



The Association Between Dysphagia Symptoms, DIGEST Scores, and Severity Ratings in Individuals with Parkinson’s Disease

Mehak Noorani¹ · Rachel K. Bolognone¹ · Donna J. Graville¹ · Andrew D. Palmer¹

Received: 23 September 2021 / Accepted: 10 January 2023 / Published online: 24 January 2023
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

The Swallow Disturbance Questionnaire (SDQ) is a screening tool developed to identify patients with Parkinson’s Disease (PD) in need of objective swallowing evaluation. In a previous investigation, the SDQ did not predict abnormal airway protection on a videofluoroscopic swallowing study (VFSS). This investigation was undertaken to determine whether SDQ scores were more accurate when a global measure was used. The Dynamic Imaging Grade for Swallowing Toxicity (DIGEST) is a validated measure that provides a safety, efficiency, and total severity grade based on VFSS. A secondary analysis was performed using data from 20 patients with PD who had participated in a standardized VFSS protocol. The study sample was predominantly male (80%) with an average age of 71 years, and an average PD duration of 9 years. Using an established cut-off score, participants were subdivided into those with “normal” ($n = 10$) and “abnormal” SDQ scores ($n = 10$). Recordings were scored using the DIGEST protocol by two blinded raters who also rated overall dysphagia severity from the VFSS. There was good agreement between the two raters on the DIGEST and strong correlations between DIGEST scores and clinician perceptions of dysphagia severity. Higher SDQ scores were associated with poorer Efficiency on the DIGEST but not Safety or Total scores. Consistent with other PD studies, subjective perceptions of dysphagia were poorly predictive of objective findings on VFSS. There is little information about the validity of the DIGEST for rating neurogenic dysphagia. Our study provides preliminary support for the use of the DIGEST in the PD population.

Keywords Dysphagia · Deglutition · Screening · Parkinson’s disease · Survey · DIGEST

Introduction

Background

Parkinson’s disease (PD) is a progressive neurodegenerative disorder that affects up to 3% of elderly adults and is projected to affect more than 9 million adults worldwide by 2030 [1]. It has been estimated that over 80% of individuals with Parkinson’s disease will experience dysphagia at some point in their disease progression [2]. Swallowing difficulties can lead to poorer quality of life, reduced ability to take medications on schedule, and increased risk for malnutrition and dehydration [3]. Aspiration pneumonia is currently the

leading cause of death in the PD population, accounting for the cause of mortality in approximately 70% of individuals with the disease [4].

Estimates of dysphagia prevalence in the PD population vary widely. A prevalence of 35% has been reported based on subjective dysphagia complaints but the prevalence approaches 85% when objective dysphagia measures are used [5]. There is a clear mismatch between patient perceptions and findings from objective testing, with silent aspiration occurring in some cases within the first two years of diagnosis [6]. Collectively, these factors make clinical assessment of dysphagia and aspiration-risk in individuals with PD difficult [7]. Currently there is no well-established screening protocol for dysphagia in PD. In one study of PD patients hospitalized with pneumonia, only 13% had previously undergone a swallowing evaluation, suggesting that referral criteria are not well established [8]. A swallowing screening test is a quick and inexpensive test that can be administered by individuals without specialized expertise to determine the need for further evaluation [9]. Several

✉ Andrew D. Palmer
palmeran@ohsu.edu

¹ Department of Otolaryngology-Head & Neck Surgery, NW Center for Voice & Swallowing, Oregon Health & Science University, 3181 SW Sam Jackson Park Rd, Portland, OR 97239, USA

protocols have been recommended, including the use of population-specific screening questionnaires [10, 11], water swallow protocols [12, 13], clinical swallowing evaluations [14], voluntary cough measures [15], and a combination of clinical criteria [2, 7, 16]. To date, none has gained widespread acceptance.

