ACOUSTIC VOICE AND SPEECH BIOMARKERS OF TREATMENT STATUS DURING HOSPITALIZATION FOR ACUTE DECOMPENSATED HEART FAILURE

O. Murton[1,2] **G. W. Dec**[2] **R. Hillman**[1,2]
M. Majmudar[3] **J. Steiner**[4] **J. Guttag**[5] **D. Mehta**[1,2*]
[1] MGH Institute of Health Professions, Boston, MA, USA
[2] Massachusetts General Hospital, Boston, MA, USA
[3] Biofourmis, Inc., Boston, MA, USA
[4] Oregon Health & Science University, Portland, OR, USA
[5] Massachusetts Institute of Technology, Cambridge, MA, USA

ABSTRACT

Acute decompensated heart failure (ADHF) refers to a distressing episode in which individuals experience a deterioration in cardiac function due to the failure of physiological mechanisms to adequate compensate for a chronic heart condition. Symptoms of ADHF include systemic fluid volume overload throughout the body, breathlessness, and fatigue. It was hypothesized that systemic congestion would increase lung congestion and vocal fold edema that would, in turn, affect measures of voice and speech function. Data were collected from 52 patients who were hospitalized and treated following an ADHF episode. For each patient, daily inpatient acoustic recordings were obtained during sustained phonation and connected speech as patients recovered to a stable, homeostatic state. Acoustic voice and speech measures were input into a machine learning model to perform supervised binary classification of the recordings as coming from the day of admission or day of discharge. The accuracy of this binary classification task reached 69% using a model that incorporated maximum phonation time, fundamental frequency variability, and average duration of phrases during spontaneous speech. Future work could use voice and speech measures to ultimately serve as early-warning indicators of impending ADHF episodes, thus preventing re-hospitalization.

Keywords: *voice and speech biomarkers, heart failure, voice monitoring*

1. INTRODUCTION

This paper summarizes work that developed a machine learning classifier to detect a particular heart-related condition called acute decompensated heart failure (ADHF) using voice and speech biomarkers [1]. Individuals with heart failure experience impaired cardiac output that reduces adequate blood supply to the body. Many of these individuals develop compensatory physiological mechanisms to remain cardiac stability for extended periods of time; however, episodes of decompensation can lead to ADHF [2], which increases the need for hospitalization and risk of mortality. Preventing ADHF, and associated hospitalization, is therefore a major goal for these patients.
The goal of this project was to determine whether ADHF produces a detectable and measurable impact on voice and/or speech production. The results would lay the foundation for assessing the potential of using non-invasive

*__Corresponding author__: *mehta.daryush@mgh.harvard.edu*

Robert E. Hillman and Daryush D. Mehta have a financial interest in InnoVoyce LLC, a company focused on developing and commercializing technologies for the prevention, diagnosis, and treatment of voice-related disorders. Dr. Hillman's and Mehta's interests were reviewed and are managed by Massachusetts General Hospital and Mass General Brigham in accordance with their conflict-of-interest policies.

measures in an automated system for early detection of heart failure so that medical intervention can occur sooner to prevent re-hospitalization. The concept is that this system would be eventually implemented on a smartphone, either alone or in conjunction with other non-invasive sensors (e.g., heart rate sensors and activity monitors).

2. METHODS

Fifty-two participants (20 female, 32 male) hospitalized for ADHF were enrolled in the study using convenience sampling from two clinical sites. The body weight of each participant was above their "target weight," referring to their usual weight without extra fluid volume. Patients with other non–cardiac diseases or a voice disorder were excluded. Participants were 72 years of age on average, ranging from 34 to 96 years. The average duration of hospitalization was 7.5 days, weight change (admission to discharge) was −5.5 kg, and change in NT-proBNP level (a blood biomarker of heart failure) was −567 pg/mL.

Each day of hospitalization, a speech recording was obtained using an acoustic microphone, including a maximum phonation time task (available for a subset), sustained vowels, reading passages, and spontaneous speech. Acoustic voice and speech features were extracted from the acoustic recordings. Features included summary statistics of speech phrase duration, fundamental frequency, cepstral peak prominence, and creak from the connected speech tasks. The sustained vowel tasks yielded measures of perturbation (jitter, shimmer, harmonics-to-noise ratio) and cepstral peak prominence.

As an initial step, a supervised machine learning paradigm was implemented to classify speech samples labeled as from "day of admission" or "day of discharge." In this supervised learning setup, an L1-regularized logistic regression model was applied using the acoustic features in leave-one-participant-out cross-validation. Each participant thus provided two sets of features (at admission and discharge). Classifier performance metrics included accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve. In addition, effect sizes (Cohen's d) for each measure were also computed.

3. RESULTS

For the reading passages, the total phrase duration resulted in a moderate effect size of −0.5 (speech rate tended to increase from admission to discharge). Maximum phonation time also yielded a moderate effect size of 0.49. Measures of cepstral peak prominence and fundamental frequency resulted in effect sizes that ranged from 0.17 to 0.35. Results of the binary classification task yielded three features with statistically significant odds ratios: maximum phonation time (1.34), fundamental frequency standard deviation (1.25), and median phrase duration in spontaneous speech (1.11). Classification accuracy was 69%, with sensitivity (admission) of 71%, specificity (discharge) of 67%, and AUC of 0.65.

4. CONCLUSION

This study is a step toward discovering acoustic voice and speech biomarkers of ADHF that are clinically interpretable and potentially linked to systemic fluid accumulation in the lungs and larynx. Voice and speech characteristics at admission were associated with shorter maximum phonation time, less stable phonation, slower speech rate, and increased pausing compared to acoustic characteristics at discharge after patients were treated for ADHF. Future work could continue to investigate acoustic biomarkers in larger study samples, combining voice and speech monitoring with other physiological measures, and obtaining information using wearable technologies and/or at-home monitoring systems.

5. ACKNOWLEDGMENTS

This project was supported by the Center for Assessment Technology and Continuous Health (CATCH) at Massachusetts General Hospital; the NIH National Institute on Deafness and Other Communication Disorders Grant T32 DC000038 (awarded to Bertrand Delgutte for trainee Olivia Murton); the NIH National Heart, Lung, and Blood Institute Grant F31 HL143824 (awarded to Olivia Murton); and the Voice Health Institute. The contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health.

6. REFERENCES

[1] O. Murton, G. Dec, R. Hillman, M. Majmudar, J. Steiner, J. Guttag, D. Mehta: "Acoustic voice and speech biomarkers of treatment status during hospitalization for acute decompensated heart failure," *Applied Sciences*, vol. 13, no. 3, article 1827, 2023.

[2] S. Joseph, A. Cedars, G. Ewald, E. Geltman, & D. Mann: "Acute decompensated heart failure: Contemporary medical management," *Texas Heart Institute Journal*, vol. 36, no. 6, pp. 510–520, 2009.