

Quality comparison of remote vs. in-person digital speech assessment for dementia

Jessica Robin¹, Jekaterina Novikova¹, Sasha Sirotkin¹, Maria Yancheva¹, Liam D. Kaufman¹, William Simpson^{1,2}

(1) Winterlight Labs, Toronto, ON, Canada

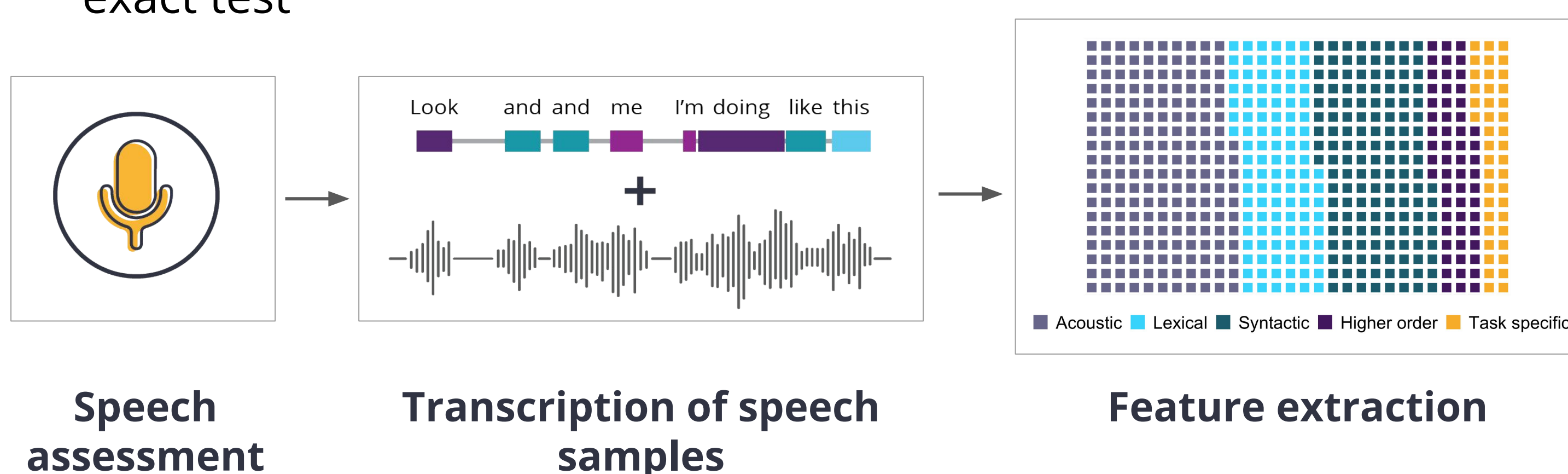
(2) Department of Psychiatry and Behavioural Neuroscience, McMaster University, Hamilton, ON, Canada

Background

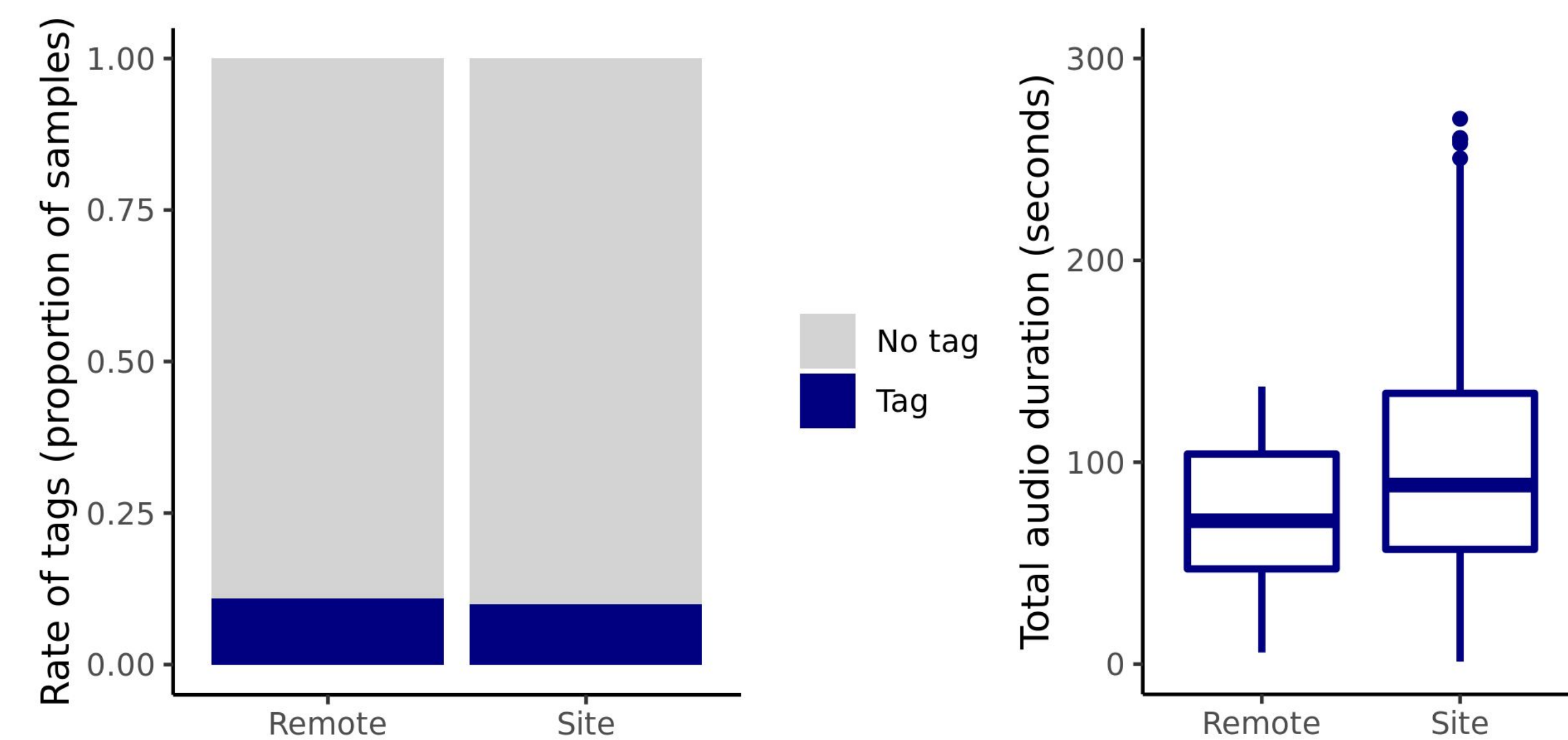
- Advancements in digital health technologies have the potential to enable remote patient assessment and monitoring^{1,2,3}
- Remote testing lowers the burden on patients and caregivers, and may enable more frequent and naturalistic assessment
- Digital speech assessments are an example of a digital health tool that can be remotely administered and offers insight into neurological and psychiatric health^{4,5,6}
- In this study, we compared the quality of speech assessments administered at home, with the help of a caregiver, to those administered in a clinical setting, in two samples of individuals with dementia (Alzheimer's Disease and Frontotemporal Dementia)

