


# BMJ Open Protocol for psychometric evaluation of the Amyotrophic Lateral Sclerosis - Bulbar Dysfunction Index (ALS-BDI): a prospective longitudinal study

Yana Yunusova <sup>1,2</sup>, Ashley Waito,<sup>2</sup> Carolina Barnett,<sup>3,4</sup> Anna Huynh,<sup>2,5</sup> Rosemary Martino,<sup>1,6</sup> Agessandro Abrahao,<sup>4,7</sup> Gary L Pattee,<sup>8</sup> James D Berry,<sup>9</sup> Lorne Zinman,<sup>4,7</sup> Jordan R Green<sup>10,11</sup>

**To cite:** Yunusova Y, Waito A, Barnett C, *et al.* Protocol for psychometric evaluation of the Amyotrophic Lateral Sclerosis - Bulbar Dysfunction Index (ALS-BDI): a prospective longitudinal study. *BMJ Open* 2022;**12**:e060102. doi:10.1136/bmjopen-2021-060102

► Prepublication history for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2021-060102>).

Received 13 December 2021  
Accepted 07 February 2022



© Author(s) (or their employer(s)) 2022. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

## Correspondence to

Dr Yana Yunusova;  
[yana.yunusova@utoronto.ca](mailto:yana.yunusova@utoronto.ca)

## ABSTRACT

**Introduction** Early detection and tracking of bulbar dysfunction in amyotrophic lateral sclerosis (ALS) are critical for directing management of the disease. Current clinical bulbar assessment tools are lacking, while existing physiological instrumental assessments are often inaccessible and cost-prohibitive for clinical application. The goal of our research is to develop and validate a brief and reliable, clinician-administered assessment tool—the ALS-Bulbar Dysfunction Index (ALS-BDI). This publication describes the study protocol that has been established to ascertain the tools' psychometric properties.

**Methods and analysis** The ALS-BDI's development closely follows guidelines outlined by the COSMIN-based Standards for the selection of health Measurement Instruments (COSMIN). Through the proposed study protocol, we expect to establish psychometric properties of both individual test items of the ALS-BDI as well as the final version of the entire tool, including test-retest and inter-rater reliability, construct validity using gold-standard assessment methods and responsiveness.

**Ethics and dissemination** This study has been reviewed and approved by research ethics boards at two data collection sites: Sunnybrook Health Science Centre, primary (Toronto, Canada; ID3080) and Mass General Brigham (#2013P001746, Boston, USA). Prior to participation in the study, the participants sign the informed consent in accordance with the Declaration of Helsinki. Once validated, the ALS-BDI will be disseminated to key stakeholders. Following validation, the ALS-BDI and any required training material will be implemented for clinical use in a context of a multidisciplinary ALS clinic and used as an outcome measure for clinical trials in ALS research.

## INTRODUCTION

Amyotrophic lateral sclerosis (ALS) is a fast-progressing disease with a lifespan post-diagnosis of 2–5 years.<sup>1</sup> The disease is characterised by the degeneration of motor neurons in the brain, brainstem and spinal cord as well as extramotor (eg, cognitive-linguistic) brain pathways.<sup>2–4</sup> The degeneration of

## Strengths and limitations of this study

- The Amyotrophic Lateral Sclerosis - Bulbar Dysfunction Index (ALS-BDI) development process adheres to the established methodological guidelines for developing a formative assessment measure—Consensus-based Standards for the selection of health Measurement Instruments (COSMIN).
- Initial development and design of the ALS-BDI has involved extensive consultation of projected end-users (ie, neurologists, speech-language pathologists) to ensure that it is comprehensive, clinically feasible and aligned with current practice goals and workflow.
- Psychometric evaluation of the ALS-BDI will include test-retest and inter-rater reliability, construct validity and responsiveness of each test item and the overall tool.
- Patient anchors will be used to inform interpretability and determine the clinical significance of ALS-BDI scores, including the minimal important difference and low activity disease states.
- Many individual items on the drafted ALS-BDI involve a subjective rating of severity (ie, mild, moderate, severe), which may challenge inter-rater reliability.

motor neurons results in progressive muscle weakness, atrophy and eventual paralysis. Nearly 90% of individuals diagnosed with ALS will experience speech and swallowing dysfunction either at the onset of the disease or when the disease spreads to the bulbar motor system, which controls muscles of the head and neck.<sup>5</sup> The presence of bulbar motor impairment is associated with a shorter survival (<2 years) and increases the risk of death by almost eightfold.<sup>6,7</sup> The resulting speech impairment has been rated as the worse aspect of the disease by patients with ALS.<sup>8</sup> Despite the devastating consequences of bulbar dysfunction on the survival and

quality of life, there are currently no validated tools for its assessment in a clinic. In the absence of a standardised assessment of bulbar motor dysfunction, current practices in ALS clinics are idiosyncratic, piecemeal and rely primarily on symptom checklists.<sup>9</sup>

The overall aim of our work is to develop a reliable and psychometrically validated clinician-administered assessment tool of bulbar dysfunction, called the ALS-Bulbar Dysfunction Index (ALS-BDI), that is efficient, standardised, clinically feasible, comprehensive and responsive to change over time. The ALS-BDI is currently being designed and validated for two diagnostic use cases: (1) to discriminate among patients with different severities of bulbar impairment (ie, discriminative purpose); and (2) to evaluate the changes over time in a person or group (ie, evaluative purpose).<sup>10</sup> This tool will have immediate implications for tracking disease progression and monitoring changes in ALS clinics or in clinical trials. In the future, it may serve as the core assessment that can be adapted for improving ALS diagnosis and prognosis; it could also eventually serve as a screening tool for determining the need for full dysphagia and speech assistive technology evaluations.

Our process for developing the ALS-BDI adheres to established guidelines for developing a formative measure (COnsensus-based Standards for the selection of health Measurement INstruments, COSMIN).<sup>11 12</sup> Within a formative model, all items are considered relevant to represent the overall construct of interest and are not interchangeable; items may or not be correlated, and they may have different patterns of change when the construct changes. Formative measures, as opposed to reflective measures, are not assessed with respect to internal consistency or item-total correlations such as Cronbach's alpha, factor analysis or the item response theory. The decision to retain or reduce items is primarily based on their clinical relevance in defining the construct (content and face validity), although reliability and responsiveness can also inform item retention.

