	Daily recordings during admission - Sustained vowels - Speech passages	**Acoustic analysis** - Measures of vocal perturbations	After treatment - Changes in creaky voice - Increased fundamental frequency
Kiran Reddy et al. Computer Speech & Language 2021	45	25 healthy speakers 20patients with acute HF	Single recording - Text-reading task x 3 and spontaneous speech x 1	**Machine Learning** - Glottal features (9 time-domain and 3 frequency-domain) - Mel frequency cepstral coefficients (MFCCs)	- Glottal features classified HF accurately - MFCCs were more accurate than glottal features
Murton et al. Appl Sci (Basel) 2023	52	Acute HF undergoing diuresis	Daily recordings during admission - Speech passages	**Machine Learning** - Several acoustic features	After treatment, speakers exhibited: - Faster speech rates - Longer phrases - More stable phonation

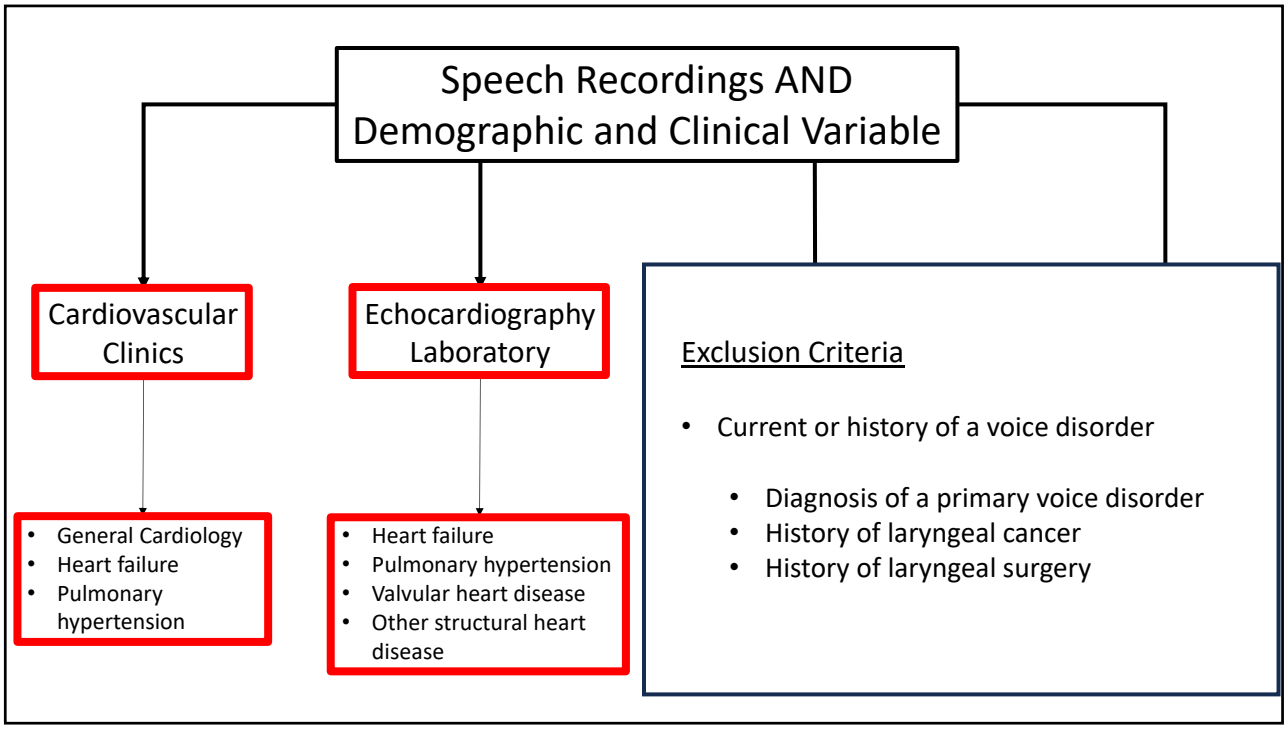
31

Specific Aims

i) Association between <u>acoustic features</u> derived from speech <u>using acoustic analysis</u> and CHRONIC STABLE HF and PH separately and in combination

ii) To train and test <u>deep learning models</u> derived from speech analysis to discriminate CHRONIC STABLE HF and PH separately and in combination

