Development of an Accurate Bedside Swallowing Evaluation Decision Tree Algorithm for Detecting Aspiration in Acute Respiratory Failure Survivors

Marc Moss, MD; S. David White, MA; Heather Warner, PhD; Daniel Dvorkin, PhD; Daniel Fink, MD; Stephanie Gomez-Taborda, BS; Carrie Higgins, BSN; Gintas P. Krisciunas, MPH, MA; Joseph E. Levitt, MD; Jeffrey McKeenan, MS; Edel McNally, MS; Alix Rubio, BS; Rebecca Scheel, MS; Jonathan M. Siner, MD; Rosemary Vojnik, BS; and Susan E. Langmore, PhD

BACKGROUND: The bedside swallowing evaluation (BSE) is an assessment of swallowing function and airway safety during swallowing. After extubation, the BSE often is used to identify the risk of aspiration in acute respiratory failure (ARF) survivors.

RESEARCH QUESTION: We conducted a multicenter prospective study of ARF survivors to determine the accuracy of the BSE and to develop a decision tree algorithm to identify aspiration risk.

STUDY DESIGN AND METHODS: Patients extubated after ≥ 48 hours of mechanical ventilation were eligible. Study procedures included the BSE followed by a gold standard evaluation, the flexible endoscopic evaluation of swallowing (FEES).

RESULTS: Overall, 213 patients were included in the final analysis. Median time from extubation to BSE was 25 hours (interquartile range, 21-45 hours). The FEES was completed 1 hour after the BSE (interquartile range, 0.5-2 hours). A total of 33% (70/213; 95% CI, 26.6%-39.2%) of patients aspirated on at least one FEES bolus consistency test. Thin liquids were the most commonly aspirated consistency: 27% (54/197; 95% CI, 21%-34%). The BSE detected any aspiration with an accuracy of 52% (95% CI, 45%-58%), a sensitivity of 83% (95% CI, 74%-92%), and negative predictive value (NPV) of 81% (95% CI, 72%-91%). Using recursive partitioning analyses, a five-variable BSE-based decision tree algorithm was developed that improved the detection of aspiration with an accuracy of 81% (95% CI, 75%-87%), sensitivity of 95% (95% CI, 90%-98%), and NPV of 97% (95% CI, 95%-99%).

INTERPRETATION: The BSE demonstrates variable accuracy to identify patients at high risk for aspiration. Our decision tree algorithm may enhance the BSE and may be used to identify patients at high risk for aspiration, yet requires further validation.

TRIAL REGISTRY: ClinicalTrials.gov; No.: NCT02363686; URL: www.clinicaltrials.gov; CHEST 2020; 158(5):1923-1933

KEY WORDS: acute respiratory failure; aspiration; dysphagia; intubation; speech language pathologist

FOR EDITORIAL COMMENT, SEE PAGE 1806

ABBREVIATIONS: 3-WST = 3-oz water swallow test; APACHE = Acute Physiology And Chronic Health Evaluation; ARF = acute respiratory failure; BSE = bedside swallowing evaluation; FEES = flexible endoscopic evaluation of swallowing; IDDSI = International Dysphagia Diet Standardization Initiative; NPV = negative predictive value; PAS =

penetration-aspiration scale; PPV = positive predictive value; SLP = speech language pathologist; VFSS = videofluoroscopic swallow study

AFFILIATIONS: From the Division of Pulmonary Sciences and Critical Care Medicine (Drs Moss and Dvorkin; Ms Higgins; and Mr McKeenan) and the Department of Otolaryngology (Dr Fink),

Take-home Point

We conducted a multicenter prospective study of acute respiratory failure survivors to determine the accuracy of the speech language pathologist's bedside evaluation (BSE) and to develop a decision tree algorithm to identify aspiration risk. Using recursive partitioning analyses, a five-variable BSE-based decision tree algorithm was developed that improved the detection of aspiration. This decision tree algorithm may be used to identify patients at high risk for aspiration, yet requires further validation.

For the 455,000 acute respiratory failure (ARF) survivors each year, aspiration is a devastating complication that can develop after starting oral nutrition.^{1,2} Occurring in as many as 44% of ARF survivors, aspiration is associated with many deleterious consequences, including pneumonia, percutaneous feeding tube placement, long-term care facility admission, and increased hospital mortality.³⁻⁶

Although aspiration can develop before starting oral nutrition, the risk of aspiration can be exacerbated by feeding and must be balanced against inappropriately delaying the initiation of oral nutrition. For ARF survivors who already are nutritionally deficient, delayed oral nutrition causes further nutritional abnormalities while increasing caregiver burden, patient dissatisfaction, and health-related costs.⁷⁻¹⁰ Therefore,

improving the ability to diagnose aspiration accurately in ARF survivors could reduce the frequency and severity of aspiration, could result in a more rapid diet advancement for those patients who are not aspirating, could limit complications associated with delaying oral nutrition, could cause more appropriate diet recommendations, and could initiate appropriate rehabilitation techniques for those patients who are aspirating.¹¹⁻¹³

Speech language pathologists (SLPs) usually determine when ARF survivors can resume oral feeding.¹⁴ Based on a national survey, most SLPs rely primarily on the bedside swallow evaluation (BSE) to provide feeding recommendations.¹⁴ The BSE consists of a comprehensive history and detailed physical examination of the structure and function of the oropharyngeal musculature (including sensory and motor cranial nerve examination), followed by assessments of the patient's ability to swallow foods and liquids of different volumes and consistencies. Previous meta-analyses and systematic reviews demonstrated that the BSE is relatively inaccurate in detecting aspiration when compared with gold standard instrumental examinations such as the flexible endoscopic evaluation of swallow (FEES) examination or the videofluoroscopic swallow study (VFSS).¹⁵⁻²⁸ However most of these studies included patients with primarily neurological disorders and excluded ARF survivors.²⁷ Therefore, we conducted a large multicenter prospective study to assess the accuracy of the BSE for FEES-detected aspiration and dietary recommendations in ARF survivors. In addition, we sought to develop a streamlined BSE-based clinical decision tree algorithm to detect aspiration more accurately in these patients.