Our test development process is divided into three development cycles: Cycle 1 includes generation of the candidate item pool and demonstration of their content and face validity; Cycle 2 is focused on establishing the item and overall tool's reliability; and Cycle 3 establishes the item and overall tool's construct validity in relation to well established instrumental and clinical measures, as well as responsiveness to change over time. The work in Cycle 1, which produced the first iteration of the ALS-BDI, has been now completed. The initial development steps have been described in detail elsewhere.<sup>13</sup> Briefly, Cycle 1 involved convening a panel of experts to identify key assessment domains for bulbar dysfunction. The panel met during a Northeastern ALS Consortium (NEALS) Bulbar Subcommittee meeting, members of which were experts in neurology, speech–language pathology (SLP) and measurement science.<sup>14</sup> Through consensus, the panel identified three core bulbar assessment domains: (1) Cranial Nerve Exam; (2) Auditory-Perceptual Assessment

(of speech and voice); and (3) Functional Assessment (of speech intelligibility, swallowing, chewing and coughing). Following the meeting, a literature review was conducted by our development team to generate candidate items for each domain. Expert surveys were then conducted to probe each item's face and content validity. Based on this procedure, an initial draft of ALS-BDI was created. A group of SLPs further vetted the testable draft of the instrument via pilot administration of the test draft and completion of a cognitive interview with a member of our development team. As a result, a beta of ALS-BDI has been prepared for its psychometric evaluation.

The process for establishing psychometric properties of the ALS-BDI has been outlined in the recommendations by the COSMIN group.<sup>11</sup> To our knowledge, no comprehensive bulbar assessment tool has yet been developed and validated using this rigorous approach.<sup>15</sup> The key psychometric properties of the ALS-BDI that must be assessed under this framework include inter-rater and test–retest reliability (Cycle 2 of the development), as well as the construct validity against gold-standard measures of bulbar dysfunction in ALS and responsiveness to change and minimal important difference (MID), minimal detectable change (MDC) and low disease state of ALS-BDI to enhance interpretability of its scores (Cycle 3 of the development). Here we present a detailed protocol for the psychometric evaluation of ALS-BDI through Cycle 2 and Cycle 3.

## METHODS AND ANALYSIS

### Prospective study design

A prospective longitudinal study design will be used to assess the ALS-BDI.

### Participants and recruitment

Data for this study will be collected at the Bulbar Function Laboratory at the Sunnybrook Health Sciences Centre in Toronto, Ontario, Canada, and the Speech and Feeding Disorders Lab at the MGH Institute for Health Professions in Boston, Massachusetts, USA. The study has been initiated in May 2021 with anticipated completion by June 2023.

Possible, probable or definite ALS was diagnosed in 100 adults ( $\geq 18$  years) as defined by the Revised El Escorial Criteria will be recruited consecutively at the ALS/MND Clinics in Toronto, Canada and Boston, USA.<sup>16</sup> Individuals will be excluded if they have a history of other neurological conditions (eg, stroke) or head and neck cancer; take medications known to affect speech production; are unable to comply with the study protocol due to significant cognitive impairment (ie, Montreal Cognitive Assessment score  $< 18$ );<sup>17</sup> or fail a hearing screen in both ears (ie, 500–4000 Hz, thresholds  $> 40$  dB).

To ensure that the ALS-BDI is tested across a broad sampling of disease severity, we will recruit an equal number of patients who fall within four categories of bulbar symptom severity, based on their ALS Functional

Rating Scale - Revised (ALSFRS-R) bulbar subscore: normal bulbar function (ie, score of 12/12); mild (9–11/12); moderate (6–8/12); or severe (1–5/12).<sup>18</sup>

#### Inter-rater and test–retest reliability

To estimate the inter-rater reliability of each candidate item, two SLPs will rate each patient's performance on the ALS-BDI on the same visit. The order of SLP raters will be randomised, and they will be blinded to each other's ratings. The first administration will be conducted at the beginning of the visit, and the second will take place at the very end. To minimise effects of fatigue, participants will be given sufficient time to rest before the second ALS-BDI assessment.

To determine test–retest reliability, each participant will be reassessed using the ALS-BDI, within 1 week after the initial visit. The second visit will be scheduled approximately at the same time of day as the first visit, to mitigate any daytime variability in fatigue. Given the natural history of ALS, it is unlikely that there will be significant clinical change in the period of 2 weeks. However, on the second visit, the participants will answer whether they think their bulbar function has changed compared with the first assessment (ie, yes/no). Participants who report any perceived differences in bulbar symptoms compared with the initial visit will be excluded from the test–retest reliability calculations.

We will calculate reliability statistics for each item on the ALS-BDI using weighted Kappa, because individual items were scored ordinally. Items with weighted Kappa <0.6 on either inter-rater or test–retest reliability will be flagged for reduction.<sup>19</sup> We will also calculate the overall reliability of the total score of the ALS-BDI (ie, the sum of all items), using intra-class correlation coefficients (ICCs), considering raters as random effects (ICC formula 2,1).<sup>20 21</sup> ICCs of  $\geq 0.80$  will be set as the minimal acceptable standard for evaluating groups of patients, whereas values of  $\geq 0.90$  will be required for individual patient use.<sup>22</sup> Item reduction based on reliability will be completed prior to construct validation of individual items and the full ALS-BDI.

Data from the first 50 consecutive individuals diagnosed with ALS and presenting with a range of bulbar impairment will be used for the reliability analyses. This exact sample size was estimated using the approach by Kraemer and Korner,<sup>23</sup> whereby for a minimum ICC of 0.8 and a lower 95% CI of 0.7, 47 analysable participants should be required. We will increase this number if some patients fail to attend the second visit (test–retest) within 2 weeks of the initial visit recording, or if they report a change in clinical status during this time interval.