Methods

- Speech samples were examined from two ongoing studies:
 - A clinical trial for Alzheimer's Disease, with speech assessments conducted in a clinical setting, by a trained rater
 - An observational study of Frontotemporal Dementia, with speech assessments conducted at home, with the assistance of a caregiver trained on administering the assessment
- Speech assessments included picture description, phonemic and semantic fluency tasks
- In total, 575 speech samples from clinical sites (AD clinical trial) were compared to 574 speech samples from the remote study (observational FTD study)
- All speech samples were manually transcribed by trained transcriptionists
- As part of the transcription process, samples are tagged for any possible quality issues (see Table 1)
- The rates of all tags and each tag type were compared across samples from the clinical sites and remote sites using Fisher's exact test

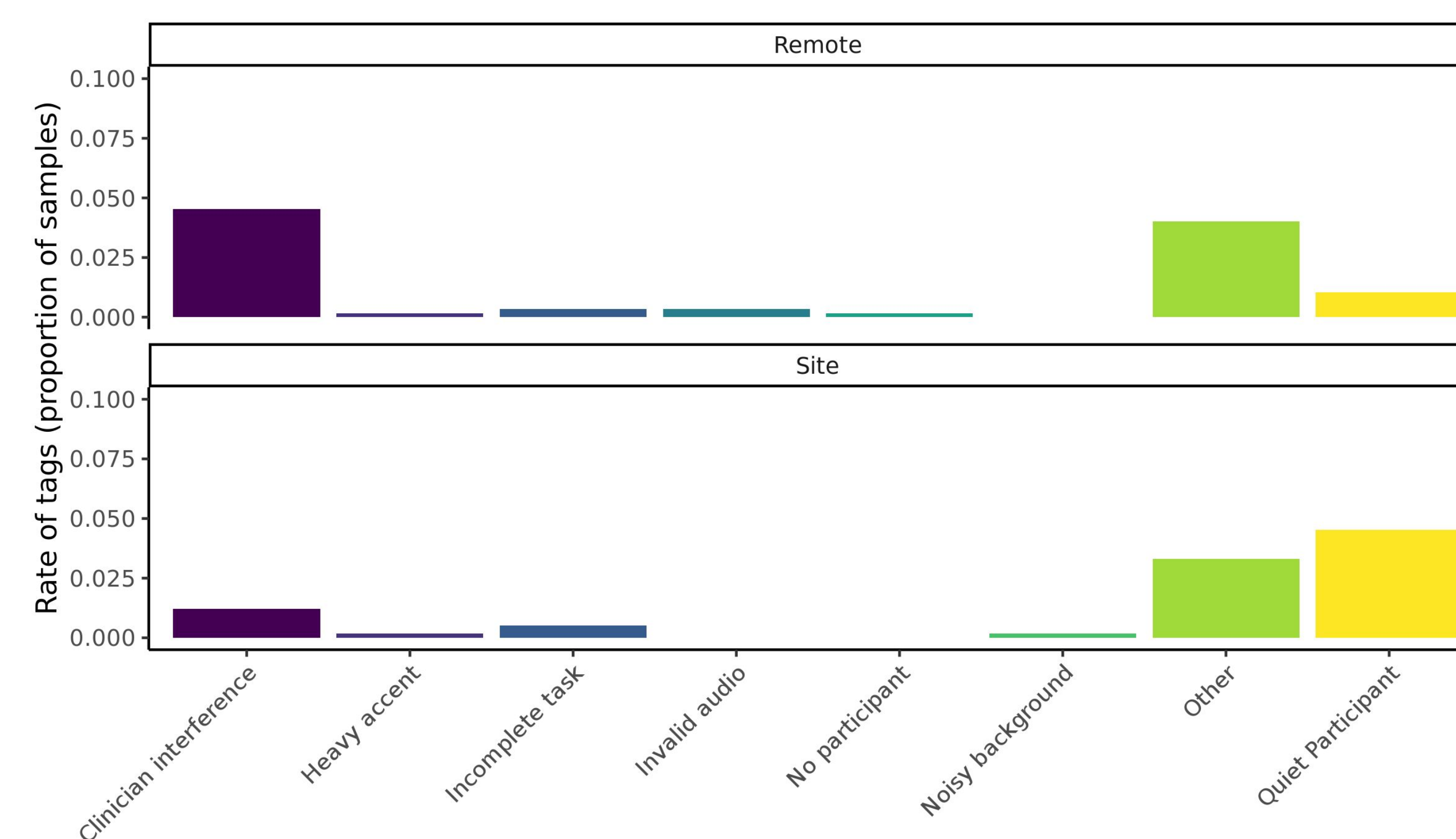


Figures 1 & 2: Comparison of overall incidence of tags and sample durations



- Overall rate of any tag did not differ significantly ($p > 0.6$) between speech samples collected at Remote (10.9%) and Clinical Sites (9.9%)
- Speech samples for picture description tasks were significantly longer in duration ($p < 0.001$) at Clinical Sites (mean duration = 101 seconds) compared to Remote samples (mean duration = 73 seconds)

Figure 3: Comparison of tag types across samples



- Clinician (caregiver) interference was more frequent for Remote samples ($p < 0.01$)
- Quiet participants were more frequent in the Site samples ($p < 0.001$)
- All other tags were equal or infrequent across samples

Table 1: Tags for possible quality issues

Transcription tags	Usage
Clinician interference	The person administering the speech assessment interferes with the task by providing assistance, encouragement or other commentary
Heavy accent	The participant has a heavy accent (including dialects and non-native accents)
Incomplete task	The participant did not complete the task or did not follow instructions
Low audio quality	The audio quality of the sample is poor, due to artifacting or distortion
Invalid audio	There is no audio recorded or the audio cannot be used
No participant	There is no participant audible in the sample