Methods

Study Design

We prospectively enrolled patients at four university-based tertiary medical centers. Screening occurred only during weekdays and when a second SLP was available to perform the FEES examination. Two of the sites (Yale and Stanford) were added to the study and enrolled patients only for the final 13 and 11 months, respectively. All consecutive patients were considered for study enrollment before a speech language pathology consult was considered by the primary team and regardless of whether the team did so. Patients were eligible if they were at least 18 years of age and were admitted to an ICU with ARF, defined as requiring mechanical ventilation with an endotracheal tube for longer than 48 hours. Exclusion criteria included contraindications to enteral nutrition; preexisting or acute primary, central, or peripheral neuromuscular disorder; preexisting

University of Colorado School of Medicine and the University of Colorado Denver Rehabilitation Therapy Services (Mr White), University of Colorado Hospital, Aurora, CO; the Section of Otolaryngology, Department of Surgery (Dr Warner), and the Section of Pulmonary, Critical Care, and Sleep Medicine (Dr Siner), Yale School of Medicine, and the Department of Communication Disorders (Dr Warner), Southern Connecticut State University, New Haven, CT; The Bioinformatics CRO, Inc. (Dr Dvorkin), Niceville, FL; Department of Otolaryngology (Mss Gomez-Taborda, McNallay, and Rubio; Mr Krisciunas; and Dr Langmore), Boston Medical Center, Boston, MA; the Department of Otolaryngology (Mr Krisciunas and Dr Langmore), Boston University School of Medicine, Boston, MA; the Division of Pulmonary and Critical Care (Dr Levitt and Ms Vojnik), Stanford University, Stanford, CA; the Sargent College of Health and Rehabilitation Sciences (Mss McNally and Scheel; and Dr Langmore), Boston University, Boston, MA; and Division of Speech Language Pathology (Ms Scheel), Massachusetts General Hospital, Boston, MA.

FUNDING/SUPPORT: M. M. is supported by National Institutes of Health grants [K24 HL 089223 and R21 NR 015886].

CORRESPONDENCE TO: Marc Moss, MD, Division of Pulmonary Sciences and Critical Care Medicine, University of Colorado School of Medicine, AMC, RC2, C-272, 12700 E 19th Ave, Aurora, CO 80045; e-mail: marc.moss@CUAnschutz.edu

Copyright © 2020 American College of Chest Physicians. Published by Elsevier Inc. All rights reserved.

DOI: <https://doi.org/10.1016/j.chest.2020.07.051>

history of dysphagia, head and neck cancer, or related surgery; presence of a tracheostomy; coagulopathy resulting in uncontrolled nasal or pharyngeal bleeding; and altered mentation (defined as positive Confusion Assessment Method for the ICU results) prohibiting compliance with the BSE and FEES. Patients were followed up for up to 72 hours after meeting the inclusion criteria to determine if the altered mentation resolved. If the altered mentation did resolve, patients were approached for possible enrollment. The type of intubation (nasal or oral), type of endotracheal tube, how the cuff pressure was managed, and placement of oral or nasal gastric tubes was left to the discretion of the primary team. The patient or surrogate decision maker signed an informed consent for participation. After enrollment, chart review was performed and basic demographic information, Acute Physiology And Chronic Health Evaluation (APACHE) II score, and Charlson Comorbidity Index were collected using data from the time of study enrollment. This study was approved by the institutional review board at each of the four centers (see [e-Appendix 1](#) for institutional review board approval numbers).

Bedside Swallowing Evaluation

Experienced SLPs performed the BSE evaluation according to a standardized protocol. Initially, the SLP reviewed the patient's medical record and performed a physical examination, starting with their judgement of lip movement and seal, tongue movement and strength, volitional cough strength, dentition, voice quality, and amount of assistance required with feeding. The patient then was seated as upright as possible. Before bolus administration, the SLP recorded the patient's oxygen saturation level and respiratory rate.

Then, the SLP administered five standard consistencies: ice chips (International Dysphagia Diet Standardization Initiative [IDDSI] consistency, starting at 7 and then transition to 0), nectar thin liquids (prethickened apple juice; IDDSI consistency, 2), pureed solids (applesauce; IDDSI consistency, 4), thin liquids (water; IDDSI consistency, 0), and solids (graham cracker; IDDSI consistency, 7) in successive boluses that increased in size.²⁹⁻³² [e-Table 1](#) outlines the order, type, and quantity of each bolus. All boluses were administered unless considered unsafe by the treating SLP. This safety determination was left to the discretion of the SLP based on the patient response to the prior bolus. If clinical signs of aspiration were evident with a smaller bolus of one consistency, the remaining boluses of that consistency could be skipped. The smallest bolus of the next consistency then would be administered. Additional descriptions of SLPs' prior clinical experience and the BSE examination are included in [e-Appendix 1](#).