#### Construct validity

Items remaining following reliability assessment will be evaluated for construct validity, relative to gold-standard instrumental physiological measures of bulbar dysfunction. For this phase of the study, data from 100 participants with ALS will be recorded. On the day of ALS-BDI administration, each patient will complete a battery of

instrumental assessments and selected patient-reported outcomes indicating overall ALS and bulbar-related disability, namely the ALSFRS-R and the Center for Neurologic Study Bulbar Function Scale (CNS-BFS).<sup>18 24</sup>

Table 1 shows the detailed mapping between individual items of ALS-BDI and specific instrumental validators (ie, tasks and measures). The entire session will be audio-recorded at high resolution (44.1 kHz, 16-bit) using a unidirectional lapel microphone, with a fixed mouth-to-mic distance; the audio signal will be calibrated using a 1000 Hz tone at the start of each session to ensure accurate measures of signal intensity.<sup>25</sup> In addition to audio recordings of all speech tasks, the Phonatory Aerodynamic System (Pentax Medical) will be used to measure nasal airflow and oral pressure during speech and cough.<sup>25 26</sup> Tongue and lip/facial muscle strength will be measured using the Iowa Oral Performance Instrument (IOPI Medical LLC) during maximum pressure generation tasks.<sup>27 28</sup> Facial movements for speech and non-speech (oromotor) will be recorded using a three-dimensional video camera (Intel RealSense),<sup>29</sup> and tongue movement will be assessed using the Wave Speech System (NDI).<sup>30 31</sup> To validate swallow tasks, each participant will undergo a standard videofluoroscopic swallow study. The full validation protocol is estimated to take approximately 2 hours.

All speech tasks will be analysed using well-established spectral and timing measures (eg, voice onset time, acoustic vowel space, segment durations; see table 1 for examples).<sup>25</sup> The standard passage reading will be analysed using an automated Speech-Pause Analysis algorithm, extracting measures of phrase and pause durations.<sup>32</sup> In addition, participants will complete the Speech Intelligibility Test to quantify speech intelligibility (ie, % words heard correctly) and speaking rate (ie, words per minute) in sentences.<sup>33</sup> Measures of swallow safety (ie, penetration/aspiration),<sup>34</sup> efficiency (ie, residue)<sup>35</sup> and key physiological events (eg, pharyngeal constriction,<sup>36 37</sup> chewing duration, laryngeal vestibule closure)<sup>35</sup> will be obtained from a videofluoroscopic swallow study. Cough volume acceleration will be measured from the Phonatory Aerodynamic System waveform of voluntary cough.<sup>38 39</sup>

We will study the validity of each reliable item, as determined in the reliability assessment described above, in relation to its corresponding instrumental measure. All instrumental measures for this study are continuous, so we will calculate the correlations (Pearson's or Spearman's as appropriate) between each item and its instrumental equivalent. We have set an a priori correlation cut point of  $\geq 0.6$  for criterion validity; items that correlate with instrumental measures of  $>0.6$  will be flagged for inclusion.<sup>40</sup> Inter-item correlations will also be examined for redundancy. In a formative model, while low or absent correlations between items are not an issue, very high inter-item correlations ( $>0.9$ ) can indicate redundancy; pairs of items with correlations  $>0.9$  will be assessed for potential reduction. If additional items are eliminated, we will re-calculate the reliability coefficients for the final version of the ALS-BDI.

**Table 1** Validation procedures for ALS-BDI items

ALS-BDI item(s)	Validation method/instrumentation	Validation task(s)	Measurement(s)
Items 1,2: Overall Severity of Dysarthria and Reduced Speech Intelligibility	Audio recordings (44.1K, 16-bit resolution) Speech Intelligibility Test <sup>33</sup>	Sentences	Speech intelligibility (%words correct); <sup>33</sup> speaking rate (words per minute, WPM) <sup>33 52</sup>
Items 3,7: Tongue Weakness and Fasciculations and Lip/Face Weakness	Iowa Oral Performance Instrument (IOPI Medical LLC)	Max isometric pressure of tongue; max isometric lip pressure <sup>27 28</sup>	Peak pressure, in kilopascals (kPa) <sup>27</sup>
Items 5,6,8,9: Reduced Tongue Range of Motion (ROM) and Slow Movement; Reduced Lips ROM and Slow Movement	3D video camera (Intel RealSense) and Wave Speech System (NDI)	Max mouth opening, rapid movements of Smile-Pucker; tongue from side-to-side outside of the mouth <sup>53</sup>	Kinematic measures of lip, jaw and tongue movements, including range of motion and speed <sup>31 54–58</sup>
Item 12: Nasal Emissions	Phonatory Aerodynamic System, PAS (PENTAX Medical)	Syllable and sentence repetitions	Max nasal flow during pressure consonants <sup>54 59 60</sup>
Item 27: Hypernasality	Nasometer (PENTAX Medical)	Oral–nasal sentences	Nasalance distance <sup>61</sup>
Items 13–18 (Phonatory): Strained Voice; Breathiness; Roughness; Voice Breaks; Overall Dysphonia; and Inability to Elevate Pitch	Audio recordings (as above); Analysis of Dysphonia in Speech and Voice (PENTAX Medical)	Sustained phonation; CAPE-V <sup>62</sup> sentences; Pitch glide	Cepstral peak prominence (CPP); <sup>63</sup> Spectral energy ratio (L/H ratio); <sup>64</sup> Cepstral spectral index of dysphonia <sup>25 64</sup>
Items 19–22 (Respiratory): Loudness Decay; Reduced Loudness; Short Phrases; and Effortful Breathing	Audio recordings; Acoustic analysis using PRAAT and automatic Speech-Pause Analysis <sup>65</sup>	Standard passage reading <sup>66</sup>	Intensity (dB SPL) mean, range, variability; tone unit ratio; mean phrase duration; <sup>32</sup> # of pauses <sup>32 54 65 66</sup>
Items 23–26 (Prosody): Monopitch and/or Monoloudness; Excess and Equal Stress; Reduced Stress; and Overall Dysprosody	Audio recording (44.1K, 16-bit resolution); Post-hoc analysis using PRAAT	Standard passage reading <sup>66</sup>	Mean fundamental frequency; <sup>25</sup> Phonatory range (min/max F0); <sup>67</sup> Pitch variability (SD F0) <sup>25 54</sup>
Items 28–30 (Articulatory): Imprecise Articulation; Slow Articulation Rate; and Reduced Overall Speaking Rate	Audio recording (as above); Post-hoc analysis Speech-Pause Analysis <sup>65</sup>	Standard passage reading <sup>66</sup>	Acoustic vowel space; <sup>25</sup> Vowel and fricative durations; <sup>25</sup> Voice onset time; <sup>25</sup> Speaking rate (WPM) <sup>36 53</sup>
Items 31–34: Slow DDK Rate; Irregular DDK Rhythm; Slow/pataka/Rate; and Irregular/pataka/Rhythm	Audio recording (as above); Post-hoc analysis using a validated in-house algorithm <sup>68</sup>	Syllable repetitions (ie, 'puh', 'tuh', 'kuh', 'pataka') <sup>69</sup>	DDK rate (syl/second); <sup>55 69–72</sup> DDK regularity (SD of syllable duration) <sup>72</sup>
Item 35: Impaired Voluntary Cough	PAS (PENTAX Medical)	Voluntary cough <sup>38</sup>	Cough volume acceleration <sup>38</sup>
Items 36–37: Swallowing Difficulties 3oz Water Swallow; Increased Chewing Time	Videofluoroscopic Swallow Study, Toshiba Ultimax System MDX-8000A at 30 pulses per second; uncompressed image captured with TIMS 2000 DICOM system (Forest Imaging)	Swallows of standardised barium mixtures (Varibar 40% weight-to-volume) of the following order/texture: 2×5 mL thin liquid; 2× natural sip thin liquid; 2× sequential sip thin liquid; 2×5 mL extremely thick liquid; one bite size cookie	Penetration-Aspiration Scale (8-point); <sup>34</sup> Pharyngeal constriction; <sup>35–37</sup> Residue; <sup>35 73 74</sup> Time-to laryngeal vestibule closure; <sup>35 75</sup> Degree of laryngeal vestibule closure <sup>35 75</sup>
Items 4, 11: Tongue Atrophy and Jaw Jerk			Cannot be instrumentally assessed/validated

