INTERSPEECH 2019
September 15–19, 2019, Graz, Austria

Use of Beiwe Smartphone App to Identify and Track Speech Decline in Amyotrophic Lateral Sclerosis

Kathryn P. Connaghan[1], Jordan R. Green[1], Sabrina Paganoni[2,3,4], James Chan[5], Harli Weber[3], Ella Collins[3], Brian Richburg[1], Marziye Eshghi[1], JP Onnela[6], James D. Berry[2,3]**

[1]Speech and Feeding Disorders Lab, MGH Institute of Health Professions, USA
[2]Harvard Medical School, School of Medicine, USA
[3]Neurological Clinical Research Institute, Department of Neurology, Massachusetts General Hospital, USA
[4]Department of Physical Medicine and Rehabilitation, Spaulding Rehabilitation Hospital, USA
[5]Massachusetts General Hospital Department of Biostatistics, USA
[6]T.H. Chan Harvard School of Public Health, USA

{kconnaghan, jgreen2, brichburg, meshghi}@mghihp.edu
{spaganoni, jchan12, hweber1, ecollins15, jdberry}@mgh.harvard.edu,
onnela@hsph.harvard.edu

Abstract

The capacity for smartphones to remotely capture speech data affords significant clinical and research opportunities for degenerative neurologic diseases such as amyotrophic lateral sclerosis (ALS). Longitudinal data may inform ALS disease prognosis, facilitate timely intervention, and document response to treatment [1]. A recent study established the feasibility of the Beiwe smartphone-based digital phenotyping to track the clinical progression of ALS across multiple domains [2]. The current investigation extends this work to address the utility of Beiwe to identify and track speech decline in ALS. Twelve participants with ALS used the Beiwe app weekly to record reading passages and self-report (ALSFRS-R) ratings of bulbar (speech) function. Speaking rate and pause variables were automatically extracted from recordings offline [3]. Speech function measures at baseline were significantly different for participants with and without bulbar symptoms. In addition, the rate of decline of all measured speech functions was greater for participants with bulbar symptoms. The successful use of Beiwe for speech function analysis suggests that smartphone-based capture of speech has potential for diagnostic screening and disease progress monitoring in ALS. Further large sample investigation across a comprehensive set of speech variables is warranted.

Index Terms: Amyotrophic lateral sclerosis, Beiwe mobile app, speech acoustics, ALSFRS-R

1. Introduction

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disease characterized by loss of muscle strength and function. A particularly devastating aspect of the disease is the involvement of the bulbar (speech, swallowing) systems. Though a minority of patients present with bulbar symptoms upon diagnosis (approximately 30%), most patients will develop bulbar symptoms throughout the course of the disease. Not only do bulbar symptoms significantly reduce quality of life, but their onset signifies reduced survival (see [1] for a review).

The early identification of the onset of bulbar symptoms is critical to providing optimal care in ALS. Additionally, the discovery and tracking of biomarkers, such as those found in the speech signal, to monitor disease progression is crucial to appropriate treatment planning and for establishing response to treatment. Subjective measures, such as clinician-ratings and patient-report, are the traditionally accepted approaches to identifying speech impairment and tracking speech decline. However, there is a clear need for objective measures, as was observed in a study finding that instrumental speech measures detected the onset of bulbar symptoms in ALS before speech-language pathologists or the patients themselves [4].

Recently, a number of instrumental measures that leverage the speech signal have been identified as ideal candidates for identifying the onset of bulbar symptoms and tracking decline in ALS. Speech features such as speaking rate, percent pause time, and articulator movement velocity have all been shown to demonstrate decline before perceptible changes in intelligibility (see [4]). In particular, speaking rate and pause analyses have proven to be robust indicators of bulbar decline in ALS. For instance, percentage pause time has been shown to differentiate between ALS speakers with and without intelligibility deficits. In the same study, speaking rate was found to not only differentiate between ALS speakers with and without intelligibility deficits, but also between ALS speakers without intelligibility deficits and healthy controls [1,5].