The SLP watched for five predefined clinical signs concerning for aspiration after each bolus: cough, throat clearing, change in vocal quality, wet breath sounds, or stridor. Based on the results of the BSE, the SLP determined whether the patient was a high aspiration risk for any of the five consistencies and then made clinically appropriate liquid and solid dietary National Dysphagia Diet recommendations.^{29,31,32}

FEES Examination

Within 4 hours after the BSE, a second independent blinded SLP performed the FEES examination. Both fiber-optic and distal chip endoscopes were used across study sites. At the discretion of the SLP

Figure 1 – Patient screening, enrollment, and follow-up. BSE = bedside swallowing evaluation; FEES = flexible endoscopic evaluation of swallowing.

TABLE 1] Patient Demographics (N = 213)

Category	Distribution
Site, No. (%)	
Boston University Medical Campus	97 (46)
University of Colorado Hospital	75 (35)
Stanford Hospital	21 (10)
Yale New Haven Hospital	20 (9)
Patient age, median (IQR), y	57 (47-66)
Male sex, No. (%)	131 (62)
Race, No. (%)	
White	117 (55)
Black	54 (25)
Other	42 (20)
Primary service, No. (%)	
Medical ICU	147 (69)
Cardiac ICU	45 (21)
Surgical ICU	20 (9)
Burn ICU	1 (1)
Comorbidity, No. (%)	
COPD	56 (26)
Diabetes	53 (25)
Congestive heart failure	46 (22)
Myocardial infarction	36 (17)
Liver disease	23 (11)
Peripheral vascular disease	19 (9)
Cerebral vascular disease	18 (8)
Dementia	7 (3)
Rheumatology disease	5 (2)
Peptic ulcer disease	5 (2)
Oxygen therapy at the time of evaluation (n = 194)	
Room air	48 (25)
Nasal cannula	117 (60)
Heat high flow or other oxygen delivery systems	29 (15)
Oxygen saturation at the time of evaluation, median (IQR), %	96 (94-98)
Time from admission to intubation, median (IQR), h	8 (0-48)
Duration of intubation, median (IQR), h	126 (72-206)
Endotracheal tube size, No. (%), mm	
≥ 8.0	102 (49)
7.5	74 (35)
≤ 7.0	34 (16)
Time from extubation to BSE examination, median (IQR), h	25 (21-45)
Time from BSE examination to FEES examination, median (IQR), h	1 (0.5-2)
Time from extubation to FEES examination, median (IQR), h	26.5 (22-47)
Charlson Comorbidity Index score, median (IQR)	3 (2-6)
APACHE II score, median (IQR)	19 (14-27)
BMI, median (IQR), kg/m ²	27.8 (23.6-34.1)

APACHE = Acute Physiology And Chronic Health Evaluation; BSE = bedside swallowing evaluation; FEES = flexible endoscopic evaluation of swallowing; IQR = interquartile range.

and as dictated by individual institutional protocols, a limited amount (0.2 mL) of 4% intranasal lidocaine spray could be administered intranasally before laryngoscopic insertion. This amount of lidocaine increases patient comfort without altering the swallow.³³⁻³⁵ The SLP then followed the same bolus administration protocol that was administered during the BSE. All FEES examinations were video and sound recorded. After watching the recordings, one of the study investigators (S. A. L.) performed a blinded review of the FEES examination and determined a penetration-aspiration scale (PAS) score for each of the boluses administered.³⁶ A PAS score of 6 or more represents aspiration, a PAS score of 6 or 7 represents nonsilent aspiration, and a PAS score of 8 represents silent aspiration. Based on the overall FEES examination, the second SLP determined clinically appropriate liquid and food dietary recommendations.

Statistical Analysis

Using the FEES as the gold standard, the accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were determined for the swallowing trials and overall BSE recommendations with 95% CIs. Because the BSE is used as a screening test and we wanted to limit the number of false-negative results, we focused on the sensitivity and NPVs. We then used recursive partitioning decision trees to predict the undesirable outcomes of overall aspiration, silent

aspiration, and nonsilent aspiration on the FEES.^{37,38} Recursive partitioning was selected as the statistical model specifically to develop a decision tree model. To maintain sensitivity, the decision trees were not trimmed by cross-validation. Initially, all 175 possible demographic and BSE predictor variables were used as candidate predictors for each outcome. We included only the results of the consistency boluses in the model and not the results of each of the five clinical signs (cough, throat clearing, change in vocal quality, wet breath sounds, or stridor) for each consistency bolus. Of the original 213 patients, 146 had sufficient data for all variables to be used in the model selection process. After choosing the optimal decision tree for each outcome, the set of all variables chosen as predictors for any model was taken as the set of candidate predictors for the final models. For the outcomes and selected predictors, 160 patients had sufficient data to use in further analysis. The final models selected different predictors from the 12 candidates for each outcome. For recursive partitioning, no specific sample size calculations exist. In a simulated data analysis, a sample size of 200 yielded an area under the receiver operating characteristic curve of 0.75 for the resultant tree algorithm. To be conservative, we enrolled an additional 13 patients. We also tested models for accuracy using all data and used leave-one-out cross-validation to determine the general applicability and internal validation of the models.

Results

From August 2015 through July 2018, a total of 248 patients were enrolled into the study. A total of 22 patients withdrew from the study before the FEES examination was performed, and an additional 13 patients had incomplete FEES examinations. Therefore, 213 patients were included in the final analysis (Fig 1).³⁹ All of the study procedures (BSE and FEES examination) were performed in the ICU.