ALS-BDI, Amyotrophic Lateral Sclerosis-Bulbar Dysfunction Index; 3D, three-dimensional; DDK, Diadochokinetic rate task; L/H ratio, Low to High spectral ratio; SD, Standard deviation; SPL, Sound Pressure Level.

Once we have the final version of the ALS-BDI, we will study the construct validity of the entire tool. We will test two hypotheses for construct validity: (1) the ALS-BDI total score will be highly correlated ( $r \geq 0.7$ ) with ALSFRS-R bulbar subscore and the CNS-BFS, which measures bulbar-specific disability; (2) the correlation will be weaker ( $r = 0.4–0.6$ ) with the ALSFRS-R total score, which measures ALS disability as a whole and includes the effect of limb weakness and respiratory dysfunction. We will also compare the ALS-BDI scores between subgroups of patients divided by bulbar

function severity. We will classify patients as having normal, mild, moderate and severe bulbar dysfunction, based on their ALSFRS-R bulbar subscores.<sup>18</sup> We will compare the mean ALS-BDI scores on each bulbar severity group using analysis of variances. We hypothesise that the different severity groups will have statistically significant different mean ALS-BDI scores. Following the COSMIN guidelines, we will consider that the new scale has construct validity, if we confirm  $\geq 75\%$  of the predefined hypotheses.<sup>41</sup>

## Establishing responsiveness, MID, MDC and low disease state of ALS-BDI

For this analysis, patients will have a second full assessment 6 months after the initial visit. On the second visit, patients will undergo the ALS-BDI, instrumental testing (as described above), as well as the ALSFRS-R and CNS-BFS. Patients will also answer a patient impression-of-change questionnaire (PIC), to indicate if their speech, voice and/or swallowing have changed since the first assessment. The PIC will have four possible answers: 0=no change, 1=minimally worse, 2=much worse, 3=very much worse; this is a modification from the method by Juniper, that used seven options for improvement or worsening.<sup>42</sup> Considering the natural history of ALS, we do not expect patients to report improvement. We will calculate the difference in ALS-BDI scores between visits 1 and 2, as well as the change scores for the instrumental measures, the ALSFRS-R and CNS-BFS. There is no single way to determine responsiveness, so current guidelines recommend using a range of methods to document sensitivity to change including (1) assessing statistically significant change—such as in a clinical trial—by comparing mean effects with a t-test, and (2) assessing change that is meaningful for patients.<sup>41</sup>

We will estimate statistically significant change in four ways. First, we will compare the mean ALS-BDI scores at session 1 and session 2 through paired t-tests; we expect to see a statistically significant difference ( $p < 0.05$ ), indicating worsening of bulbar function. Second, we will calculate the magnitude of change through the standardised response mean (SRM), which is the ratio between the mean change score to the SD of said change score. We hypothesise that the SRM will be  $\geq 0.3$ . Third, we will study the correlations between the change in the ALS-BDI and the comparison measures, as a form of longitudinal validity.<sup>43</sup> We expect to find moderate correlations between the change scores ( $r = 0.4–0.7$ ). Fourth, we will study the efficiency to detect statistically significant change of the ALS-BDI compared with the instrumental measures, the ALSFRS-R bulbar score and CNS-BFS. For this, we will calculate the ratio of the paired t-test statistics, as follows:  $(t - statistic_{ALS-BDI} / t - statistic_{comparison\ measure})$ .<sup>44,45</sup>

A ratio  $> 1$  indicates that the ALS-BDI is more efficient than the comparison measure of interest; in other words, it is more sensitive with detecting a given effect size for the same sample size. We hypothesise that the ALS-BDI will be more efficient than the ALSFRS-R bulbar score and the CNS-BFS; we expect that the ALS-BDI subscores will have equal or slightly lower efficiency than the corresponding instrumental measures.

We will determine if the ALS-BDI detects change that is meaningful to patients by using the PIC as an anchor. To be meaningful from the patients' perspective, the ALS-BDI change-scores should correlate with the PIC. For instance, patients reporting to be much worse should have a higher change in score than those reporting moderate, mild or no change. We will use analysis of covariance to compare the mean change in scores among PIC groups, correcting for differences in baseline scores. We expect a

significant difference in change scores across PIC groups, with larger change scores in those patients reporting to be much worse.