In a recent study, we investigated the predictive value of three swallow screening tools in a cohort of consecutive patients with PD with no confounding medical conditions to determine their accuracy in predicting abnormal airway protection on a videofluoroscopic swallowing study (VFSS) [17]. As described in greater detail below, the screening protocol included three different modes of assessment, each of which has been used previously as a screening tool in the PD population, namely a self-report dysphagia survey, a water swallow test, and a measure of voluntary cough strength. The survey measure used was the Swallow Disturbance Questionnaire (SDQ), a validated self-report tool specifically developed for the identification early dysphagia in individuals with PD [18]. In its original validation, the questionnaire was shown to be able to accurately detect symptoms of dysphagia in the PD population with a sensitivity and specificity of 81%. It had been previously recommended that patients with a total SDQ score of 11 or more should be referred for a comprehensive swallowing evaluation that includes objective testing. Unfortunately, we were not able to replicate this finding in our study and were also not able to find an alternate cut-off score that was more accurate. There were a number of limitations in the previous study, including the fact that airway protection was the only measure of swallowing function that was assessed. Abnormal airway protection was identified using the worst score on the Penetration–Aspiration Scale (PAS) [19] across a series of bolus trials. Although this is common in both clinical practice and research [20], there is some debate about the accuracy of this methodology [21, 22]. In addition, our previous investigation considered only airway protection and other aspects of swallowing, including swallowing efficiency, was not considered. One further limitation was that the previous investigation examined the SDQ total score but did not explore the survey's two subdomains (i.e., Oral and Laryngopharyngeal scores) to see if either of these might be more strongly associated with objective measures. The current study was undertaken to address these questions.

The Dynamic Imaging Grade of Swallowing Toxicity (DIGEST) is a measure of global swallowing severity that uses established metrics of swallowing safety and efficiency in order to assign a global grade of pharyngeal swallow function based on VFSS findings [23]. Based on the frequency, pattern, and amount of penetrated or aspirated material a score is assigned to provide an overall safety grade. The maximum percentage of pharyngeal residue from the first swallow of each bolus consistency is used to determine

an efficiency grade. The DIGEST was validated for a MBS protocol that included two trials each of 5-mL, 10-mL, and self-administered cup sip volumes of thin liquid barium, barium pudding, and a cracker coated in barium paste and is not applied to swallow attempts where compensatory strategies were trialed. Even though the DIGEST was originally developed for use in the head and neck cancer (HNC) population, a number of studies have used it to classify dysphagia in patients with neurologic disease including amyotrophic lateral sclerosis (ALS) [24, 25], anoxic brain injury [26], oculopharyngeal muscular dystrophy (OPMD) [27], and PD [28]. To date, however, the DIGEST protocol has not been formally evaluated for use with neurogenic dysphagia and one of the aims of this study was to further examine the utility of this measure for individuals with PD.

Study Aims

In this study, we sought to address the following aims:

Aim 1: To evaluate the intra- and inter-rater reliability of DIGEST scores and whether DIGEST scores were associated with clinician assessments of overall severity on VFSS.

Aim 2: To determine whether those with normal and abnormal SDQ scores differ in their safety, efficiency, or overall severity using the DIGEST, or with regard to other characteristics.

Aim 3: To examine whether the SDQ subscores (i.e., Oral and Laryngopharyngeal) or other patient characteristics were associated with DIGEST scores.

Methods

Participants

The current study is a secondary analysis of data previously collected using a standardized collection protocol in a large cohort of individuals with PD that has been described previously [17]. All procedures were approved by the Oregon Health and Science University Institutional Review Board (IRB#17516). The data collected during that study were entered into a database which was then used for the current investigation. Data were systematically collected on consecutive individuals who were referred for a swallowing evaluation for symptoms of dysphagia and met the study's inclusion criteria from 1/1/18 to 2/28/20. The participants in that study were individuals with idiopathic PD who had no comorbid conditions that could affect swallowing (including previous radiation or surgery to the head and neck, trauma, or other neurological conditions). A self-reported dysphagia measure, the SDQ [18], was completed by each participant

prior to undergoing VFSS using a standardized protocol, as described in greater detail below. All participants were in the clinical “on” state with patients being asked to take their PD medications before the examination. Demographic and disease-related variables were collected from a patient interview as well as the electronic medical record. The variables in the database included age, gender, body mass index (BMI), previous VFSS, history of pneumonia, diet status, and duration of PD. Diet status was measured using the Functional Oral Intake Scale (FOIS) [29]. PD severity was based on the Hoehn and Yahr (H&Y) scale [30]. A clinical measure of swallowing efficiency was calculated using the 100 ml Timed Water Swallow Test (TWST) [12]. Swallowing efficiency on the TWST is calculated by dividing the volume of water consumed by the amount of time to complete the task with a value of less than 10 ml/second considered abnormal [31]. A hand-held peak flow meter (TruZone; Monaghan Medical Corporation, Syracuse, New York, USA) with a measurement range of 60 to 800 L/min was used to measure volitional peak cough flow (PCF).