Patient demographics are included in Table 1. Median age was 57 years (interquartile range, 47-66 years). Median duration of mechanical ventilation was 126 hours (interquartile range, 72-206 hours). A total of 131 patients (62%) were men. Median admission APACHE II score was 19 (interquartile range, 14-27), and median Charlson Comorbidity Index was 3 (interquartile range, 2-6). Additional demographic information is included in e-Appendix 1. Median time from extubation to the BSE was 25 hours (interquartile range, 21-45 hours), and the FEES was completed on average 1 hour after the BSE was completed (interquartile range, 0.5-2 hours). During the FEES examination, most patients received at least one trial of each bolus consistency test: ice, 100% (213/213); puree, 96.2% (205/213); nectar, 98.6% (210/213); liquid, 92.5% (197/213); and solids, 82.6% (176/213). Based on the PAS scoring on the FEES examination, overall aspiration (PAS score of 6-8) was detected in 70 patients (33%) with at least one bolus of any volume or

consistency. Nonsilent aspiration (PAS score of 6 or 7) was detected in 52 patients (24%), and silent aspiration (PAS score of 8) was detected in 29 patients (14%). Eleven patients (5%) demonstrated both nonsilent and silent aspiration on different consistencies. The percentages of overall, nonsilent, and silent aspiration for each individual consistency are displayed in Table 2. Five adverse events occurred during the study. One patient vomited during the FEES, and one vomited the night after the study; both of these episodes were considered to be associated directly with study-related events. Three patients demonstrated respiratory failure within 24 hours of the study; these were not considered study-related events.

When compared with the gold standard FEES examination, the accuracy of the BSE examination for detecting overall, nonsilent, and silent aspiration was 52% (110/213; 95% CI, 45%-58%), 48% (102/213; 95% CI, 41%-55%), and 39% (83/213; 95% CI, 32%-46%). The sensitivity, specificity, PPV, and NPV for overall, nonsilent, and silent aspiration and for each consistency testing for FEES-documented aspiration are displayed in Table 3.

Based on the interpretation of the BSE examination, the liquid dietary recommendations were: regular thin liquids in 46%, nectar-thickened liquids in 29%, honey-thickened liquids in 2%, and nothing by mouth in 23% of patients. When compared with liquid dietary

TABLE 2] Percentage of Different Types of Aspiration by Consistency

Consistency	Overall Aspiration	Nonsilent Aspiration	Silent Aspiration
Ice	8% (n = 18/213)	5% (n = 11/213)	3% (n = 7/213)
Puree	3% (n = 7/205)	1% (n = 3/205)	2% (n = 4/205)
Nectar	8% (n = 17/210)	5% (n = 11/210)	3% (n = 6/210)
Liquid	27% (n = 54/197)	19% (n = 38/197)	10% (n = 19/197)
Solid	2% (n = 3/176)	2% (n = 3/176)	0% (n = 0/176)

Unique patients demonstrated nonsilent aspiration with more than one consistency. Similarly, unique patients demonstrated silent aspiration with more than one consistency. So, the total number of episodes of aspiration with each consistency is greater than the unique number of patients who aspirated with at least one consistency.

recommendations based on the FEES examination (stratified as regular vs modified recommendations or nothing by mouth), the accuracy of the BSE dietary recommendations was 70% (149/213; 95% CI, 64%-

mechanical National Dysphagia Diet 2 in 29%, puree National Dysphagia Diet 1 in 11%, and nothing by mouth in 22% of patients. When compared with solid dietary recommendations based on the FEES

TABLE 3] Diagnostic Accuracy of the Aspiration Types, Consistencies, and Decision Tree Algorithms

Aspiration Types, Consistencies, and Decision Tree Algorithms	Sensitivity (95% CI) (n/N)	Specificity (95% CI) (n/N)	Positive Predictive Value (95% CI) (n/N)	Negative Predictive Value (95% CI) (n/N)
BSE results				
Overall aspiration	83% (90%-98%) (58/70)	36% (28%-44%) (52/143)	39% (31%-47%) (58/149)	81% (72%-91%) (52/64)
Nonsilent aspiration	87% (78%-96%) (45/52)	35% (28%-42%) (57/161)	30% (23%-37%) (45/149)	89% (81%-97%) (57/64)
Silent aspiration	83% (69%-97%) (24/29)	32% (25%-39%) (59/184)	16% (10%-22%) (24/149)	92% (85%-99%) (59/64)
BSE consistencies				
Ice	55% (32%-78%) (10/18)	74% (68%-80%) (144/195)	16% (7%-25%) (10/61)	95% (94%-96%) (144/152)
Puree	28% (0%-61%) (2/7)	82% (77%-87%) (166/202)	5% (0%-12%) (2/38)	97% (94%-100%) (166/171)
Nectar, thick	64% (41%-87%) (11/17)	68% (61%-75%) (132/195)	15% (8%-23%) (11/74)	96% (94%-98%) (132/138)
Thin	74% (62%-86%) (38/51)	50% (42%-58%) (74/148)	34% (25%-43%) (38/112)	85% (77%-93%) (74/87)
Solid	100% (1/1)	82% (76%-88%) (150/182)	3% (0%-9%) (1/33)	100% (150/150)
Decision algorithms				
Overall aspiration	95% (90%-98%)	67% (60%-73%)	54% (47%-60%)	97% (95%-99%)
Nonsilent aspiration	96% (93%-100%)	65% (57%-72%)	41% (33%-48%)	99% (98%-99%)
Silent aspiration	100%	75% (68%-82%)	35% (28%-43%)	100%

76%). Of the 30% of cases where disagreement occurred, the BSE resulted in overly conservative liquid diet recommendations in 20% of cases and in overly liberal recommendations in 10% of cases. The BSE solid dietary recommendations were: regular in 38%,

examination (stratified as regular vs National Dysphagia Diet 2, National Dysphagia Diet 1, or nothing by mouth), the accuracy of the BSE dietary recommendations was 80% (170/213; 95% CI, 75%-85%). Overall, 10% of patients received a conservative

altered solid diet recommendation based on the BSE results when the FEES examination results recommended a regular solid diet. In contrast, 10% of patients received a liberal diet recommendation of regular solid diet when the FEES examination results recommended an altered solid diet.