32

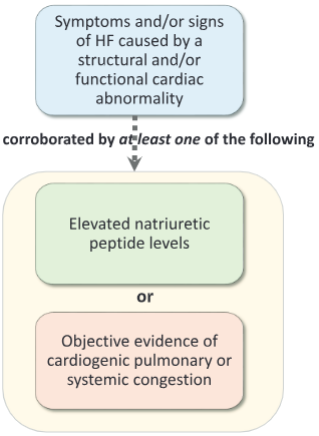


33

Patient Information

- Demographics
- Cardiovascular risk factors and history
- Echocardiography results

HEART FAILURE



Bozkurt et al. Eur J Heart Fail. 2021

PULMONARY HYPERTENSION

	2018 WHO Update
PH	
Mean pulmonary artery pressure	>20 mmHg
Group II PH	
Pulmonary artery wedge pressure	>15 mmHg
Isolated Post-capillary PH	
Diastolic pressure gradient	N/A
Pulmonary vascular resistance	<3 WU

mPAP > 20 mmHg

Maron et al. JAHA 2023

34

Patient Groups

Pulmonary Hypertension – Heart Failure –	Pulmonary Hypertension + Heart Failure –	Pulmonary Hypertension – Heart Failure +	Pulmonary Hypertension + Heart Failure +

35

One morning, Dorothy crossed the hall of the palace and knocked on the door of another girl named Trot. When told to enter, Dorothy found that Trot had company, an old sailor man with a wooden leg who was sitting by the open window puffing smoke from a pipe. The sailor man was named Captain Bill, and he had accompanied Trot to the land of Oz and was her oldest and most faithful comrade. Dorothy liked Captain Bill, and after she had greeted him, she said to Trot, "You know it's Ozma's Birthday next month, and I've been wondering what I can give her as a birthday present. She's so good to us all that we certainly ought to remember her birthday."

The Magic of Oz / Chapter 5 by Frank Baum (1919)

36

Acoustic Analysis – using Praat

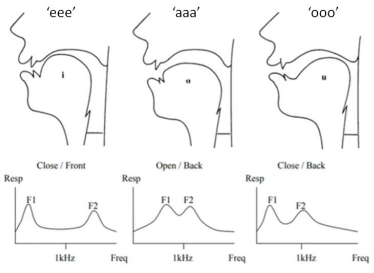
- Number of syllables/duration (sec) – SPEAKING RATE
- Cepstral Peak Prominence – VOICED SEGMENTS (excluding pauses and voiceless consonants)
- Cepstral Peak Prominence – ALL SPEECH (voiced segments, pauses and voiceless consonants)

