modality.ai

# Analytical validation of facial metrics in Amyotrophic Lateral Sclerosis (ALS) extracted using a multimodal remote patient monitoring platform

*Cathy Zhang, Oliver Roesler, Jackson Liscombe, Reva Bajjuri, Hardik Kothare,*
***Vikram Ramanarayanan***

# Acknowledgements

This work was supported by the **NIH NIDCD grant R42DC019877**.

We thank our collaborators at EverythingALS and the Peter Cohen Foundation for participant recruitment and data collection.





# Motivation: Remote Assessment of ALS

- Amyotrophic Lateral Sclerosis (ALS) primarily affects motor neurons, leading to muscle weakness and atrophy.
- Remote audiovisual allows patients to be assessed from the comfort of their homes
- Most ALS patients experience some form of bulbar symptoms throughout their disease progression, which has symptoms starting from the face and the neck.
- Thus, facial metrics can provide insights into the progression and impact of the disease by measuring changes in orofacial movement, facial expressions, facial symmetry, etc. (e.g., Guarin et al., 2022; Neumann et al., 2021)

# How Do I Know My Measures Are Useful?


**Verification** evaluates sample-level sensor outputs

**Analytical validation** evaluates the performance of an algorithm to convert sensor outputs into physiological metrics using a defined data capture protocol in a specific subject population

**Clinical validation** evaluates whether the physiological metric acceptably identifies, measures, or predicts a meaningful clinical, biological, physical, functional state, or experience, in the stated context of use and specified population

**The V3 framework allows us to evaluate whether a digital measurement product is fit-for-purpose**

*Robin et al. (Digital Biomarkers 2020)*

# Research Question

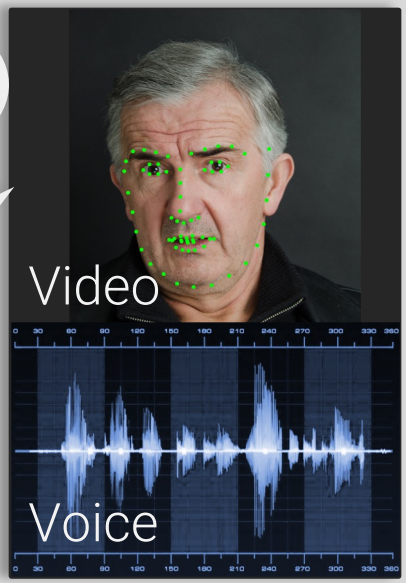
**Can analytically valid facial landmarks and metrics be extracted from standard tasks administered through a multimodal dialog agent?**

# Modality's Assessment Platform

## Virtual Guide

Tina

## Multimodal



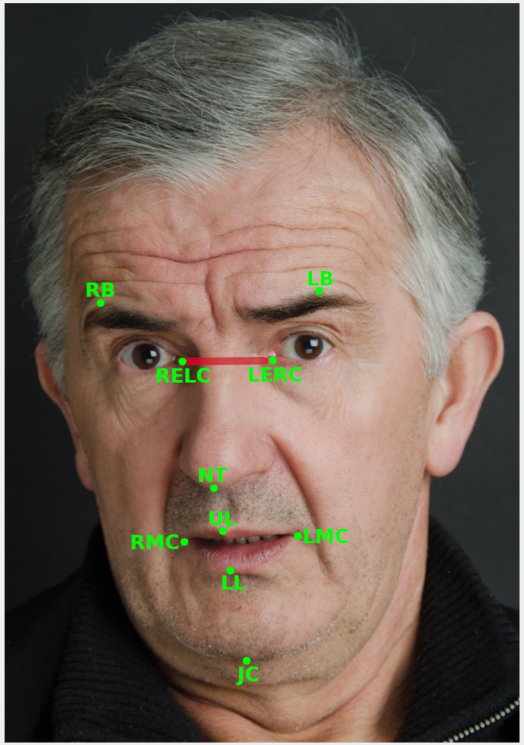

## Metrics

- ❑ Speech
- ❑ Language
- ❑ Facial
- ❑ Limb
- ❑ Cognitive
- ❑ Eye gaze
- ❑ Activities of daily living
- ❑ Patient Report of Problems™