Development of a Streamlined BSE-Based Clinical Decision Tree Algorithm

For overall aspiration, the decision tree algorithm (derived from the 160 patients for whom complete data existed for all of the predictor and outcome variables) identified five factors (duration of intubation, the results of the thin liquid consistency bolus 2-oz (60-mL) volume, APACHE II score, voice

quality, and type of ICU patient diagnosis [medical or cardiac vs other]) that resulted in a sensitivity of 95%, specificity of 67%, PPV of 54%, and a NPV of 97% (Fig 2, Table 3). Voice quality was determined as a single assessment during the BSE examination (normal vs abnormal defined as aphonic, hoarse, or wet). If the diagnostic algorithm was used as a clinically useful screening test, then nearly half (49%) of all patients would be identified accurately as not being an aspiration risk and would require no further diagnostic testing to detect aspiration. The leave-one-out cross-validation of the decision tree yielded different predictive values: sensitivity of 74% (95% CI, 68%-80%), specificity of 67% (95% CI, 60%-73%), PPV of 47% (95% CI, 41%-54%), and NPV of

Figure 2 – Decision tree algorithm for overall aspiration (N = 160). APACHE = Acute Physiology And Chronic Health Evaluation.

86% (95% CI, 82%-91%). Decision tree algorithms also were developed for nonsilent and silent aspiration. However, both of these algorithms were more complex than the algorithm for overall aspiration, and therefore, we considered them not to be clinically useful (algorithms not shown) (Table 3).

Discussion

In this prospective multicenter study, we evaluated the accuracy of the BSE for detecting aspiration in ARF survivors that previously had required intubation and mechanical ventilation. One third of patients aspirated at least one of the consistencies on their gold standard FEES examination, with thin liquids being the most commonly aspirated consistency. Unfortunately, the overall BSE was fairly inaccurate for predicting FEES-detected aspiration as well as for determining the proper liquid and solid dietary recommendations. We were able to develop a five-variable decision tree algorithm that maintains a high sensitivity and NPV for FEES-detected overall aspiration. Of the five variables, three are easily obtained clinical data (length of intubation, APACHE II score, and medical or cardiac cause of critical illness), one is a physical examination finding (voice quality), and the fifth is a 2-oz (60-mL) thin liquid consistency test. Because of its excellent NPV and relative ease to perform, our decision tree algorithm could serve as a clinically useful screening test for aspiration in ARF survivors. In addition, the BSE examination is designed to elicit behaviors that identify clinical signs suggestive of aspiration and serves other important purposes in the assessment of dysphagia and the determination of airway safety during swallowing.

Our results concerning the accuracy of the BSE for detecting aspiration are consistent with prior studies in other patient populations. In stroke patients, the BSE yielded a sensitivity of 80% to 91% and a specificity of 30% to 68% when compared with either FEES or a VFSS.^{21,40-42} In our prior study of ARF survivors, an SLP's recommendation for altered diet yielded a sensitivity of 86% and specificity of 52%.⁴³ In light of these results of the BSE for predicting aspiration, other screening assessments have been proposed to identify a patient's risk of aspiration. These other screening assessments ideally would identify more accurately those patients who could have their diet advanced or would require more formal evaluation by an SLP, such as an FEES before oral feeding.⁴⁴ One such proposed

screening assessment is the 3-oz water swallow test (3-WST).⁴⁵ Two large studies of heterogeneous populations reported that the 3-WST yielded a high sensitivity (96.5% and 98%) for aspiration. However, the assessments were not blinded, because the same individual performed the 3-WST directly after the FEES.^{46,47} In one of these studies, nearly 10% of patients did not undergo the 3-WST because the FEES already had demonstrated significant aspiration. This exclusion of high-risk patients may have eliminated some patients at risk for silent aspiration who may have failed the 3-WST.⁴⁷ Future studies will be required to validate our algorithm and to compare its accuracy for aspiration against the 3-WST and should consider other screening measures such as cough peak flow strength. In addition, further diagnostic examinations are strongly encouraged for patients identified to be at risk of aspiration to determine the specific cause and to develop an appropriate management plan to treat all elements contributing to dysphagia.

A growing body of literature reports poor sensitivity of the BSE for patients who are deemed to be at high risk of aspiration or who demonstrate other signs of severe dysphagia. In patients at high risk of aspiration, SLPs should consider bypassing at least part of the BSE and proceeding directly to an instrumental assessment of swallowing and dysphagia, such as the FEES and VFSS. After being validated, our algorithm could prove to be a clinically useful screening tool (Fig 1). The algorithm would require the user to gather a finite amount of information from the medical record (length of intubation, APACHE II score, and medical unit), to determine vocal quality, and to administer a simple water swallow challenge. This information would allow the user to determine if further instrumental assessment is needed or if the patient could be recommended an oral diet safely. If validated, the use of our diagnostic algorithm may not need to be limited to SLPs. Other health care professionals such as registered nurses may be trained effectively to perform these screening examinations for aspiration. In fact, it has been shown that registered nurses can administer and interpret the 3-WST effectively.⁴⁸