With the ubiquity of smartphones and the development of mobile technology, the acquisition of speech signals for clinical and research opportunities is convenient for patients, and can be implemented remotely, frequently, and in a minimally invasive manner. These qualities are particularly

*these authors share senior authorship

http://dx.doi.org/10.21437/Interspeech.2019-3126

well-suited for patients with ALS, who may experience barriers to accessing health care and research opportunities [6]. We have previously defined digital phenotyping as the "moment-by-moment quantification of the individual-level human phenotype *in situ* using data from personal digital devices, in particular smartphones." [7,8] As part of our efforts in this area, we have developed the Beiwe research platform for smartphone-based digital phenotyping. The platform front-end consists of native smartphone applications for Android and iOS devices, and the back-end is based on Amazon Web Services (AWS) cloud computing infrastructure, making it scalable and globally deployable. The Beiwe front-end app can capture both passive (e.g., phone sensors and logs) and active (e.g., surveys / EMA and voice recordings). To facilicate reproducibility of research and enable re-analyses of collected data, the Beiwe front-end applications collect raw (vs. pre-processed) sensor data and all data collection settings are captured in JSON-formatted configuration files. Data are secured through encryption while stored on the phone awaiting transit, during transit, and while stored on the server. The Beiwe platform is available as open source software (https://github.com/onnela-lab).

The feasibility of Beiwe to track ALS disease progression across multiple domains was established in a previous investigation [2]. Participants completed the Revised ALS Functional Rating Scale (ALSFRS-R), a valid and reliable self-report measure of functioning for daily living for individuals with ALS [9]. The ALSFRS-R spans multiple domains and includes questions about bulbar function (speech, swallowing, salivation), gross and fine motor skills (e.g., handwriting, cutting food, walking) and other activities of daily living (dressing and hygiene). ALSFRS-R scores acquired using Beiwe were compared with those acquired through clinic-based administration. In-clinic and Beiwe scores showed a high level of agreement at baseline. In addition, frequent recordings of reading passages were analyzed for their pause content, a measure that tracks bulbar decline [1,5]. A statistically nonsignificant increase in mean pause time over the study period was observed. This finding may be explained by the pooling of participants with and without bulbar symptoms. While patients with bulbar symptoms show relatively rapid speech decline, patients without bulbar symptoms may experience long periods of stable speech function. It is anticipated that stratifying participants based on the presence of bulbar symptoms would reveal differing rates of change.

While the feasibility of Beiwe to track ALS disease progression remotely across domains has been demonstrated [2], the current preliminary study evaluates its utility to identify the group differences between participants with and without bulbar symptoms and to track speech decline. We explored its feasibility for frequent collection of speech recordings and self-ratings. Expanding upon previous study of Beiwe for ALS disease monitoring, we further described differences in variables based on the presence or absence of bulbar symptoms

2. Methods

2.1. Participants

Data were collected from 12 participants (age range = 44-63 years, Mean=54.3, SD = 6.2 years; 6F, 6M*)* meeting the El Escorial Criteria for diagnosis of ALS as part of a larger study of smartphone-based digital phenotyping of ALS progression [2]. All participants were able to provide informed consent and were judged as likely to comply with the procedures by study staff. Only individuals who owned a smartphone running the iOS or Android operating system and reported at least moderate phone use were included in the study. Study enrollment was conducted on a rolling basis, with the length of enrollment varying from 3 to 51 weeks across participants. A total of 181 sessions were analyzed.

2.2. Procedures

Details about data collection procedures have been reported elsewhere [2]. Briefly, following consenting procedures, the Beiwe app was installed on each participant's smartphone and study staff provided training on the app and procedures for filling out surveys and making speech recordings. For the speech (passage reading) recording task, participants were instructed to "recite the following passage in your usual voice".

2.2.1. Materials & Data Acquisition

Each week, Beiwe sent a phone notification to participants to complete the ALSFRS-R questionnaire and passage reading task on the phone. Participants then used the phone to record themselves reading the Bamboo passage, which was displayed on the screen. The Bamboo passage is a short passage developed for the 5th grade reading level. The passage was designed to aid in the automatic detection of pause boundaries [3,5]. For instance, voiced consonants were placed at word boundaries to better highlight pause-speech boundaries than unvoiced consonants [5]. Speech samples were recorded at 44.1kHz sampling rate with no compression.

The ALFSRS-R is a rating scale designed to measure physical function in carrying out activities of daily living by individuals with ALS [9]. While the ALSFRS-R provides an overall score based on 12 questions to assess function across multiple domains, the current investigation focuses only on the speech and bulbar total function (speech, swallowing, saliva) questions. For the speech score, participants are asked to rate their speech from 4 (normal speech processes) to 0 (loss of useful speech). The bulbar total score of the ALSFRS-R is calculated from responses to the speech, swallowing, and salivation questions, for a total possible score of 12.

2.2.2. Data Processing

Speech recording files were pre-processed offline using Audacity 2.2.2. to reduce noise and to remove word insertions, word repetitions, filled pauses (e.g., un, uh) and extraneous non-speech sounds, (e.g., throat clearing, laughter). Pauses prior to the onset of passage reading and following the end of passage reading were further removed. Speech Pause Analysis (SPA) software, a semi-automated MATLAB speech pause segmentation procedure [3], was implemented to extract the target speech and pause variables. The minimum pause threshold was set to 300 ms and minimum speech threshold set to 25 ms [5].