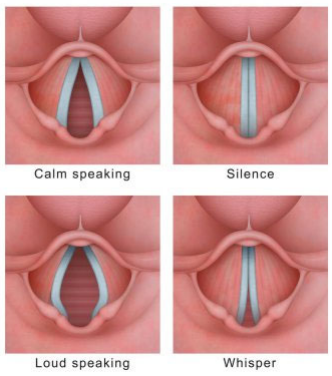
Boersma, P. Praat, 2001

37

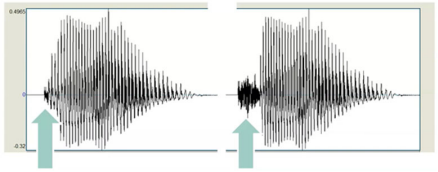
Voiced time

VOWELS





VOICED CONSONANTS



'd' – 10ms aperiodicity

't' – 60ms aperiodicity

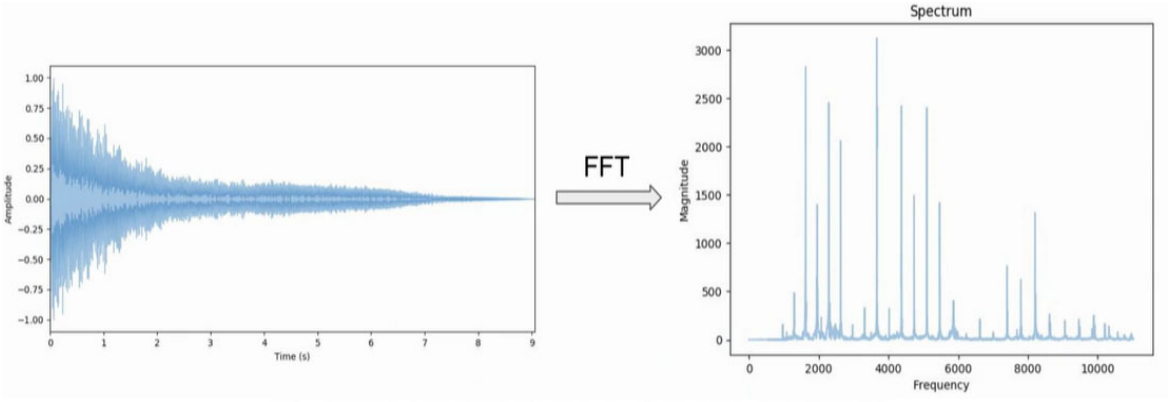
38

Deep Learning Features

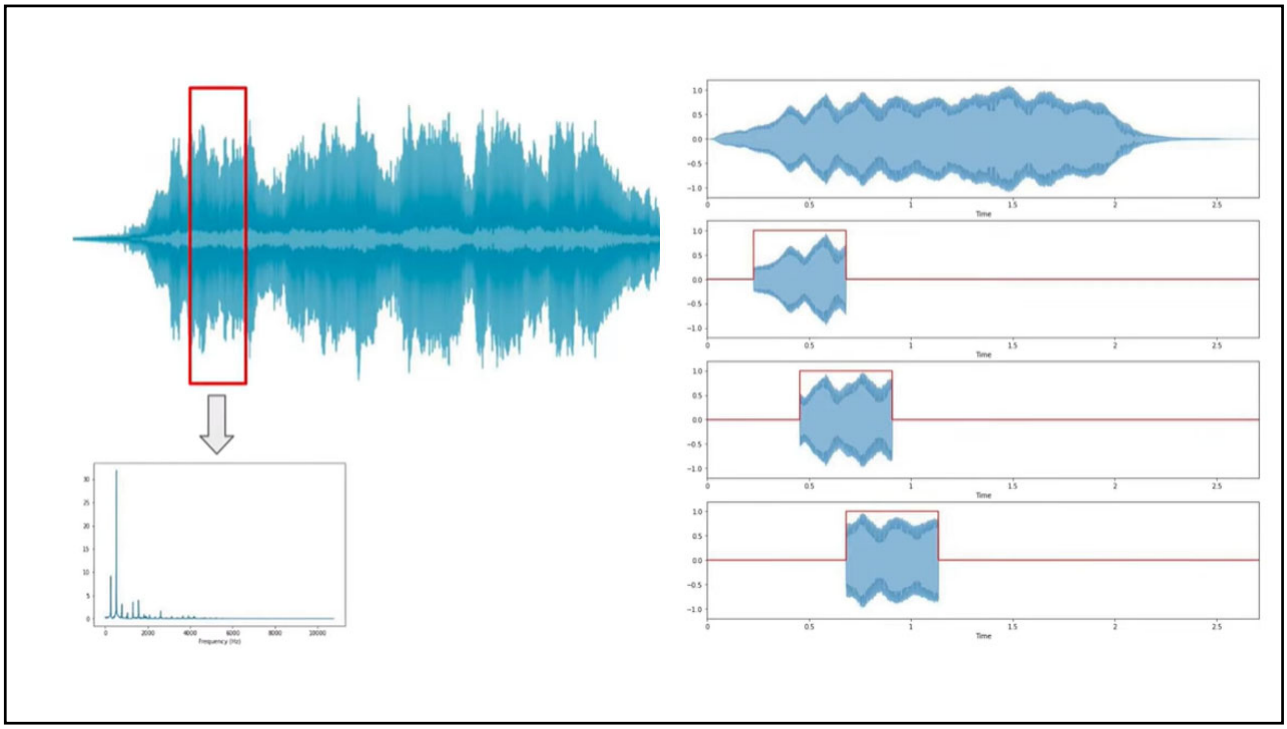
- Mel Spectrograms
- Mel Frequency Cepstral Coefficient (MFCCs)