To enhance the interpretation of the change in scores, we will estimate the MID for worsening. This is the *smallest* change in score that is meaningful for patients, and we will use the PIC category of 'a little worse' as the anchor of minimal worsening. This follows current recommendations of using anchor-based methods to determine the MID, as opposed to distribution-based methods.<sup>46</sup> We will estimate the MID at the individual level—that is, to classify one individual as having at least minimal worsening—using a receiver operator characteristic curve. Patients who have at least minimal worsening will be classified as 'changers.' The MID at the individual level is the point of highest sensitivity and specificity. We will also estimate the MID at the group level—that is, the smallest mean change in scores that would be meaningful in a group, such as in a clinical trial. For this, we will calculate the mean change in ALS-BDI scores in the group of patients that reported being 'a little worse';<sup>42</sup> we will also estimate the 75% percentile for potential misclassification bias. For the MID estimates to be interpretable, they should be above error of measurement. Therefore, we will calculate the MDC, which is the smallest change between assessments that is very likely above error of measurement. The MDC (with 95% CI) is calculated using the test–retest reliability statistic as follows:

$$1.96 \times \sqrt{2} \times (SD \times \sqrt{1 - ICC}),$$

where *SD* is the SD of the sample, and *ICC* is the test–retest reliability coefficient.<sup>46</sup> For the MID scores to be meaningful, they should be above error; the MID should be larger than the MDC.

We will use the patient anchor to determine an ALS-BDI score threshold for the low disease activity states.<sup>47</sup> This metric is similar to the patient acceptable symptom states (PASS), which reflect the threshold on a patient-reported outcome where people feel generally well. At the same time, it differs from the MID, described above, as it is not a measure of change. To determine this threshold, we will ask patients on both visits, 'Considering all the difficulties on speaking and swallowing that you experience due to ALS, if you were to remain at this same level of function for the following months, would you consider your speaking or swallowing function satisfactory?'. This is a modification of PASS/low disease state questions that have been used in other diseases to help with interpretability of the scale scores.<sup>47–50</sup>

## PATIENT AND PUBLIC INVOLVEMENT

We envision the end users of ALS-BDI to be clinicians delivering care to patients with ALS in the context of a multidisciplinary ALS clinic. As such, patients have not been directly involved in the development of the tool but were engaged in the trial assessments with the tool in Cycle 1 of its development. Patients have not been directly involved in the design of the reliability or validation studies presented here



(Cycles 2 and 3). Patients were involved in the pilot of the validation study to ensure that the study is well tolerated and optimised for the abilities of those patients with a more advanced presentation of ALS. Patients will be involved in conduct of the study. The results will be disseminated to the patient participants through sharing study publications and poster presentations with those who will express interest in receiving this information. The results will also be presented to patient groups.

## DISCUSSION

Bulbar motor dysfunction is among the most devastating consequences of ALS, a fast-progressing motor neuron disease. Although the assessment of bulbar dysfunction is a core component of the ALS clinical examination, a well-designed, easy to perform and efficient tool for its assessment is lacking.<sup>14 15</sup> This work aims to fill this gap by developing and validating a clinician-administered tool, the ALS-BDI, to serve as a primary bulbar assessment tool in a clinic and in clinical trials. The innovation of this work is not only in its uniqueness and methodological rigour, but the state-of-the-art instrumentation that will be used for its validation.

The ALS-BDI is being designed for SLPs because of their specialised training in the assessment of voice, speech and swallowing impairments. The tool's administration is consistent with clinical assessment procedures used by SLPs, which are based on clinician ratings of speech, swallowing and oral structure and function. Although clinician ratings are the foundation of an SLP assessment, the efficacy of these measures has rarely been tested using rigorous experimental designs.<sup>51</sup>

Because the tool will be based on the clinician's ratings, particular attention will be paid to creating a training module for its standardised administration. Further, the tool's efficiency will be addressed by careful evaluation and selection of the items. A useful bulbar assessment will need to be brief, such that it can be administered from start to finish within 10–15 min, even for patients who are severely affected. A brief assessment will reduce the burden on the patients, who are prone to fatigue and are frequently evaluated by multiple practitioners within a single clinic visit. Standardisation and efficiency will be achieved by developing a tool with a minimal set of test items—removed based on reliability, validity and responsiveness evaluations—that selectively target key clinical characteristics associated with bulbar ALS and is accompanied by explicit instructions for administration.

## NEXT STEPS AND IMPACT

The proposed study was placed on hold between March 2020 and March 2021 due to the global COVID-19 pandemic. During this time, the tool was redesigned for an online/remote administration, and the summary of the remote version of the tool is forthcoming. Meanwhile,

the return to in-person study administration outlined in the current protocol is in progress.

To the best of our knowledge, the ALS-BDI will be the first standardised and psychometrically validated clinician-administered bulbar assessment tool. The ALS-BDI aims to meet the key requirements for the development of new clinical assessment tools: (1) supported by the multi-disciplinary ALS expert consensus; (2) developed with considerations of efficiency, specificity and standardisation; and (3) established evidence of strong psychometric properties. The long-term impacts of the ALS-BDI includes improving detection of bulbar ALS, expediting diagnosis, improving clinical decision-making and accelerating ALS clinical trials and drug discovery. Because the tool is consistent with current assessment practices in SLP, which are based on clinician ratings of speech, swallowing and oral structure and function, it is likely to have uptake within clinical practice settings focused on ALS.

## ETHICS AND DISSEMINATION

This study has been reviewed and approved by research ethics boards at both data collection sites: Sunnybrook Health Science Centre, primary (Toronto, Canada; ID3080) and Mass General Brigham (Boston, USA; #2013P001746). Prior to participation in all elements of the study, the participants sign the informed consent in accordance with the Declaration of Helsinki. Once validated, the ALS-BDI will be disseminated to key stakeholders at clinical ALS and SLP conferences (eg, American Speech Hearing Association (ASHA) Convention, International Symposium on ALS/MND). The tool and any relevant training materials will be made publicly available to SLPs and neurologists who provide ALS care. The ALS-BDI will also be launched as a validated outcome tool for clinical research trials in ALS.