SPA analysis calculates a variety of variables, we analyzed the following three; speaking rate, articulation rate, and percentage of pause time. See Table 1 for operational definitions.

2.2.3 Stratification of Participants

Participants were grouped into those presenting with (Bulbar) and without (No-Bulbar) bulbar symptoms at the onset of the study. Participants with speaking rate at or below 150 words per minute (WPM) for at least the first two sessions were categorized as Bulbar (n=5), while those with speaking rates above 150 WMP were categorized as No-Bulbar (n=7) [4].

Table 1. *List of speech and pause variables*

Variable	Operational Definition
Speech rate (WPM)	# words / total duration of reading task
Articulation rate (Syll/Sec)	# syllables / duration of speaking time of reading task
Percentage pause time (%)	(total pause duration/total duration of reading task) * 100

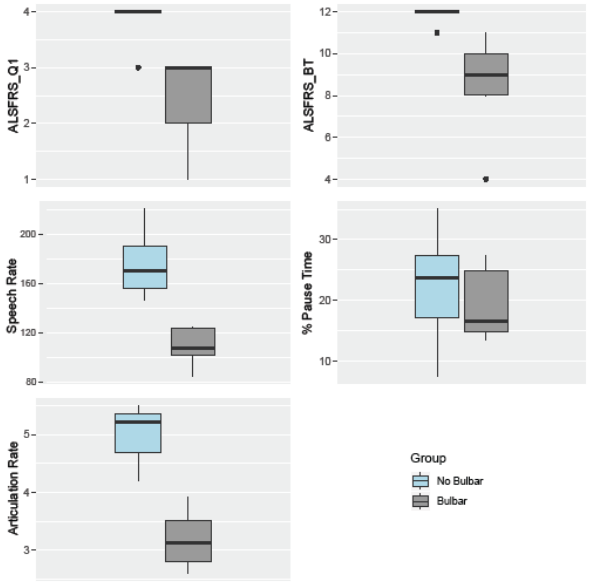

Figure 1. *Boxplots of Non Bulbar and Bulbar groups speech function (ALSFRS-R Question 1, ALSFRS-R bulbar total score, speech rate, percentage pause time, articulation rate) at baseline.*

2.2. Analysis

We conducted two types of analysis. First, we examined group (Bulbar vs No Bulbar) differences in speech function at a single time-point, baseline. The ability to identify group differences lays the foundation for diagnostic screening for the early detection of the onset of bulbar symptoms. A Mann-Whitney-Wilcoxon Test was conducted for the group comparison of the ALSFRSQ1 (ALSFRS-R speech question score). Separate Welch two-sample t-tests were conducted for the remaining dependent variables, including speech rate, percentage pause time, articulation rate, and ALSFRSBT (ALSFRS-R bulbar total score). Next, we fitted separate linear mixed-effects models with fixed effects for week, bulbar group and an interaction between week and bulbar group (week*group), and a random intercept for each participant. The interaction effect from these models was used to evaluate whether participants with bulbar symptoms changed at a different rate than those without bulbar symptoms at baseline.

3. Results

3.1. Group Differences at Baseline

The boxplots in Figure 1 describe the No Bulbar (blue) and Bulbar (grey) data for each dependent variable. Significant differences were observed between groups for ALSFRSQ1 (ALSFRS-R speech question) ($W = 33.5$, $p = 0.006$) and ALSFRSBT (ALSFRS-R bulbar total score) ($t(4.11) = 2.84$, $p = 0.05$). Baseline speaking rate was significantly different between the groups, as anticipated given stratification based upon this variable ($t(9.90) = 5.35$, $p = 0.001$). Articulation rate ($t(5.57) = 5.21$, $p = 0.002$) also differed between the groups. The variable percentage pause time was not different between the groups at baseline ($t(10.00) = 0.61$, $p = .56$).

3.2. Changes in Speech Measures over Time

Figures 2-4 illustrate each participant's speech function across time for the acoustic variables. Because only two data points were collected beyond 38 weeks, data up to 40 weeks are presented. Participants in the No Bulbar group are depicted with a black square and in the Bulbar group are shown with grey circle markers. Descriptively, for speaking and articulation rate, participants in the Bulbar group demonstrated declining speech function, where the participants in the No Bulbar group demonstrated relative stability.