# Dataset

- Used data collected in collaboration with EverythingALS
- Selected 90 sessions from 30 participants (3 sessions per participant)
  - 10 bulbar onset patients
  - 10 non-bulbar onset patients
  - 10 controls
- For this study, every second video frame from a video recording of the participant reading the last sentence of the Bamboo passage was used

# Annotations



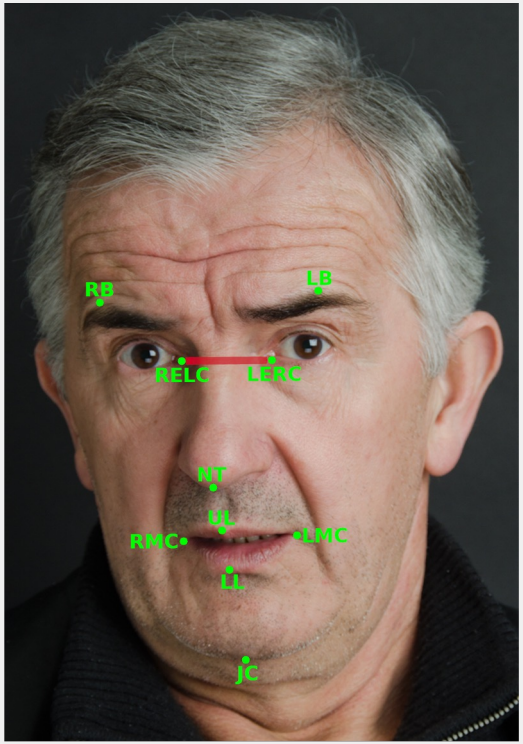

1. Annotation of 10 facial landmarks
    - for 90 sessions using MediaPipe Face Mesh (AP)
    - for 90 sessions by a first annotator (H1)
    - for 30/90 sessions by a second annotator (H2)
2. Calculation of 63 facial metrics for all annotation sets

| Domain | Metrics |
|---|---|
| Mouth<br>Movement<br>Eyes<br>Eyebrows | lip aperture, mouth surface area<br>velocity of lower lip and jaw center<br>eye opening<br>vertical eyebrow displacement |

# Analytical Tests

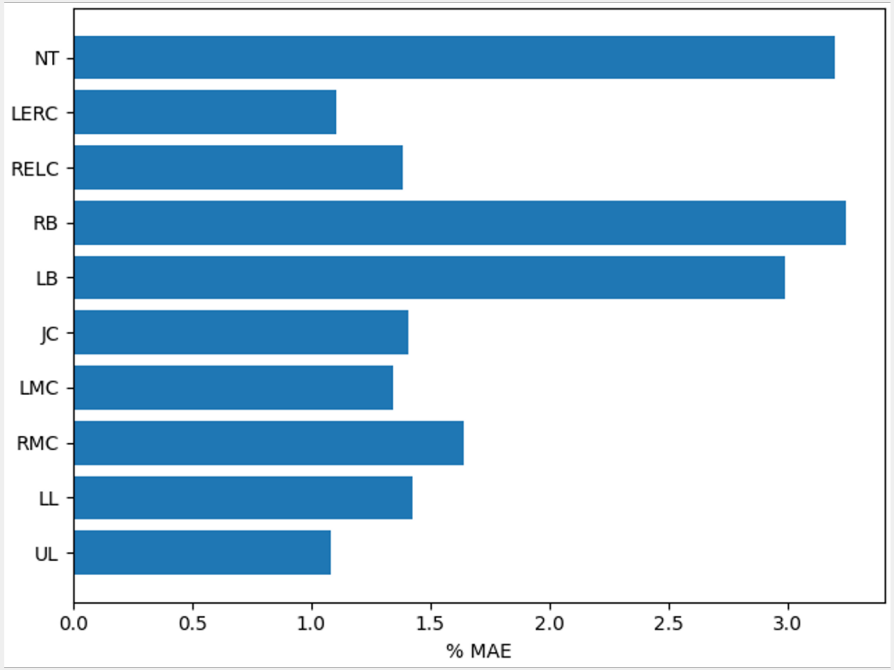
| | |
|---|---|
| Are automatically computed values significantly different from human-computed ones? | • correlation test<br>• distribution difference test<br>• mean absolute error (MAE)<br>• MAE normalized as a percentage<br>for both landmarks and metrics |
| Do the metrics show sign of being clinically valid? | Compute effect sizes on metrics from H1 and AP on pairs of cohorts and compute the Kruskal-Wallis test on all cohorts. Metrics that showed p-value <= 0.05 were then tested by Dunn's test. |