Noisy background	There is high amounts of noise in the background either from the other people speaking or from the environment
Overwhelming noise	There is background noise to the point that the participant cannot be heard
Quiet participant	The participant is hard to hear due to low volume, whispering or mumbling
Other	Any other possible quality issue with the audio, not covered by the categories above

Conclusions

This study suggests that remote speech assessments yield recordings of comparable quality to in-person assessments. We found higher, though still low, rates of caregiver interference for remote assessment, which should be monitored and mitigated in future remote assessment. Surprisingly, recordings from clinical sites had higher instances of quiet participants, which could be due to microphone placement. Remote assessments yielded shorter recordings, but this may be due to the different dementia diagnoses across groups. Future work should compare the same participants across both assessment settings.

References

1. Piau A, Wild K, Mattek N, Kaye J. Current State of Digital Biomarker Technologies for Real-Life, Home-Based Monitoring of Cognitive Function for Mild Cognitive Impairment to Mild Alzheimer Disease and Implications for Clinical Care: Systematic Review. *J Med Internet Res*. 2019 Aug 30;21(8):e12785.
2. Bakker JP, Goldsack JC, Clarke M, Coravos A, Geoghegan C, Godfrey A, et al. A systematic review of feasibility studies promoting the use of mobile technologies in clinical research. *npj Digit Med*. 2019 Dec;2(1):47.
3. Kourtis LC, Regele OB, Wright JM, Jones GB. Digital biomarkers for Alzheimer's disease: the mobile/wearable devices opportunity. *npj Digital Med*. 2019 Dec;2(1):9.
4. Fraser KC, Meltzer JA, Rudzicz F. Linguistic Features Identify Alzheimer's Disease in Narrative Speech. *JAD*. 2015 Oct 15;49(2):407-22.
5. Fraser KC, Meltzer JA, Graham NL, Leonard C, Hirst G, Black SE, et al. Automated classification of primary progressive aphasia subtypes from narrative speech transcripts. *Cortex*. 2014 Jun;55:43-60.
6. Boschi V, Catricalà E, Consonni M, Chesi C, Moro A, Cappa SF. Connected Speech in Neurodegenerative Language Disorders: A Review. *Front Psychol*. 2017 Mar;8.