Dietary recommendations do not depend entirely on a finding of aspiration. In our study, dietary recommendations varied substantially when depending on the interpretations of the BSE or FEES. This suggests that other swallowing parameters, beyond aspiration, were

noted and taken into consideration. Whether these other signs of dysphagia were related to adverse outcomes was not determined and is fruitful ground for further studies. Although more hospitals have access to VFSS equipment, we used the FEES as our gold standard instrumental examination.¹⁴ Previous studies have reported excellent agreement and good interrater and intrarater reliability between the VFSS and FEES, and the FEES has been reported to be more sensitive for penetration (based on a PAS score of 2-5) as well as aspiration (PAS score of 6-8).⁴⁹⁻⁵⁴

Our study has several important strengths. This was a multicenter trial performed at four sites with a relatively large sample size of 213 patients. The study also included a well-constructed protocol for both the BSE and FEES that resulted in a paucity of missing data. We ensured that two different SLPs performed the BSE and the FEES and that the SLP who performed the subsequent FEES examination was blinded to the results of the BSE. However, the SLPs were aware of the goal of the study, and this may have resulted in a more conservative interpretation of their BSE findings. Additionally, although no true standardization of BSE exists in the field of speech-language pathology, the established BSE protocol included widely accepted components of the BSE. The BSE included consistency trials with higher volumes. This is important because silent aspiration has been shown to be volume dependent, and screens with higher volumes like the 3-WST have been shown to have strong validation in detecting aspiration risk.^{46,47}

Our study also had some limitations. The study was performed only in academic institutions. Therefore, it is unknown how those evaluations would perform in different nonacademic settings. We chose a design with sequential assessment of aspiration initially with the BSE and then the FEES. A nonrandomized ordering of the two examinations was not chosen because this is the order that commonly would be performed in the clinical setting, and no evidence suggests that the performance of the BSE impacts the performance or results of the FEES. We did not collect data regarding how the endotracheal cuff pressures were managed or where patients had nasal or oral gastric tubes in place. Future studies will be necessary to determine the impact of these factors of the risk of

aspiration. We did not test the accuracy of our algorithm compared with the 3-WST. Our study design dictated that the 3-WST be performed after the other three liquid boluses were administered. If patients showed clear signs of aspiration on the previously administered liquid boluses, the 3-WST did not have to be performed. As a result, only 56% of participants underwent the 3-WST. Therefore, future studies will be required to compare our algorithm against the 3-WST. Most of the patients (69% [146/213]) received 0.2 mL of intranasal lidocaine at the discretion of the SLP. However, the use of topical lidocaine at this dose during the FEES can improve patient comfort and has not been reported to increase the risk of penetration or aspiration.³³⁻³⁵ We used single cutoff values for each of the predictor variables. It is possible that using different gradients for some variables (like APACHE scores) could enhance the predictive accuracy of the decision tree algorithm. Most patients underwent the BSE 1 day after extubation. Whether temporal resolution of postextubation swallowing dysfunction results in aspiration also is currently unknown.⁴⁵ We did not include critically ill patients with burn injuries, for whom additional testing may be necessary because of inhalational injuries.⁵⁵ We dichotomized the assessment of voice quality as simply normal or abnormal. Future studies could further delineate voice quality using clinician-based scales such as the Grade, Roughness, Breathiness, Asthenia, Strain or the Consensus Auditory Perceptual Evaluation-Voice.⁵⁶ Finally, our internal validation with using a leave-one-out cross-validation yielded less robust results than the original model, identifying the need for future external validation of the algorithm.

Conclusions

Clinical BSEs frequently are used after mechanical ventilation to detect aspiration and to guide dietary and other recommendations. However, these BSEs are relatively inaccurate for the detection of aspiration. We were able to develop a relatively simple five-variable decision tree algorithm that could serve as a screening tool for aspiration in ARF survivors. However, future studies are necessary to validate the accuracy of our algorithm in ARF survivors and possibly other critically ill patients at risk of aspiration.

Acknowledgments

Author contributions: All authors contributed substantially to the conception and design of the study, the acquisition of data, or the analysis and interpretation of the data; drafted or provided critical revision of the article; provided final approval of the version submitted for publication; and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. M. M., S. D. W., H. W., D. D., G. P. K., and S. L. performed material preparation and analysis. M. M. wrote the first draft of the manuscript, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Financial/nonfinancial disclosures: None declared.

Role of sponsors: The sponsor had no role in the design of the study, the collection and analysis of the data, or the preparation of the manuscript.

Additional information: The e-Appendix and e-Table can be found in the Supplemental Materials section of the online article.