Procedure

The SDQ is a validated self-report tool for the identification of early dysphagia in individuals with PD [18]. The survey consists of 15 questions and asks the respondent to rate the frequency of each of item. Five items relate to the oral phase of swallowing (e.g., “Do you experience difficulty chewing solid food, like an apple, cookie or a cracker?”) and ten items relate to the pharyngeal phase (e.g., “Do you cough while swallowing liquids?”). In its original validation study, the questionnaire was reported to be sensitive and specific for the detection of symptoms of dysphagia in this patient population with a sensitivity and specificity of 81%. The authors recommended that PD patients with a total SDQ score of 11 or more should be referred for a comprehensive swallowing evaluation that includes objective testing. Consequently, a score of 11 on the SDQ was used in the current study as a cut-off value to differentiate those with “normal” and “abnormal” scores. The SDQ also includes two subscales: an Oral score can be calculated by summing the first five items and a Laryngopharyngeal score by summing the remaining ten items [11].

Each participant had undergone VFSS using a standardized protocol. During the VFSS, images of the oropharynx and cervical esophagus were recorded in the lateral plane, while the participant swallowed standardized consistencies of Varibar© barium contrast in the following sequence: honey (5 mL), nectar-thick liquid (5 ml, 10 ml, 20 ml), thin liquid (5 ml, 10 ml, 20 ml), pudding (5 ml), and ¼ of a graham cracker with pudding barium (3 ml). The 5 ml boluses were presented via teaspoon, and the 10 ml and 20 ml liquid boluses were presented in a medicine cup. For the 20 ml

liquid boluses, participants were instructed to, “Drink the liquid until it is gone,” which resulted in either sequential swallows or one large swallow. For all other bolus sizes and consistencies, participants were cued to orally contain and then swallow the entire bolus when cued. Images were captured using a continuous image and recorded at rate of 30 frames per second. Video recordings were stored digitally to a picture archiving system (IMPAX), and then downloaded for subsequent analysis.