39

Spectrum

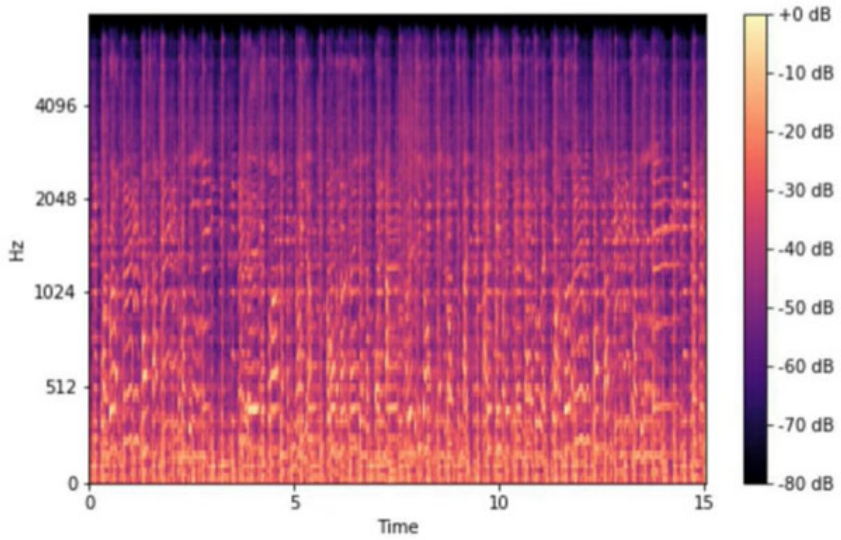


40



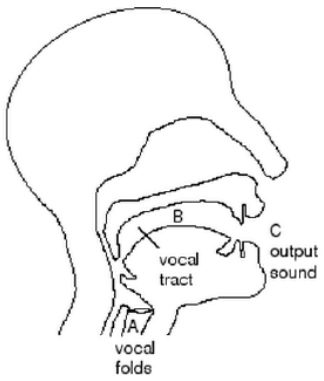
41

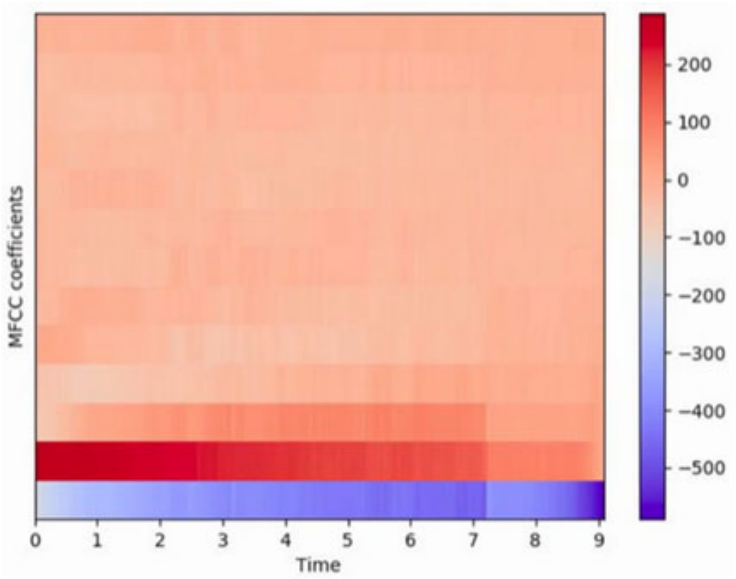
Mel Spectrograms



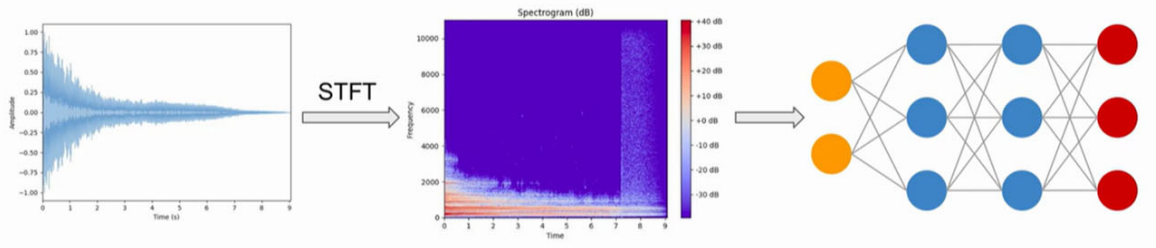
42

MFCCs

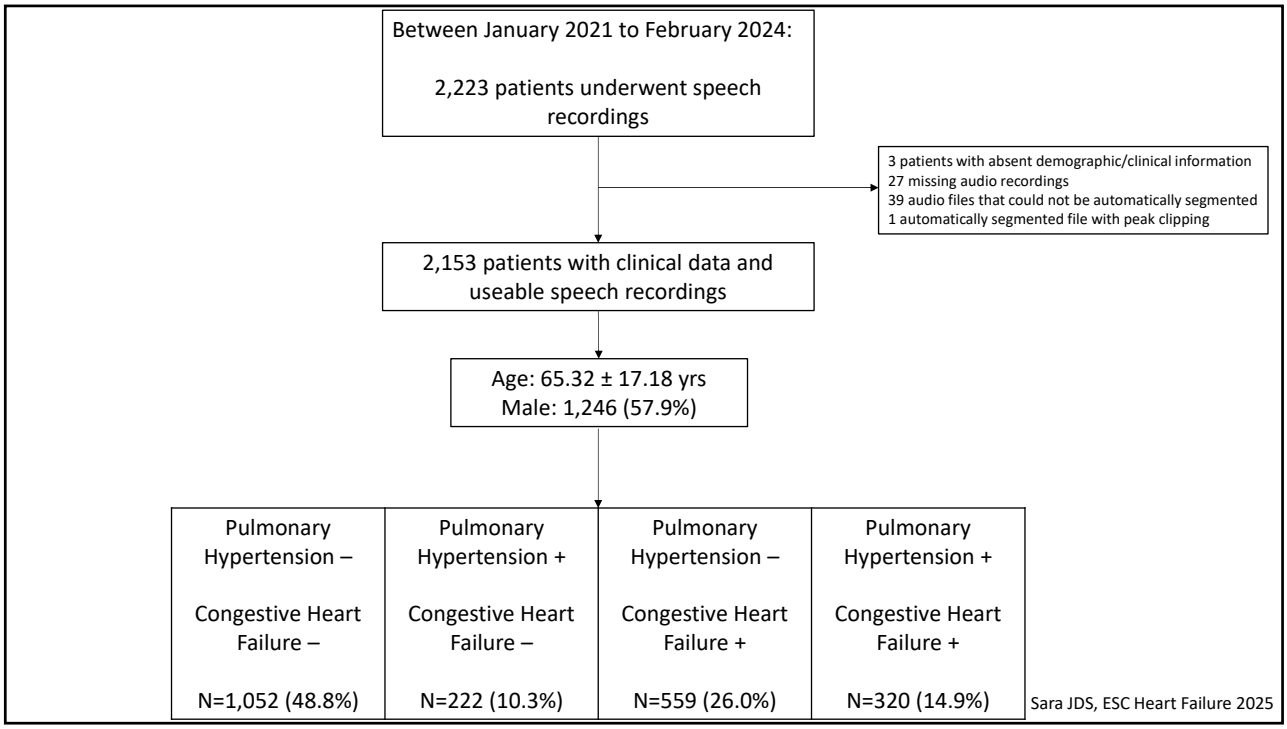

B
C
output
sound
vocal
tract
A
vocal
folds


MFCC coefficients
Time
0
1
2
3
4
5
6
7
8
9
200
100
0
−100
−200
−300
−400
−500

43


STFT
Spectrogram (dB)
Time
Frequency

44



45

	Pulmonary HTN – Heart Failure – (n = 1052)	Pulmonary HTN + Heart Failure – (n = 222)	Pulmonary HTN – Heart Failure + (n = 559)	Pulmonary HTN + Heart Failure + (n = 320)	p-value
Age (years)	60.5 (14.6)	67.9 (13.4)	69.2 (12.8)	71.6 (14.5)	<0.0001*
Female Sex %	455 (43.3)	118 (53.2)	189 (33.8)	145 (45.3)	<0.0001*
Ethnicity %					
White	1009 (95.9)	216 (97.3)	539 (96.4)	307 (95.9)	
Asian	12 (1.1)	1 (0.5)	4 (0.7)	1 (0.3)	
African American	12 (1.1)	1 (0.5)	5 (0.9)	6 (1.9)	0.9460
Other	6 (0.6)	2 (0.9)	4 (0.7)	2 (0.6)	