## Author affiliations

<sup>1</sup>Department of Speech-Language Pathology, University of Toronto Temerty Faculty of Medicine, Toronto, Ontario, Canada

<sup>2</sup>Hurvitz Brain Sciences Research Program, Sunnybrook Research Institute, Toronto, Ontario, Canada

<sup>3</sup>Department of Medicine, University Health Network, Toronto, Ontario, Canada

<sup>4</sup>Division of Neurology, Department of Medicine, University of Toronto Temerty Faculty of Medicine, Toronto, Ontario, Canada

<sup>5</sup>Rehabilitation Sciences Institute, University of Toronto Temerty Faculty of Medicine, Toronto, Ontario, Canada

<sup>6</sup>Department of Otolaryngology – Head and Neck Surgery, University of Toronto Temerty Faculty of Medicine, Toronto, Ontario, Canada

<sup>7</sup>Division of Neurology, Department of Medicine, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada

<sup>8</sup>Neurology Associates P.C, Lincoln, Nebraska, USA

<sup>9</sup>Sean M. Healey and AMG Center for ALS, Massachusetts General Hospital, Boston, Massachusetts, USA

<sup>10</sup>Department of Communication Sciences and Disorders, MGH Institute of Health Professions, Boston, Massachusetts, USA

<sup>11</sup>Speech and Hearing Biosciences and Technology, Harvard University, Cambridge, Massachusetts, USA

**Twitter** Yana Yunusova @YunusovaLab

**Acknowledgments** We would like to thank all members of the Northeastern ALS Consortium Bulbar Subcommittee, who participated in focus groups, as well as

the speech–language pathologists and neurologists who participated in online surveys and cognitive interviews, for providing important input into the domain conceptualisation and informing design and drafting of the Amyotrophic Lateral Sclerosis – Bulbar Dysfunction Index. We also extend appreciation to Madhura Kulkarni and Brian Richburg for managing the study, Diego Guarin and Andrea Bandini for designing key tools for the validation protocol and Amanda Martino and Reeman Marzouqah for pilot testing the validation protocol.

**Contributors** YY, CB, RM, AA, LZ, JDB, GLP and JRG conceived the study and prepared the funding application. All authors contributed to the design of the study protocol. AH and AW coordinated pilot testing of the validation protocol. All authors reviewed feedback and came to consensus on tool design. YY, AW, CB and JRG wrote the manuscript. All authors read, edited and approved the final manuscript.

**Funding** This work has been supported by National Institute on Deafness and Other Communication Disorders (NIDCD) grant number R01DC017291.

**Competing interests** None declared.

**Patient and public involvement** Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

**Patient consent for publication** Not applicable.

**Provenance and peer review** Not commissioned; peer reviewed for ethical and funding approval prior to submission.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

