Abstract for the Virtual Symposium on ALS/MND 2021

Multimodal dialog based speech and facial biomarkers capture differential disease progression rates for ALS remote patient monitoring

Michael Neumann, Oliver Roesler, Jackson Liscombe, Hardik Kothare, David Suendermann-Oeft, James D. Berry, Ernest Fraenkel, Raquel Norel, Aria Anvar, Indu Navar, Alexander V. Sherman, Jordan R. Green and Vikram Ramanarayanan

Background:

Continuous monitoring of ALS symptoms is crucial to improve quality of life because both survival and progression rate can vary substantially between individuals. Speech and facial biomarkers can serve as useful proxies for disease progression and can be monitored remotely and automatically.

Objective:

The main objective of this longitudinal study is to analyze acoustic and facial speech metrics extracted from a web-based conversational assessment over time. Our hypothesis is that such multimodal remote patient monitoring allows us to measure and track changes in certain speech and facial biomarkers (i) more frequently and cost-effectively, while (ii) remaining as informative (if not more) than current clinical standard scales in capturing differences in ALS disease progression between slow and fast progressors.

Methods:

Speech and video data was collected from 56 people with ALS (pALS) using a cloud-based multimodal dialog platform between October 2020 and July 2021. In addition to structured speech tasks such as sentence repetition and diadochokinesis (DDK) tasks, participants also filled out the ALSFRS-R, a standard scale for monitoring ALS progression. Based on these ALSFRS-R scores, we stratified participants into two cohorts: (a) slow progressors (<0.47 points/month decline; 17 males, 20 females), and (b) medium to fast progressors (>= 0.47 points/month; 12 males, 7 females). The rate of change was calculated based on the first and last observation for each participant.

Acoustic (timing, frequency, and energy-related) and visual features (movement, surface, and velocity-related) were automatically extracted from the recordings. Their rate of change was computed as linear regression slopes over time. To normalize for sex-specific differences in metrics (such as fundamental frequency), we z-scored all metrics by sex

group. Statistical tests were conducted to identify features for which these slopes are significantly different between the two cohorts.

Results:

Acoustic speech features related to timing (percent pause time), frequency (fundamental frequency), and voice quality showed statistically significant differences in their rate of change between slow and fast progressors at a significance level of 0.05. Effect sizes (Cohen's d) were moderate for most features. Among facial kinematic features, we observed significant differences for the change over time in higher order statistics of the jaw and lips (such as acceleration of the jaw center). Effect sizes were moderate. Additionally, we show that splitting up the analysis by different speech task types (such as read speech, DDK, or vowel phonation) yields a higher responsiveness of certain metrics as measured by effect size.

Conclusions:

Our findings demonstrate the efficacy of remote patient monitoring via a cost-effective and scalable dialog platform to extract informative speech and facial biomarkers, allowing the potential capture of information that standard scales like the ALSFRS-R (which is not very granular) might not register.