46

	Pulmonary HTN – Heart Failure – (n = 1052)	Pulmonary HTN + Heart Failure – (n = 222)	Pulmonary HTN – Heart Failure + (n = 559)	Pulmonary HTN + Heart Failure + (n = 320)	p-value
Diabetes mellitus %	151 (14.4)	52 (23.4)	149 (26.7)	80 (25)	<0.0001*
Hypertension %	619 (58.8)	179 (80.6)	452 (80.9)	266 (83.1)	<0.0001*
Hyperlipidemia %	610 (58)	151 (68)	442 (79.1)	238 (74.4)	<0.0001*
History of CAD %	336 (31.9)	92 (41.4)	342 (61.2)	177 (55.3)	<0.0001*
History of Stroke %	66 (6.3)	24 (10.8)	79 (14.1)	41 (12.8)	<0.0001*

47

Acoustic Analyses Results

48

	Univariable Linear Model			Multivariable Linear Model (Adjusted for Age and Sex)		
	Estimated Coefficient	95% CI	p-value	Estimated Coefficient	95% CI	p-value
CPP VOICED SPEECH						
PH+ HF- vs. No disease	0.35	0.14, 0.56	0.0011*	0.37	0.16, 0.57	0.0004*
PH- HF+ vs. No disease	-0.12	-0.27, 0.03	0.1079	0.06	-0.08, 0.21	0.4054
PH+ HF+ vs. No disease	0.01	-0.17, 0.19	0.8802	0.14	-0.04, 0.32	0.1233
CPP ALL SPEECH						
PH+ HF- vs. No disease	0.25	0.09, 0.4	0.0020*	0.23	0.08, 0.38	0.0025*
PH- HF+ vs. No disease	-0.04	-0.15, 0.07	0.4994	0.08	-0.03, 0.19	0.1498
PH+ HF+ vs. No disease	0	-0.14, 0.13	0.9433	0.06	-0.08, 0.19	0.4081
SPEECH RATE (syllables/seconds)						
PH+ HF- vs. No disease	-0.14	-0.22, -0.06	0.0006*	-0.14	-0.21, -0.06	0.0006*
PH- HF+ vs. No disease	-0.18	-0.23, -0.12	<0.0001*	-0.11	-0.17, -0.05	0.0002*
PH+ HF+ vs. No disease	-0.21	-0.28, -0.14	<0.0001*	-0.17	-0.24, -0.1	<0.0001*