#### ORCID iD

Yana Yunusova <http://orcid.org/0000-0002-2353-2275>

#### REFERENCES

- Armon C, Graves MC, Moses D, *et al*. Linear estimates of disease progression predict survival in patients with amyotrophic lateral sclerosis. *Muscle Nerve* 2000;23:874–82.
- Abrahams S, Goldstein LH, Suckling J, *et al*. Frontotemporal white matter changes in amyotrophic lateral sclerosis. *J Neurol* 2005;252:321–31.
- Abrahams S, Leigh PN, Goldstein LH. Cognitive change in ALS: a prospective study. *Neurology* 2005;64:1222–6.
- Goldstein LH, Abrahams S. Changes in cognition and behaviour in amyotrophic lateral sclerosis: nature of impairment and implications for assessment. *Lancet Neurol* 2013;12:368–80.
- Haverkamp LJ, Appel V, Appel SH. Natural history of amyotrophic lateral sclerosis in a database population. validation of a scoring system and a model for survival prediction. *Brain* 1995;118 (Pt 3):707–19.
- Turner MR, Scaber J, Goodfellow JA, *et al*. The diagnostic pathway and prognosis in bulbar-onset amyotrophic lateral sclerosis. *J Neurol Sci* 2010;294:81–5.
- Chiò A, Logroscino G, Hardiman O, *et al*. Prognostic factors in ALS: a critical review. *Amyotroph Lateral Scler* 2009;10:310–23.
- Hecht M, Hillemacher T, Gräsel E, *et al*. Subjective experience and coping in ALS. *Amyotroph Lateral Scler Other Motor Neuron Disord* 2002;3:225–31.
- Plowman EK, Tabor LC, Wymer J, *et al*. The evaluation of bulbar dysfunction in amyotrophic lateral sclerosis: survey of clinical practice patterns in the United States. *Amyotroph Lateral Scler Frontotemporal Degener* 2017;18:351–7.
- Kirshner B, Guyatt G. A methodological framework for assessing health indices. *J Chronic Dis* 1985;38:27–36.
- Mokkink LB, Prinsen CAC, Bouter LM, *et al*. The consensus-based standards for the selection of health measurement instruments (COSMIN) and how to select an outcome measurement instrument. *Braz J Phys Ther* 2016;20:105–13.
- Diamantopoulos A, Siguaw JA. Formative versus reflective indicators in organizational measure development: a comparison and empirical illustration. *Br J Nurs* 2006;17:263–82.
- Yunusova Y, Waito A, Barnett-Tapia C. A novel ALS bulbar dysfunction index (ALS-BDI): establishing face and content validity. *20th Biennial Conference on Motor Speech: conference proceedings*, Santa Barbara, California, USA: Madonna Rehabilitation Hospitals, 2020:19–23.
- Pattee GL, Plowman EK, Brooks BR, *et al*. Best practices protocol for the evaluation of bulbar dysfunction: summary recommendations from the NEALS bulbar Subcommittee symposium. *Amyotroph Lateral Scler Frontotemporal Degener* 2018;19:311–2.
- Yunusova Y, Plowman EK, Green JR, *et al*. Clinical measures of bulbar dysfunction in ALS. *Front Neurol* 2019;10:106.
- Brooks BR, Miller RG, Swash M, *et al*. El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis. *Amyotroph Lateral Scler Other Motor Neuron Disord* 2000;1:293–9.
- Nasreddine ZS, Phillips NA, Bédirian V, *et al*. The Montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc* 2005;53:695–9.
- Cedarbaum JM, Stambler N, Malta E, *et al*. The ALSFRS-R: a revised ALS functional rating scale that incorporates assessments of respiratory function. BDNF ALS Study Group (phase III). *J Neurol Sci* 1999;169:13–21.
- Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther* 2005;85:257–68.
- Koo TK, MY L. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 2016;15(2):155–63. 10.1016/j.jcm.2016.02.012 [published Online first: 2016/06/23] Erratum in. *J Chiropr Med* 2017;16:346.
- Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979;86:420–8.
- Nunnally JC. *Psychometric theory*. New York, NY: McGraw-Hill, 1967.
- Kraemer HC, Korner AF. Statistical alternatives in assessing reliability, consistency, and individual differences for quantitative measures: application to behavioral measures of neonates. *Psychol Bull* 1976;83:914–21.
- Smith RA, Macklin EA, Myers KJ. Assessment of bulbar function in amyotrophic lateral sclerosis: validation of a self-report scale (Center for Neurologic Study Bulbar Function Scale). *Eur J Neurol* 2018;25(7):907–e66. 10.1111/ene.13638 [published Online First: 2018/03/27] Erratum in. *Eur J Neurol* 2018;25:1303.
- Ludlow CL, Kent RD, Gray LC, *et al*. *Speech, and swallowing in the clinic and laboratory*. San Diego, CA: Plural Publishing, Inc, 2018.
- Tabor-Gray LC, Gallestagui A, Vasilopoulos T, *et al*. Characteristics of impaired voluntary cough function in individuals with amyotrophic lateral sclerosis. *Amyotroph Lateral Scler Frontotemporal Degener* 2019;20:37–42.
- Youmans SR, Stierwalt JAG. Measures of tongue function related to normal swallowing. *Dysphagia* 2006;21:102–11.
- Adams V, Mathisen B, Baines S, *et al*. A systematic review and meta-analysis of measurements of tongue and hand strength and endurance using the Iowa oral performance instrument (IOP). *Dysphagia* 2013;28:350–69.
- Bandini A, Green JR, Taati B. Automatic detection of amyotrophic lateral sclerosis (ALS) from video-based analysis of facial movements: speech and non-speech tasks. *13th IEEE International Conference on Automatic Face and Gesture Recognition, FG 2018: conference proceedings*, 2018:150–7.
- Berry JJ. Accuracy of the NDI wave speech research system. *J Speech Lang Hear Res* 2011;54:1295–301.
- Shellikeri S, Green JR, Kulkarni M, *et al*. Speech movement measures as markers of bulbar disease in amyotrophic lateral sclerosis. *J Speech Lang Hear Res* 2016;59:887–99.
- Barnett C, Green JR, Marzouqah R, *et al*. Reliability and validity of speech & pause measures during passage reading in ALS. *Amyotroph Lateral Scler Frontotemporal Degener* 2020;21:42–50.
- Yorkston KM, Beukelman DR, Hakel M. *Speech intelligibility test for windows. communication disorders software*, 1996.
- Rosenbek JC, Robbins JA, Roecker EB, *et al*. A penetration-aspiration scale. *Dysphagia* 1996;11:93–8.
- Waito AA, Plowman EK, Barbon CEA, *et al*. A cross-sectional, quantitative videofluoroscopic analysis of swallowing physiology and function in individuals with amyotrophic lateral sclerosis. *J Speech Lang Hear Res* 2020;63:948–62.
- Waito AA, Tabor-Gray LC, Steele CM, *et al*. Reduced pharyngeal constriction is associated with impaired swallowing efficiency in amyotrophic lateral sclerosis (ALS). *Neurogastroenterol Motil* 2018;30:e13450.
- Leonard R, Rees CJ, Belafsky P, *et al*. Fluoroscopic surrogate for pharyngeal strength: the pharyngeal constriction ratio (PCR). *Dysphagia* 2011;26:13–17.
- Plowman EK, Watts SA, Robison R, *et al*. Voluntary cough airflow differentiates safe versus unsafe swallowing in amyotrophic lateral sclerosis. *Dysphagia* 2016;31:383–90.
- Tabor-Gray L, Vasilopoulos T, Plowman EK. Differences in voluntary and reflexive cough strength in individuals with amyotrophic lateral sclerosis and healthy adults. *Muscle Nerve* 2020;62:597–600.