References

- Vincent JL, Akka S, De MA, et al. The epidemiology of acute respiratory failure in critically ill patients. *Chest*. 2002;121(5):1602-1609.
- Wunsch H, Linde-Zwirble WT, Angus DC, Hartman ME, Milbrandt EB, Kahn JM. The epidemiology of mechanical ventilation use in the United States. *Crit Care Med*. 2010;38(10):1947-1953.
- Barker J, Martino R, Reichardt B, Hickey EJ, Ralph-Edwards A. Incidence and impact of dysphagia in patients receiving prolonged endotracheal intubation after cardiac surgery. *Can J Surg*. 2009;52(2):119-124.
- El Soth A, Okada M, Bhat A, Pietrantoni C. Swallowing disorders post orotracheal intubation in the elderly. *Intensive Care Med*. 2003;29(9):1451-1455.
- Macht M, Wimbish T, Clark BJ, et al. Postextubation dysphagia is persistent and associated with poor outcomes in survivors of critical illness. *Crit Care*. 2011;15(5):R231.
- Skoretz SA, Flowers HL, Martino R. The incidence of dysphagia following endotracheal intubation: a systematic review. *Chest*. 2010;137(3):665-673.
- de Aguilar-Nascimento JE, Kudsk KA. Clinical costs of feeding tube placement. *JPN J Parenter Enteral Nutr*. 2007;31(4):269-273.
- Martino R, Foley N, Bhogal S, Diamant N, Speechley M, Teasell R. Dysphagia after stroke: incidence, diagnosis, and aspiration in adults subsequent to stroke. *J Commun. Disord*. 2001;34(1-2):55-72.
- McCullough GH, Wertz RT, Rosenbek JC, Mills RH, Ross KB, Ashford JR. Inter- and intrajudge reliability of a clinical examination of swallowing in adults. *Dysphagia*. 2000;15(2):58-67.
- O'Horo JC, Rogus-Pulia N, Garcia-Arguello L, Robbins J, Safdar N. Bedside diagnosis of dysphagia: a systematic review. *J Hosp Med*. 2015;10(4):256-265.
- Ramsey DJ, Smithard DG, Kalra L. Early assessments of dysphagia and aspiration risk in acute stroke patients. *Stroke*. 2003;34(5):1252-1257.
- Rosenthal DJ, Mendoza TR, Chambers MS, et al. The M. D. Anderson symptom inventory-head and neck module, a patient-reported outcome instrument, accurately predicts the severity of radiation-induced mucositis. *Int J Radiat Oncol Biol Phys*. 2008;72(5):1355-1361.
- Smithard DG, O'Neill PA, Park C, et al. Can bedside assessment reliably exclude aspiration following acute stroke? *Age Ageing*. 1998;27(2):99-106.
- Virvidaki IE, Nasios G, Kosmidou M, Giannopoulos S, Milionis H. Swallowing and aspiration risk: a critical review of non instrumental bedside screening tests. *J Clin Neurol*. 2018;14(3):265-274.
- Warms T, Richards J. "Wet voice" as a predictor of penetration and aspiration in oropharyngeal dysphagia. *Dysphagia*. 2000;15(2):84-88.
- Park J, Yoo W, Yoo B. Standard recipes for the preparation of thickened barium liquids used in the diagnosis of dysphagia. *Clin Nutr Res*. 2019;8(4):265-271.
- Cichero JA, Lam P, Steele CM, et al. Development of international terminology and definitions for texture-modified foods and thickened fluids used in dysphagia management: the IDDSI framework. *Dysphagia*. 2017;32(2):293-314.
- Cho HM, Yoo B. Rheological characteristics of cold thickened beverages containing xanthan gum-based food thickeners used for dysphagia diets. *J Acad Nutr Diet*. 2015;115(1):106-111.
- Seo CW, Yoo B. Steady and dynamic shear rheological properties of gum-based food thickeners used for diet modification of patients with dysphagia: effect of concentration. *Dysphagia*. 2013;28(2):205-211.
- Fife TA, Butler SG, Langmore SE, et al. Use of topical nasal anesthesia during flexible endoscopic evaluation of swallowing in dysphagic patients. *Ann Otol Rhinol Laryngol*. 2015;124(3):206-211.
- Kamarunas EE, McCullough GH, Guidry TJ, Mennemeier M, Schluterman K. Effects of topical nasal anesthetic on fiberoptic endoscopic examination of swallowing with sensory testing (FEESST). *Dysphagia*. 2014;29(1):33-43.
- pulmonary complications. *Stroke*. 2005;36(12):2756-2763.
- Metheny NA, Clouse RE, Chang YH, Stewart BJ, Oliver DA, Kollef MH. Tracheobronchial aspiration of gastric contents in critically ill tube-fed patients: frequency, outcomes, and risk factors. *Crit Care Med*. 2006;34(4):1007-1015.
- Quill TE. Utilization of nasogastric feeding tubes in a group of chronically ill, elderly patients in a community hospital. *Arch Intern Med*. 1989;149(9):1937-1941.
- Schefold JC, Berger D, Zurcher P, et al. Dysphagia in Mechanically Ventilated ICU Patients (DYNAMICS): a prospective observational trial. *Crit Care Med*. 2017;45(12):2061-2069.
- Brodsky MB, Nollet JL, Spronk PE, Gonzalez-Fernandez M. Prevalence, pathophysiology, diagnostic modalities and treatment options for dysphagia in critically ill patients [published online ahead of print April 16, 2020]. *Am J Phys Med Rehabil*. <https://doi.org/10.1097/PHM.0000000000001440>.
- Brodsky MB, Suiter DM, Gonzalez-Fernandez M, et al. Screening accuracy for aspiration using bedside water swallow tests: a systematic review and meta-analysis. *Chest*. 2016;150(1):148-163.
- Macht M, Wimbish T, Clark BJ, et al. Diagnosis and treatment of post-extubation dysphagia: results from a national survey. *J Crit Care*. 2012;27(6):578-586.
- Carnaby GD, Crary MA. Development and validation of a cancer-specific swallowing assessment tool: MASA-C. *Support Care Cancer*. 2012;22:595-602.
- Chen AY, Frankowski R, Bishop-Leone J, et al. The development and validation of a dysphagia-specific quality-of-life questionnaire for patients with head and neck cancer: the M. D. Anderson dysphagia inventory. *Arch Otolaryngol Head Neck Surg*. 2001;127(7):870-876.
- List MA, Ritter-Sterr C, Lansky SB. A performance status scale for head and neck cancer patients. *Cancer*. 1990;66(3):564-569.
- Mari F, Matei M, Ceravolo MG, Pisani A, Montesi A, Provinciali L. Predictive value of clinical indices in detecting aspiration in patients with neurological disorders. *J Neurol Neurosurg Psychiatry*. 1997;63(4):456-460.
- Martino R, Silver F, Teasell R, et al. The Toronto Bedside Swallowing Screening Test (TOR-BSST): development and validation of a dysphagia screening tool for patients with stroke. *Stroke*. 2009;40(2):555-561.
- McCullough GH, Rosenbek JC, Wertz RT, McCoy S, Mann G, McCullough K. Utility of clinical swallowing examination measures for detecting aspiration post-stroke. *J Speech Lang Hear Res*. 2005;48(6):1280-1293.
- McCullough GH, Wertz RT, Rosenbek JC. Sensitivity and specificity of clinical/ bedside examination signs for detecting

