Objective: identify audiovisual speech markers that are responsive to clinical progression of amyotrophic lateral sclerosis (ALS) and their reliability in an ALS remote patient monitoring setting.

Methods: longitudinal analysis of acoustic and visual speech measures extracted from a web-based conversational assessment of people with ALS (pALS).

Implications: multimodal remote patient monitoring allows to (i) measure changes in speech markers frequently and cost-effectively, while (ii) capturing differences between slow and fast progressors.

Methods and Materials

- 54 pALS (Table 2) completed at least three sessions between October 2020 and July 2021 using a cloud-based multimodal dialog platform (illustration on Figure 1).
- Each session consisted of structured and spontaneous speech tasks (Table 1), followed by a self-reported ALSFRS-R questionnaire.
- Rate of progression was calculated based on first and last ALSFRS-R score. pALS were stratified into slow and fast progressors based on a threshold of 0.47 points/month.
- Statistical tests were conducted to identify acoustic and visual metrics for which the rates of change are significantly different between the two cohorts.

Results and Discussion

- A variety of metrics showed statistically significant differences in their rate of change between slow and fast progressors (Figure 2).
- In particular, metrics related to timing, voice quality, fundamental frequency, and higher order kinematics of the lower lip, eye opening and eyebrow positioning showed differences.
- These differences in trajectories signify changes in articulatory motor control and their acoustic implications as well as extracortical and pericortical motor control.
- For example, PPT in the Bamboo passage was increasing by 0.09%/month (median slope) for fast progressors, while it was decreasing by -0.07%/month for slow progressors (which might be attributed to a training effect of the repeated task performance).
- To confirm robustness of the findings, data from more participants over a longer period of time will be analyzed in a future study.

Limitations:

- Duration between participants first and last sessions varies (on average 135 ± 70 days), as well as the number of sessions from each participant.
- Sustained phonation speech samples exhibited browser-driven systematic attenuation of loudness, which can affect feature extraction.

Conclusions

- As hypothesized, frequent and continuous monitoring of acoustic and visual speech markers can capture objective physiological changes that may not be captured by subjective scales like the ALSFRS-R instrument, the current clinical standard to track progression in ALS.
- Changes in these audiovisual metrics could serve as potential digital biomarkers, which could contribute towards patient stratification and tracking of outcomes following pharmaceutical interventions.