# Results: Landmarks




- Very strong correlation (Spearman correlation coefficient values between 0.9 and 1.0) between AP and H1 for all 10 landmarks
- No significant difference ($p > 0.05$) between AP and H1 as well as between H1 and H2 by the Mann-Whitney U test for all 10 landmarks

# Results: MAE between Hand-Annotated and Automated Landmarks




- The normalized MAE of landmark positions (the Euclidean R^2 distance divided by the range of the results) between AP and H1 ranges between 1.0 to 3.2%.

# Results: Statistical Differences between Hand-Annotated and Automated Metrics

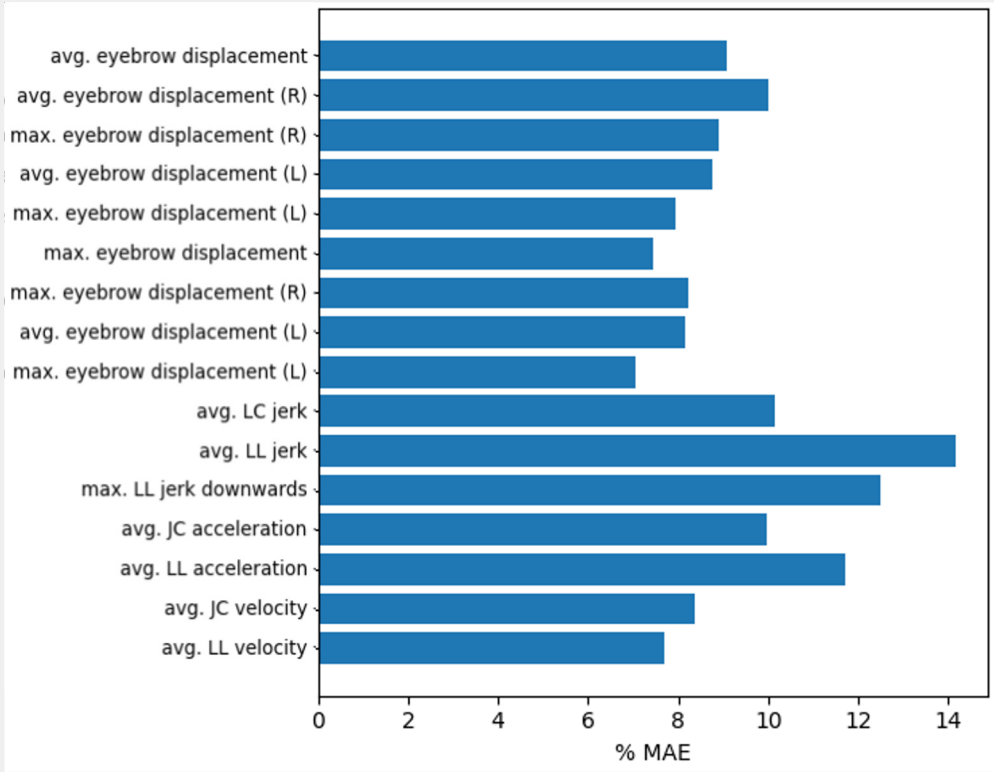
- When automatically-extracted metrics were compared to those derived using hand annotations, 16 metrics were consistently extracted (not significantly different) (Mann-Whitney U test). They fall into two groups:

  1. vertical eyebrow displacement
  2. movement (velocity, acceleration, jerk) of the lower lip or jaw center