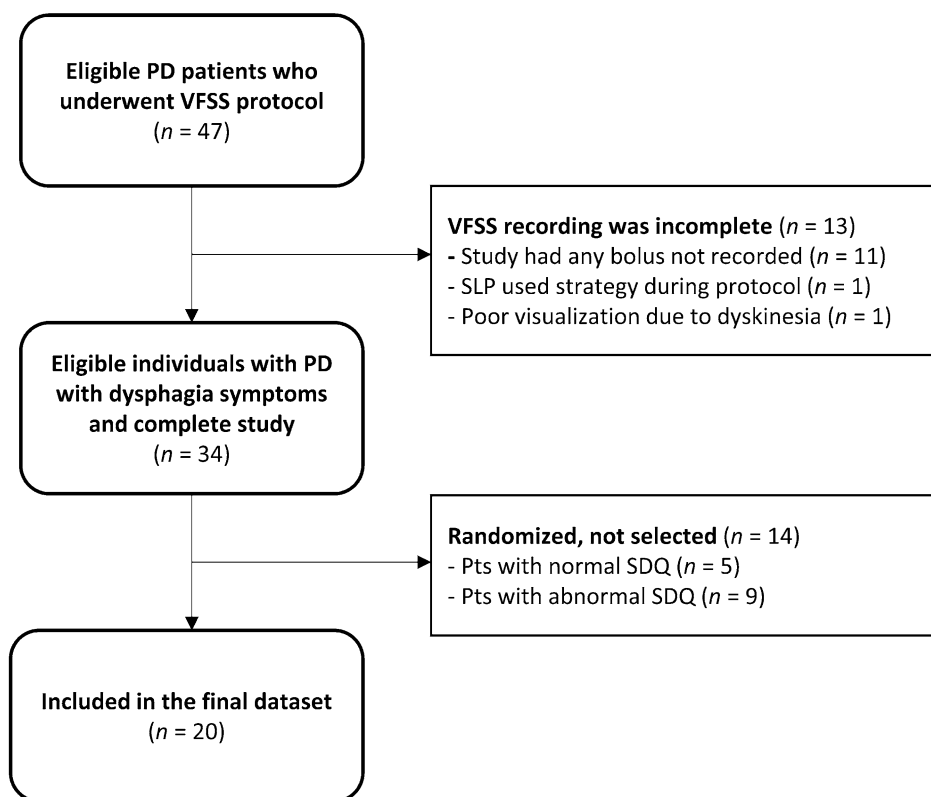
Sample Selection

As there is no normative data for the SDQ, data from our previous analysis were used to perform a power analysis [17]. With 80% power and a p value of 0.10, a sample of 20 patients was required for comparison. A number of authors have advocated for the use of higher p values in preliminary investigations to reduce the risk of a type 2 error [32]. A total of 47 eligible individuals had undergone the VFSS protocol during the initial study period. Upon review of the VFSS recordings, 13 studies were found to be incomplete or not suitable for scoring because imaging for at least one of the nine boluses administered had not been recorded or archived ($n = 11$), at least one of the boluses could not be scored using the DIGEST because the SLP had used a compensatory maneuver for that bolus ($n = 1$), or the study had inadequate visualization for scoring due to patient dyskinesia ($n = 1$). The remaining 34 studies were then subdivided into two groups based using the recommended cut-off score of 11 on the SDQ. There were 15 individuals with SDQ scores in the normal range (i.e., below 11) and 19 individuals with abnormal SDQ scores (i.e., 11 or more). Women made up a minority of the study sample (24%). In order to control for gender effects, 8 males and 2 females were randomly selected from each group to comprise the final sample of 20 participants. The selection of the study sample is outlined in Fig. 1.

Scoring the VFSS Studies Using the DIGEST

The DIGEST is a measure of global swallowing severity that uses subscores of swallowing safety and efficiency in order to arrive upon a global grade of swallow function [23]. The DIGEST was originally developed using the National Cancer Institute’s Common Terminology Criteria for Adverse Events for use in the HNC population. A panel of 9 expert clinician-scientists reviewed 100 VFSS that were completed using a standardized protocol. They assigned severity grades for both constructs of pharyngeal dysphagia (i.e., Safety and Efficiency) as well as an overall Total score for each VFSS. All three grades are scored a using a number from 0 to 4 which can be interpreted as follows: 0, no impairment; 1, mild; 2, moderate; 3, severe;

Fig. 1 Flowsheet outlining the selection of the study sample



and 4, life-threatening. Each bolus administered during the study is first scored using the PAS, originally developed by Rosenbek and colleagues [33]. If a PAS score of 5 or more is assigned, DIGEST safety modifiers of pattern and amount are then applied. The amount of penetrated/aspirated material must be rated as “trace” (i.e., “faint coating, droplets, or trickle of barium on/below TVF”), “neither trace nor gross,” or “gross” (i.e., “> 25% bolus volume”) using the DIGEST guidelines. The frequency, pattern, and amount of penetrated/aspirated material is then used to determine the Safety grade. In addition, each bolus is scored for percentage of pharyngeal residue after the initial swallow attempt. The amount of residue is scored on a four-point scale (“< 10%, minimal to no residue,” “10–49%, less than half residue,” “50–90%, majority residue,” or “> 90%, near complete residue.”) The bolus consistencies with the maximum amount of pharyngeal residue are used to determine the Efficiency grade. After both the Safety and Efficiency grade have been determined, the Total grade can be determined using the interaction of these two subgrades. More recently, the scoring for the DIGEST was amended to address concerns about the definition of mild safety impairment [34]. The DIGEST-v2 was used for the current study. As described above, the standardized DIGEST protocol includes two trials each of 5 mL, 10 mL, and self-administered cup sip volumes of thin liquid barium, barium pudding, and a cracker coated in barium paste for a total of ten bolus trials [23]. In contrast, our

protocol included fewer thin liquid trials but a greater range of textures, including both honey and nectar-thick liquids for a total of five consistencies rather than three and only measured boluses for a total of nine bolus trials. After consultation with one of the developers of the DIGEST protocol for guidance, we adjusted the rules of the DIGEST protocol in the following manner: (a) where penetration or aspiration of more than 50% of thin liquid trials was required for the classification “chronic” our protocol was based on 2 or more out of 3 trials rather than more than 4 or more out of 6 trials; (b) both nectar and thin liquids were defined as a “liquid” and these were considered a single “consistency” for both the safety and efficiency scores when findings were scored for one or more consistencies.