35. O'Dea MB, Langmore SE, Krisciunas GP, et al. Effect of lidocaine on swallowing during FEES in patients with dysphagia. *Ann Otol Rhinol Laryngol*. 2015;124(7):537-544.
36. Rosenbek JC, Robbins JA, Roecker EB, Coyle JL, Wood JL. A penetration-aspiration scale. *Dysphagia*. 1996;11(2):93-98.
37. Brown AW, Malec JF, McClelland RL, Diehl NN, Englander J, Cifu DX. Clinical elements that predict outcome after traumatic brain injury: a prospective multicenter recursive partitioning (decision-tree) analysis. *J Neurotrauma*. 2005;22(10):1040-1051.
38. Seheult JN, Anto VP, Farhat N, et al. Application of a recursive partitioning decision tree algorithm for the prediction of massive transfusion in civilian trauma: the MTPitt prediction tool. *Transfusion*. 2019;59(3):953-964.
39. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ*. 2015;351:h5527.
40. Rosenbek JC, McCullough GH, Wertz RT. Is the information about a test important? Applying the methods of evidence-based medicine to the clinical examination of swallowing. *J Commun Disord*. 2004;37(5):437-450.
41. Leder SB, Espinosa JF. Aspiration risk after acute stroke: comparison of clinical examination and fiberoptic endoscopic evaluation of swallowing. *Dysphagia*. 2002;17(3):214-218.
42. Smith HA, Lee SH, O'Neill PA, Connolly MJ. The combination of bedside swallowing assessment and oxygen saturation monitoring of swallowing in acute stroke: a safe and humane screening tool. *Age Ageing*. 2000;29(6):495-499.
43. Lynch YT, Clark BJ, Macht M, et al. The accuracy of the bedside swallowing evaluation for detecting aspiration in survivors of acute respiratory failure. *J Crit Care*. 2017;39:143-148.
44. Brodsky MB, Mayfield EB, Gross RD. Clinical decision making in the ICU: dysphagia screening, assessment, and treatment. *Semin Speech Lang*. 2019;40(3):170-187.
45. Leder SB, Warner HL, Suiter DM, et al. Evaluation of swallow function post-extubation: is it necessary to wait 24 hours? *Ann Otol Rhinol Laryngol*. 2019;128(7):619-624.
46. Suiter DM, Leder SB. Clinical utility of the 3-ounce water swallow test. *Dysphagia*. 2008;23(3):244-250.
47. Leder SB, Suiter DM, Green BG. Silent aspiration risk is volume-dependent. *Dysphagia*. 2011;26(3):304-309.
48. Warner HL, Suiter DM, Nystrom KV, Poskus K, Leder SB. Comparing accuracy of the Yale swallow protocol when administered by registered nurses and speech-language pathologists. *J Clin Nurs*. 2014;23(13-14):1908-1915.
49. Kelly AM, Leslie P, Beale T, Payten C, Drinnan MJ. Fiberoptic endoscopic evaluation of swallowing and videofluoroscopy: does examination type influence perception of pharyngeal residue severity? *Clin Otolaryngol*. 2006;31(5):425-432.
50. Kelly AM, Drinnan MJ, Leslie P. Assessing penetration and aspiration: how do videofluoroscopy and fiberoptic endoscopic evaluation of swallowing compare? *Laryngoscope*. 2007;117(10):1723-1727.
51. Wu CH, Hsiao TY, Chen JC, Chang YC, Lee SY. Evaluation of swallowing safety with fiberoptic endoscope: comparison with videofluoroscopic technique. *Laryngoscope*. 1997;107(3):396-401.
52. Madden C, Fenton J, Hughes J, Timon C. Comparison between videofluoroscopy and milk-swallow endoscopy in the assessment of swallowing function. *Clin Otolaryngol Allied Sci*. 2000;25(6):504-506.
53. Colodny N. Interjudge and intrajudge reliabilities in fiberoptic endoscopic evaluation of swallowing (fees) using the penetration-aspiration scale: a replication study. *Dysphagia*. 2002;17(4):308-315.
54. Hey C, Pluschinski P, Pajunk R, et al. Penetration-aspiration: is their detection in FEES ® reliable without video recording? *Dysphagia*. 2015;30(4):418-422.
55. Clayton NA, Ward EC, Rumbach AF, Cross RR, Kol MR, Maitz PK. Influence of inhalation injury on incidence, clinical profile and recovery pattern of dysphagia following burn injury [published online ahead of print February 27, 2020]. *Dysphagia*. <https://doi.org/10.1007/s00455-020-10098-y>.
56. Karnell MP, Melton SD, Childs JM, Coleman TC, Dailey SA, Hoffman HT. Reliability of clinician-based (GRBAS and CAPE-V) and patient-based (V-RQOL and IPVI) documentation of voice disorders. *J Voice*. 2007;21(5):576-590.