- 40 de Vet HCW, Terwee CB, Mokkink LB. *Measurement in medicine: a practical guide*. Cambridge: Cambridge University Press, 2011.
- 41 Terwee CB, Dekker FW, Wiersinga WM, *et al*. On assessing responsiveness of health-related quality of life instruments: guidelines for instrument evaluation. *Qual Life Res* 2003;12:349–62.
- 42 Juniper EF, Guyatt GH, Willan A, *et al*. Determining a minimal important change in a disease-specific quality of life questionnaire. *J Clin Epidemiol* 1994;47:81–7.
- 43 Hedeker DR, Gibbons RD. *Longitudinal data analysis*. Hoboken, N.J: Wiley-Interscience, 2006.
- 44 Ruoppolo G, Schettino I, Frasca V, *et al*. Dysphagia in amyotrophic lateral sclerosis: prevalence and clinical findings. *Acta Neurol Scand* 2013;128:397–401.
- 45 Liang MH, Larson MG, Cullen KE, *et al*. Comparative measurement efficiency and sensitivity of five health status instruments for arthritis research. *Arthritis Rheum* 1985;28:542–7.
- 46 King MT. A point of minimal important difference (mid): a critique of terminology and methods. *Expert Rev Pharmacoecon Outcomes Res* 2011;11:171–84.
- 47 Wells G, Boers M, Shea B, *et al*. MCID/Low disease activity state workshop: low disease activity state in rheumatoid arthritis. *J Rheumatol* 2003;30:1110–1.
- 48 Franklyn K, Lau CS, Navarra SV, *et al*. Definition and initial validation of a lupus low disease activity state (LLDAS). *Ann Rheum Dis* 2016;75:1615–21.
- 49 Tubach F, Ravaud P, Baron G, *et al*. Evaluation of clinically relevant states in patient reported outcomes in knee and hip osteoarthritis: the patient acceptable symptom state. *Ann Rheum Dis* 2005;64:34–7.
- 50 Puyraimond-Zemmour D, Etcheto A, Fautrel B, *et al*. Associations between five important domains of health and the patient acceptable symptom state in rheumatoid arthritis and psoriatic arthritis: a cross-sectional study of 977 patients. *Arthritis Care Res* 2017;69:1504–9.
- 51 Kent RD. Hearing and believing: some limits to the auditory-perceptual assessment of speech and voice disorders. *Am J Speech Lang Pathol* 1996;5:7–23.
- 52 Stipancic KL, Yunusova Y, Berry JD, *et al*. Minimally detectable change and minimal clinically important difference of a decline in sentence intelligibility and speaking rate for individuals with amyotrophic lateral sclerosis. *J Speech Lang Hear Res* 2018;61:2757–71.
- 53 KO SL, Ruscello DM. *Oral speech mechanism screening examination (OSMSE)*. Baltimore, MD: University Park Press, 1981.
- 54 Rong P, Yunusova Y, Wang J, *et al*. Predicting early bulbar decline in amyotrophic lateral sclerosis: a speech subsystem approach. *Behav Neurol* 2015;2015:183027–11.
- 55 Rong P, Yunusova Y, Richburg B, *et al*. Automatic extraction of abnormal lip movement features from the alternating motion rate task in amyotrophic lateral sclerosis. *Int J Speech Lang Pathol* 2018;20:610–23.
- 56 Yunusova Y, Green JR, Lindstrom MJ, *et al*. Kinematics of disease progression in bulbar ALS. *J Commun Disord* 2010;43:6–20.
- 57 Bandini A, Green JR, Wang J, *et al*. Kinematic features of jaw and lips distinguish symptomatic from presymptomatic stages of bulbar decline in amyotrophic lateral sclerosis. *J Speech Lang Hear Res* 2018;61:1118–29.
- 58 Bandini A, Green JR, Zinman L. Classification of bulbar ALS from kinematic features of the jaw and lips: towards computer-mediated assessment. *18th Annual Conference of the International Speech Communication Association (INTERSPEECH, 2017): Situated Interaction: conference proceedings*, 2017:2018.
- 59 Rong P, Yunusova Y, Green JR. Differential effects of velopharyngeal dysfunction on speech intelligibility during early and late stages of amyotrophic lateral sclerosis. *17th Annual Conference of the International Speech Communication Association (INTERSPEECH, 2016): Understanding Speech Processing in Humans and Machines: conference proceedings*, 2016:2016.
- 60 Gauster A, Yunusova Y, Zajac D. The effect of speaking rate on velopharyngeal function in healthy speakers. *Clin Linguist Phon* 2010;24:576–88.
- 61 Bressmann T, Sader R, Whitehill TL, *et al*. Nasalance distance and ratio: two new measures. *Cleft Palate Craniofac J* 2000;37:248–56.
- 62 Kempster GB, Gerratt BR, Verdolini Abbott K, *et al*. Consensus auditory-perceptual evaluation of voice: development of a standardized clinical protocol. *Am J Speech Lang Pathol* 2009;18:124–32.
- 63 Heman-Ackah YD, Heuer RJ, Michael DD, *et al*. Cepstral peak prominence: a more reliable measure of dysphonia. *Ann Otol Rhinol Laryngol* 2003;112:324–33.
- 64 Awan SN, Solomon NP, Helou LB, *et al*. Spectral-cepstral estimation of dysphonia severity: external validation. *Ann Otol Rhinol Laryngol* 2013;122:40–8.
- 65 Green JR, Beukelman DR, Ball LJ. Algorithmic estimation of pauses in extended speech samples of dysarthric and typical speech. *J Med Speech Lang Pathol* 2004;12:149–54.
- 66 Yunusova Y, Green JR, Wang J. A protocol for comprehensive assessment of bulbar dysfunction in amyotrophic lateral sclerosis (ALS). *J Vis Exp* 2010;48.
- 67 Yunusova Y, Graham NL, Shellikeri S, *et al*. Profiling speech and pausing in amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD). *PLoS One* 2016;11:e0147573.
- 68 Tanchip C, Yunusova Y, Waito AA. A complexity-based approach for automatic diadochokinesis analysis in amyotrophic lateral sclerosis (ALS). *20th Biennial Biennial Conference on Motor Speech: conference proceedings*, Santa Barbara, California, USA: Madonna Rehabilitation Hospitals, 2020.
- 69 Duffy JR. *Motor speech disorders: substrates, differential diagnosis, and management*. St. Louis: Mosby, 2005.
- 70 Nishio S, Niimi M. Changes over time in dysarthric patients with amyotrophic lateral sclerosis (ALS): a study of changes in speaking rate and maximum repetition rate (Mrr). *Clin Linguist Phon* 2000;14:485–97.
- 71 Nishio M, Niimi S. Comparison of speaking rate, articulation rate and alternating motion rate in dysarthric speakers. *Folia Phoniatr Logop* 2006;58:114–31.
- 72 Wang Y-T, Kent RD, Duffy JR, *et al*. Analysis of diadochokinesis in ataxic dysarthria using the motor speech profile program. *Folia Phoniatr Logop* 2009;61:1–11.
- 73 Hutcheson KA, Barrow MP, Barringer DA, *et al*. Dynamic imaging grade of swallowing toxicity (digest): scale development and validation. *Cancer* 2017;123:62–70.
- 74 Steele CM, Peladeau-Pigeon M, Nagy A, *et al*. Measurement of pharyngeal residue from lateral view videofluoroscopic images. *J Speech Lang Hear Res* 2020;63:1404–15.
- 75 Steele CM, Peladeau-Pigeon M, Barbon CAE, *et al*. Reference values for healthy swallowing across the range from thin to extremely thick liquids. *J Speech Lang Hear Res* 2019;62:1–26.