Two licensed, certified, speech-language pathologists (SLPs) with a specialty interest in dysphagia and more than 10 years of clinical experience (D.J.G. and R.K.B.) rated the VFSS studies using the DIGEST scoring form. At the time that the study was conducted, DIGEST trainings had not been developed. The training for the study included reading the original validation article for the study [23], attending a webinar about the DIGEST offered by M.D. Anderson, and participating in a series of training sessions where VFSS recordings were reviewed and the protocol applied in order to reach consensus. This was undertaken as part of an ongoing quality improvement process at our facility to use validated assessment instruments in clinical practice.

The scoring guidelines for the DIGEST-v2 had not been published at the time of analysis and these guidelines were learned from a presentation at the Dysphagia Research Society [35]. The two raters were blinded to the identity of each participant and viewed the studies in a random order. Both raters were randomly assigned 10% of the sample (i.e., two studies) to score twice to allow calculation of intrarater reliability. After viewing each study in its entirety, each rater was also asked to assign an overall severity score using a 100-mm visual analog scale (VAS). The use of VAS a measurement technique is well validated and been widely used in clinical practice and research in the behavioral and social sciences including the measurement of voice- and swallowing-related characteristics [36–39]. Each rater marked an “X” on a 100-mm VAS indicating how they would rate that individual’s swallowing function overall, from “Not impaired at all” to “The most severe impairment possible.” Using the scores provided by the two raters, two other members of the study team (“coders”) independently used the DIGEST protocol to generate a Safety, Efficiency, and Total grade for each participant. Both coders were licensed SLPs (A.D.P. and M.N.) and had received the same training described above for the two raters. Both coders were also randomly assigned 10% of the sample (i.e., two studies) to score twice to allow calculation of intrarater reliability for coding.

Statistical Analysis

Statistical analysis was performed using SPSS, version 25 (IBM, Armonk, New York, 2017). Interrater and intrarater reliability was reported as a percentage of agreement and as a weighted kappa statistic and Kendall’s tau [40–42]. Weighted kappa was calculated using the IBM SPSS (2015) extension software and the strength of the agreement was interpreted using Landis and Koch’s guidelines [43], where 0.21–0.40 indicates fair agreement, 0.41–0.60 moderate agreement, 0.61–0.80 substantial agreement, and 0.81–1.00 almost perfect agreement. Although the DIGEST

is an ordinal scale, previous research has reported mean scores for the measure [44]. After calculation of interrater and intrarater reliability statistics, the two DIGEST scores were averaged to create a single score for each participant for subsequent analyses. Averaging the two scores resulted in a single ordinal score ranging from 0 to 4 in 0.5 point increments for each of the DIGEST subscales. Participants were divided into two groups based on their SDQ score (i.e., under 11 = “normal”; 11 or more = “abnormal”) according to the guidelines for that instrument [18]. Between-group comparisons were made using an independent samples *t* test for continuous variables, the median test for ordinal variables, and Fisher’s exact test for binary variables. Correlations were calculated using Pearson correlations for continuous variables and Spearman correlations for ordinal variables. Given concerns about combining ordinal scores from multiple raters [45], in order to confirm the findings from the combined score from the two raters, the correlations were also compared from that of each rater in isolation. An alpha value of *p* < 0.10 was used to reduce the risk of a type 2 error. All tests were two-tailed.

Results

Aim 1: Intra- and Inter-Rater Reliability for the DIGEST and Association with Severity Scores

The 20 VFSS videos were scored independently by two raters. The two raters scored each of the nine boluses using the 8-point PAS scale and the DIGEST’s 4-point residue scale. With regard to PAS scores, the two raters showed complete agreement for the majority of the 180 VFSS clips (149/180, 83%). There were disagreements of a single point on the PAS for 24 clips (13%) and more than a single point for 7 clips (4%), as shown in Table 1. There was substantial agreement between the two raters for PAS ratings, weighted $\kappa = 0.739$ (95% CI, 0.628 to 0.850), *p* < 0.001, Kendall’s

Table 1 Comparison of PAS scores by rater

		Rater 2								Total
		1	2	3	4	5	6	7	8	
Rater 1	1	129	9	1	0	0	0	0	0	139
	2	5	11	3	0	1	0	0	0	20
	3	2	3	4	0	0	0	0	0	9
	4	0	1	0	0	0	0	0	0	1
	5	0	0	0	2	1	2	0	0	5
	6	0	0	0	0	0	0	0	0	0
	7	1	0	0	0	0	0	1	0	2
	8	0	0	0	0	1	0	0	3	4
Total		137	24	8	2	3	2	1	3	180