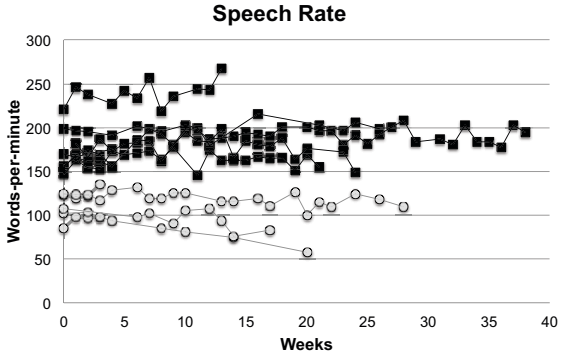

Figure 2. *Speech rate of No Bulbar (black squares) and Bulbar (grey circles) participants across time.*

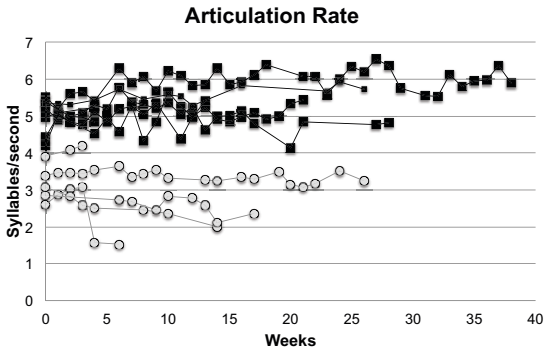

Figure 3. *Articulation rate of No Bulbar (black squares) and Bulbar (grey circles) participants across time.*

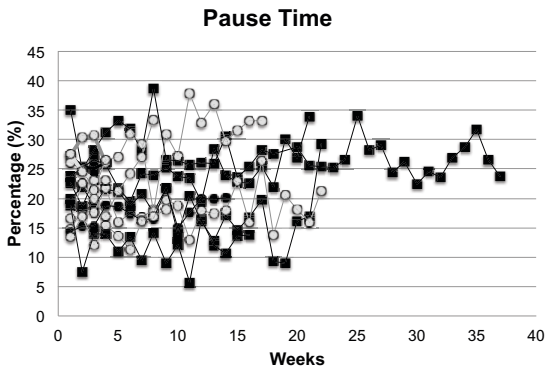

Figure 4. *Percentage pause time of No Bulbar (black squares) and Bulbar (grey circles) participants across time.*

Results from the linear mixed-effects analyses (LME) revealed that only the ALSFRS-R bulbar total score and articulation rate significantly changed across time when data were pooled across participants. However, the week*group interaction (difference in across time between groups) indicated that the Bulbar group demonstrated significantly faster rates of decline compared to the No Bulbar group for ALSFRS-R Question 1, ALSFRS-R bulbar total score, speaking rate, and articulation rate ($p<0.001$). The increase in percentage pause time across time was also statistically significantly for the Bulbar group compared to the No Bulbar group ($p<0.0001$). For the Group analysis, No Bulbar was the reference group, and the group effect reflected the comparison of the model with and without the Bulbar group. This analysis suggested that the group differences observed at baseline were maintained across the study period. Table 3 summarizes the results of the mixed-effect models, including the coefficients and level of statistical significance for the intercept, week (change across time), group, and week*group interaction.

Table 3. *Summary of LME Findings (Coefficients and p-values). Group refers to comparison of No Bulbar and Bulbar. Dependent variables include ALSFRS-R speech score, (FRSQ1), ALSFRS-R bulbar total score (FRSBT), speech rate (Sp Rate), percentage pause time (% Pause), and articulation rate (Art Rate).*

LME Model	Intercept	Week	Group	Week* Group
FRSQ1	3.98**	-0.01^{NS}	-1.52**	-0.03**
FRSBT	11.89**	-0.02*	-3.28*	-0.04**
Sp Rate	181.71**	0.11^{NS}	-70.69**	-1.02*
% Pause	22.90**	-0.03^{NS}	-3.07^{NS}	0.29**
Art Rate	5.17**	0.01*	-1.95**	-0.03**

** $p<0.001$ *$p<0.01$ ^$p<0.05$ $^{ns}p>0.05$

4. Discussion

The preliminary findings suggest the feasibility of using the Beiwe platform to identify differences between participants with and without bulbar symptoms and to track speech decline in ALS. Participants with ALS were able to comply with the research protocol, with 181 sessions successfully completed. Speech recordings were of sufficiently high quality for the implementation of the SPA acoustic analysis.