  **10 of the 16 metrics** had a correlation coefficient > 0.5.

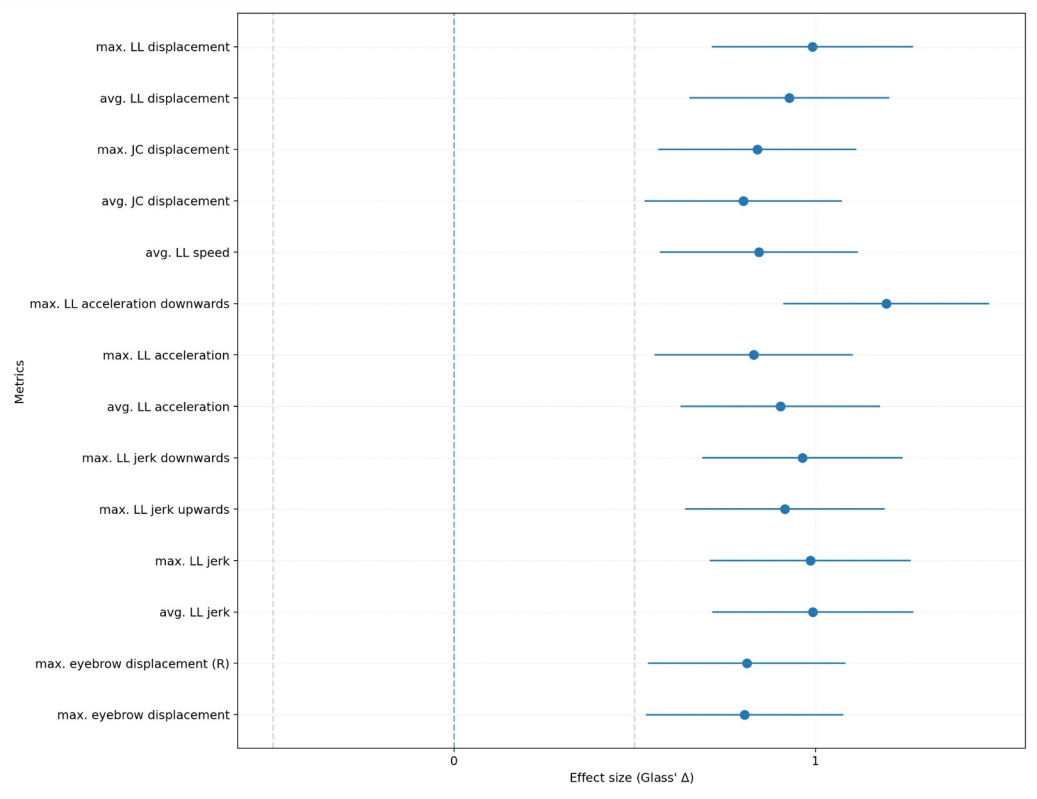
- There are more metrics not significantly different between H1 and H2 in addition to the 16 metrics, but belong to the same two groups above.

# Results: MAE between Hand-Annotated and Automated Landmarks




- The normalized MAE of all 63 metrics between AP and H1 ranges between 4.6 and 17.6
- The MAE of the 16 consistently extracted metrics ranges between 7.0 and 14.2
- Errors in landmarks are magnified in metrics as each metric is calculated from multiple landmarks

# Results: Clinically Valid Metrics




- 14 metrics survived post-hoc Dunn's tests for multiple comparison and showed effect sizes greater than 0.8 between non-bulbar and bulbar cohort pair

# Discussion

Automatically extracted facial landmarks are reasonably accurate.

A subset of facial metrics (movement features of jaw/lips, eyebrows) demonstrate analytical validity and strong and/or significant correlations between the predictions and annotations.

Caveats:

1. The 30 sessions may not be representative of the entire population.
2. Important to take steps to reduce measurement error. Compliance and adherence to task instructions can help ensure consistent performance.

# Questions?

- A subset of facial metrics (movement features of jaw/lips, eyebrows) demonstrate analytical validity and strong and/or significant correlations between the predictions and annotations.

- Important to take steps to reduce measurement error. Compliance and adherence to task instructions can help ensure consistent performance.

modality.ai

**Vikram Ramanarayanan**
v@modality.ai
Web: vikramr.com