Bold values indicate exact agreement

tau = 0.730, $p < 0.001$. Residue ratings also had high levels of agreement with complete agreement for the majority of the 180 clips (157/180, 87%). There were disagreements of a single point on the residue scale for 22 clips (12%) and more than a single point for 1 clip (0.5%), as shown in Table 2. Once again there was substantial agreement between the two raters for residue ratings, weighted $\kappa = 0.769$ (95% CI, 0.689 to 0.849), $p < 0.001$, Kendall's tau = 0.836, $p < 0.001$. Both raters had been randomly assigned 10% of the sample (i.e., two studies) to score twice. Both raters both showed perfect intrarater agreement for all 18 of their PAS scores across the two studies, weighted $\kappa = 1.000$ (95% CI, 1.000 to 1.000), $p < 0.001$. For the residue ratings, Rater 1 demonstrated almost perfect agreement for 17/18 (94%) of the ratings, weighted $\kappa = 0.931$ (95% CI, 0.803 to 1.060), $p < 0.001$, and Rater 2 demonstrated perfect agreement, weighted $\kappa = 1.000$ (95% CI, 1.000 to 1.000), $p < 0.001$.

The two raters each assigned a global severity score for each study using a 100-mm VAS. There was a strong, significant correlation between VAS ratings between the two raters, $r = 0.890$, $p < 0.001$.

After the PAS and residue ratings were generated, two coders then independently used the PAS and residue ratings to assign DIGEST scores based on that instrument's guidelines. Interrater reliability was calculated for the generation of the Safety, Efficiency, and Total DIGEST scores. There was perfect agreement between the two coders for all of the DIGEST-Safety (40/40, 100%), DIGEST-Efficiency (40/40, 100%), and DIGEST-Total scores (40/40, 100%), weighted $\kappa = 1.000$ (95% CI, 1.000 to 1.000), $p < 0.001$, Kendall's tau = 1.000, $p < 0.001$. Intrarater reliability was also calculated for two randomly selected repeated studies. Intrarater reliability for both coders was also perfect (40/40, 100%), weighted $\kappa = 1.000$ (95% CI, 1.000 to 1.000), $p < 0.001$.

As there had been discrepancies between the two original raters for the PAS and residue ratings, this also caused differences in the DIGEST scores for some individuals. For the DIGEST-Safety scale, there was perfect agreement for 14/20 (70%) of the ratings, a difference of 1 point for 5/20 (25%), and one case of a difference of more than 1 point (5%), as shown

Table 2 Comparison of residue scores by rater

		Rater 2				Total
		1	2	3	4	
Rater 1	1	121	3	1	0	125
	2	7	29	3	0	39
	3	0	7	7	2	16
	4	0	0	0	0	0
	Total	128	39	11	2	180

Bold values indicate exact agreement

Table 3 Comparison of DIGEST-Safety scores by rater

		Rater 2				Total
		0	1	2	3	
Rater 1	0	11	1	0	0	12
	1	4	1	0	0	5
	2	0	0	1	0	1
	3	0	1	0	1	2
	Total	15	3	1	1	20

Bold values indicate exact agreement

Table 4 Comparison of DIGEST-Efficiency scores by rater

		Rater 2				Total
		0	1	2	3	
Rater 1	0	10	0	0	0	10
	1	1	4	1	0	6
	2	0	0	0	0	0
	3	0	0	0	4	4
	Total	11	4	1	4	20

Bold values indicate exact agreement

Table 5 Comparison of DIGEST-Total scores by rater

		Rater 2				Total
		0	1	2	3	
Rater 1	0	7	1	0	0	8
	1	2	5	0	0	7
	2	0	0	3	0	3
	3	0	0	1	1	2
	Total	9	6	4	1	20

Bold values indicate exact agreement

in Table 3. For the DIGEST-Efficiency scale, there was perfect agreement for 18/20 (90%) of the ratings and a difference of 1 point for 2/20 (10%), as shown in Table 4. For the DIGEST-Total scale, there was perfect agreement for 16/20 (80%) of the ratings and a difference of 1 point for 4/20 (20%), as shown in Table 5.