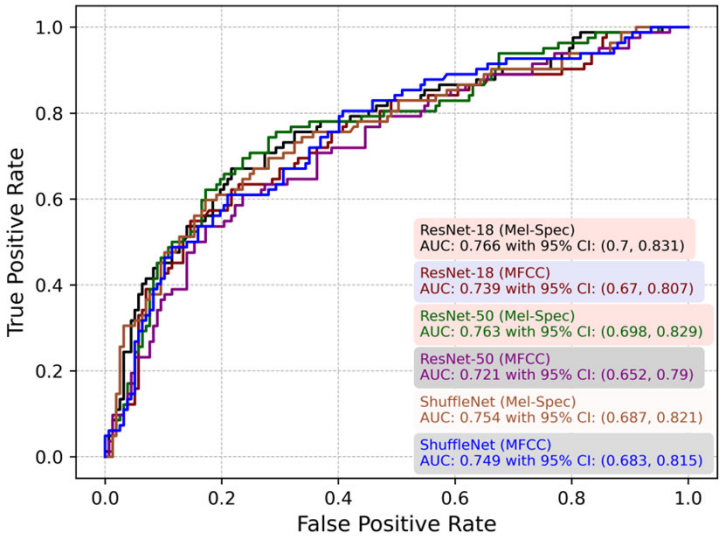
Sara JDS, ESC Heart Failure

49

Deep Learning Results

50

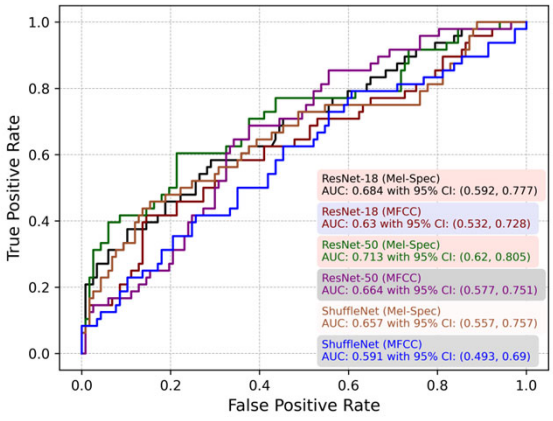
Pulmonary Hypertension vs. No Disease



AUC: 0.72 – 0.77

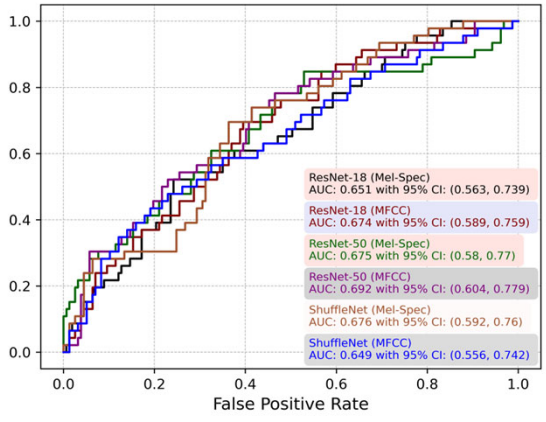
51

Heart Failure vs. No Disease



AUC: 0.59 – 0.68

Pulmonary Hypertension and Heart Failure vs. No Disease



AUC: 0.65 – 0.69

52

Summary

- Cepstral peak prominence differed in pulmonary hypertension
  - Associated with dysphonia/hoarseness
- Speech rate differed in pulmonary hypertension only, heart failure only and both combined
  - Slower speech
- Deep learning models performed moderately well
  - Best classification for pulmonary hypertension

53

Clinical Implications

1) Acoustic analysis of recorded speech to screen for heart failure and pulmonary hypertension
   - non-invasive, convenient and potentially cost-effective
2) Added value to current standard of care?
   - improvements in patient outcomes and healthcare costs?

54

Future Directions

- Standardization in voice data collection and analysis protocols
- Generating large-scale voice sample repositories
- Large prospective clinical trials
- Evaluation of voice biomarkers in diverse populations
- Integration into clinical practice
- Voice biomarkers at home

55

- Agency approval (e.g., FDA)
- Company business models
- Insurance companies payment models
- Caution about false positives and false negatives
- Guidelines outlining best practice for remote care
- Unintended consequences from the ‘digital divide’
- Data protection and privacy

56

Research Innovation Award

Allina Health

Your Voice, Your Wellness

Join Our Study on Burnout in Healthcare

Supported by: Minneapolis Heart Institute Foundation & Allina Health Foundation

Principal Investigator: Emmanouil S. Brilakis, MD, PhD
Director, Center for Complex Coronary Interventions
Chairman, Center for Coronary Artery Disease

57

What's This Study About?

Voice + Burnout

- Short speech samples + quick survey
- Explore voice biomarkers of burnout
- CalmiGo Device: 12-week stress relief trial (if eligible)

Why Join?

- Be part of innovative research
- Contribute to wellness solutions for clinicians
- Try a drug-free stress management tool for free

Dr. Emmanouil Brilakis – emmanouil.brilakis@allina.com
Dr. Michaella Alexandrou – michaella.alexandrou@allina.com

58

Slide 58

JB0 This should be plural (multiple samples and surveys)
Boe, Jennifer L, 2025-04-22T16:05:58.770

JB1 I would also recommend using the summary already approved in the ICF. A good number of participants may not be eligible for part 2 and of those - only half will get to use the device, so it is important that is clear.
Boe, Jennifer L, 2025-04-22T16:07:10.954

JB2 I'd also refrain from using this as a reason to join since most participants will not receive a device.
Boe, Jennifer L, 2025-04-22T16:07:39.349

59