We found significant group (No Bulbar vs. Bulbar) differences on speech function measured at baseline, highlighting the potential for development of a Beiwe diagnostic screening tool to identify the onset of bulbar symptoms. This early diagnosis is critical to patients with ALS, as speech decline can be rapid from onset and successful planning to address impaired and/or loss of speech can affect communication outcomes. Moreover, the LME Group findings indicated that these differences were maintained across the course of the study

Our data further showed that participants with bulbar symptoms demonstrated more rapid decline of all speech function measures over time, even though some measures (ALSFRS-R speech question, speaking rate, percentage pause time) did not significantly decline across time for the overall data set. This result supports previous findings of patients experiencing more rapid decline following the onset of bulbar symptoms [1]. In addition, percentage pause time showed a significant week*group interaction, despite not showing group differences at baseline or throughout the study.

As this pilot investigation comprised a small number of participants, the findings should be interpreted with caution. Despite this limitation, we were able to demonstrate the feasibility of employing Beiwe to collect speech recordings and self-report data, as well as identify group differences across speech measures at baseline and over time. Future investigation with a larger sample will include a comprehensive set of acoustic speech variables, as well as acoustic analysis of cough recordings to inform bulbar function. We further aim to incorporate the capability of Beiwe to capture passive measures (e.g., GPS, phone/text logs) to develop a comprehensive picture of the communicative participation changes experienced by individuals with ALS.

5. Conclusions

The Beiwe platform provides a viable approach for the identification of bulbar symptoms and to track speech decline in ALS. Its convenience for frequent data collection at any time of day and myriad locations can potentially transform health care delivery and research. The ease and non-invasiveness of recording the speech signal for use as a biomarker will potentially reduce barriers, such as limitations to travel to services, and thereby promote adherence to clinical and research protocols and increased access to services.

6. Acknowledgements

This study was supported by the Winthrop Family Fund for ALS Science at Massachusetts General Hospital, and NIH grants DP2-MH103909 and K24DC016312. We appreciate analysis support provided by Kenzie Carlson and Josh Barback. We thank the supporters of the Winthrop Family Fund for ALS Science and all of the participants in this study.

7. References

[1] J. R. Green, Y. Yunusova, M. S. Kuruvilla, J. Wang, G. L. Pattee, L. Synhorst, L. Zinman and J. D. Berry, "Bulbar and speech motor assessment in ALS: Challenges and future directions", *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration,* 14(7-8), 494-500, 2013.

[2] J. D. Berry, S. Paganoni, K. Carlson, K. Burke, H. Weber, P. Staples, J. Salinas, J. Chan, J. R. Green, K. Connaghan, J. Barback and J. P. Onnela, "Design and results of a smartphone-based digital phenotyping study to quantify ALS progression", *Annals of Clinical and Translational Neurology*, April 2019.

[3] J. R. Green, D. R. Beukelman and L. J. Ball, "Algorithmic estimation of pauses in extended speech samples of dysarthric and typical speech", *Journal of Medical Speech-language Pathology,* 12(4), 149, 2004.

[4] K. M. Allison, Y. Yunusova, T. F. Campbell, J. Wang, J. D. Berry and J. R. Green, "The diagnostic utility of patient-report and speech-language pathologists' ratings for detecting the early onset of bulbar symptoms due to ALS", *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, 18(5-6), 358-366, 2017.

[5] Y. Yunusova, N. L. Graham, S. Shellikeri, K. Phuong, M. Kulkarni, E. Rochon, D. F. Tang-Wai, T. W. Chow, S. E. Black, L. H. Zinman and J. R. Green, "Profiling speech and pausing in amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD)." *PloS one,* 11(1), e0147573, 2016.

[6] S. B. Rutkove, K. Qi, K. Shelton, J. Liss, V. Berisha and J. M. Shefner, "ALS longitudinal studies with frequent data collection at home: study design and baseline data", *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration,* 1-7, 2018.

[7] J. Torous, M. V. Kiang, J. Lorme, and JP Onnela, "New tools for new research in psychiatry: A scalable and customizable platform to empower data driven smartphone research", *JMIR Mental Health,* 3(2), e16, 2016.

[8] JP Onnela and S. L. Rauch, "Harnessing smartphone-based digital phenotyping to enhance behavioral and mental health", *Neuropsychopharmacology,* 41, 1691, 2016.

[9] J. M. Cedarbaum, N. Stambler, E. Malta, C. Fuller, D. Hilt, B. Thurmond, A. Nakanishi and BDNF ALS Study Group, "The ALSFRS-R: a revised ALS functional rating scale that incorporates assessments of respiratory function." *Journal of Neurological Sciences*, 169(1-2), 13-21, 1999.