Aim: 2 Comparison of Individuals with Normal and Abnormal SDQ Scores

The background data for the 20 participants are presented in Table 6. The study sample was predominantly male (80%) with an average age of 71.05 years (± 10.87 years), and an average PD duration of 9.1 years (± 5.22 years). Most were eating an unrestricted diet (75%), had not previously had a VFSS (80%), and had no recent history of pneumonia (90%). There were no significant differences in background characteristics of the two groups, with one exception: those with a normal SDQ had a significantly higher mean BMI (29.75 ± 4.88) than those with an abnormal SDQ (25.26 ± 3.14 , $p = 0.025$).

For the subsequent analyses, the DIGEST scores from the two raters were averaged so that each of the 20 participants had a single score ranging from 0 to 4 for each of the three DIGEST subscales, as shown in Table 7. The VAS scores from both raters were also averaged to create a single score for each participant. As shown in Table 8, the median score for each of the three DIGEST subscales did not differ significantly by group when compared using a median test. The prevalence of abnormal findings was compared between the two groups in a binary fashion with a score of 0 for each DIGEST subscale being considered “normal” and any other score being “abnormal.” As shown in Table 8, abnormal findings occurred commonly in both groups and did not differ significantly. The VAS ratings for both groups were also compared. Those with an abnormal SDQ had a higher mean VAS rating than those with a normal SDQ score (23.20 ± 26.44 vs. 14.05 ± 12.59 , respectively) but this did not differ significantly.

Table 6 Demographic and clinical characteristics of all participants and compared by group

	All participants ($n = 20$)	Normal SDQ ($n = 10$)	Abnormal SDQ ($n = 10$)	p value
Age (y), $M \pm SD$ (range)	71.05 ± 10.87 (49–90)	70.30 ± 11.16 (49–83)	71.80 ± 11.12 (52–90)	0.767
Male gender, % (n)	80% (16/20)	80% (8/10)	80% (8/10)	1.000
BMI, $M \pm SD$ (range)	27.50 ± 4.61 (21.45–36.51)	29.75 ± 4.88 (21.45–36.51)	25.26 ± 3.14 (21.92–32.07)	0.025*
Unrestricted diet, % (n)	75% (15/20)	80% (8/10)	70% (7/10)	1.000
History of pneumonia, % (n)	10% (2/20)	0% (0/10)	20% (2/10)	0.474
Previous VFSS, % (n)	20% (4/20)	10% (1/10)	30% (3/10)	0.582
PD duration (y), $M \pm SD$ (range)	9.10 ± 5.22 (2–22)	9.00 ± 4.47 (2–15)	9.20 ± 6.13 (2–22)	0.934
H&Y Scale, median (range)	2 (1–4)	3 (2–4)	2 (1–4)	0.170
TWST swallowing efficiency (ml/sec), $M \pm SD$ (range)	10.66 ± 6.28 (2.16–22.27)	12.40 ± 6.49 (2.16–19.65)	9.08 ± 5.97 (2.61–22.27)	0.262
PCF maximum (L/min), $M \pm SD$ (range)	446.50 ± 159.48 (210–750)	489.00 ± 150.96 (290–670)	404.00 ± 163.99 (210–750)	0.244

BMI body mass index, H&Y Hoehn & Yahr, M mean, PCF Peak Cough Flow, SD standard deviation, SDQ Swallowing Disturbance Questionnaire, TWST Timed Water Swallow Test, VFSS videofluoroscopic swallowing study, y years

*Significant at $p < .05$

Table 7 Mean DIGEST scores for all participants and compared by group

Score	All participants (<i>n</i> = 20)	Normal SDQ (<i>n</i> = 10)	Abnormal SDQ (<i>n</i> = 10)
DIGEST-Safety			
0	11	5	6
0.5	5	4	1
1	1	1	0
2	2	0	2
3	1	0	1
DIGEST-Efficiency			
0	10	6	4
0.5	1	1	0
1	4	2	2
1.5	1	0	1
3	4	1	3
DIGEST-Total			
0	7	3	4
0.5	3	3	0
1	5	3	2
2	3	1	2
2.5	1	0	1
3	1	0	1

SDQ swallowing disturbance questionnaire

Aim 3: Association Between Patient Characteristics, Study Scales, and DIGEST Scores

Spearman correlations were used to examine the association between the ratings on the DIGEST and the characteristics of the study participants as well as their scores on the SDQ and the VAS. As shown in Table 9, three of the study variables were significantly associated with DIGEST scores. Older age was moderately associated with worse DIGEST-Safety scores ($r_s = 0.453$, $p = 0.045$). On the FOIS, a less restricted diet was moderately associated with better DIGEST-Efficiency scores ($r_s = -0.474$, $p = 0.035$) and better DIGEST-Total scores ($r_s = -0.433$,

$p = 0.057$). Only one of the SDQ scores was significantly associated with any of the DIGEST ratings. Higher Total scores on the SDQ were associated with worse DIGEST-Efficiency scores ($r_s = 0.388$, $p = 0.091$). In contrast, the global VAS ratings from the VFSS showed statistically significant associations with all three of the DIGEST scales. There were strong, positive associations between global VAS ratings and DIGEST-Safety ($r_s = 0.793$, $p < 0.001$), DIGEST-Efficiency ($r_s = 0.711$, $p < 0.001$), and DIGEST-Total ($r_s = 0.808$, $p < 0.001$) scores. These associations were confirmed for each of the two raters individually and therefore did not seem to be an artifact from averaging the DIGEST scores for the participants.

Discussion

One of the aims of this study was to examine the validity of the DIGEST as a global measure in the PD population. Even though the DIGEST was originally developed for use in the HNC population, the metrics that it uses to generate a summary score are widely used across populations, suggesting that the DIGEST may have broader applicability. The frequency, pattern, and amount of penetrated or aspirated material are calculated to determine an overall safety grade using the PAS [19], a measure that is in widespread clinical and research use for adults with dysphagia of varying etiologies [20, 22]. The maximum percentage of pharyngeal residue and bolus consistency is used to determine an efficiency grade and a number of similar ordinal scales exist for rating pharyngeal residue in a similar fashion [46]. To date, several studies have used the DIGEST to classify dysphagia in patients with neurologic disease. Using a cut-off score of 1 to represent abnormal swallowing, previous studies have used the DIGEST to estimate the prevalence of dysphagia in cohorts of patients with OPMD [27] and also to test the predictive value of swallowing-related items on the ALS Functional Rating Scale-Revised [24]. Two interventional case studies have reported the impact of a respiratory-retraining program for an individual with anoxic brain injury [26] and

Table 8 Median DIGEST scores for all participants and comparison of abnormal scores by group

Score	All participants (<i>n</i> = 20)	Normal SDQ (<i>n</i> = 10)	Abnormal SDQ (<i>n</i> = 10)	<i>p</i> value
DIGEST-Safety, median (range)	0 (0–3)	0.25 (0–1)	0 (0–3)	1.000
DIGEST-Efficiency, median (range)	0.25 (0–3)	0 (0–3)	1 (0–3)	0.656
DIGEST-Total, median (range)	0.75 (0–3)	0.5 (0–2)	1 (0–3)	0.656
Abnormal DIGEST-Safety, % (<i>n</i>)	45% (9/20)	50% (5/10)	40% (4/10)	1.000
Abnormal DIGEST-Efficiency, % (<i>n</i>)	50% (10/20)	40% (4/10)	60% (6/10)	0.656
Abnormal DIGEST-Total, % (<i>n</i>)	65% (13/20)	70% (7/10)	60% (6/10)	1.000
VAS rating, $M \pm SD$ (range)	18.63 ± 20.69 (1–63.5)	14.05 ± 12.59 (2–38.5)	23.20 ± 26.44 (1–63.5)	0.336

M mean, *SD* standard deviation, *SDQ* Swallowing Disturbance Questionnaire

Table 9 Correlation between patient characteristics and study scales with DIGEST scores

Characteristic	DIGEST-Safety		DIGEST-Efficiency		DIGEST-Total	
	r_s	p value	r_s	p value	r_s	p value
Age	0.453**	0.045	0.214	0.365	0.329	0.157
BMI	0.197	0.405	0.165	0.488	0.238	0.313
FOIS score	-0.270	0.250	-0.474**	0.035	-0.433*	0.057
PD duration	-0.020	0.933	0.158	0.505	0.017	0.943
H&Y scale	0.094	0.701	0.323	0.177	0.305	0.204
TWST swallowing efficiency	0.223	0.359	0.014	0.954	0.166	0.497
PCF maximum	-0.171	0.471	0.014	0.954	0.048	0.840
SDQ-Oral score	0.115	0.640	0.322	0.178	0.207	0.396
SDQ-Laryngopharyngeal score	-0.028	0.910	0.235	0.332	0.137	0.577
SDQ-Total score	0.140	0.557	0.388*	0.091	0.296	0.204
VAS rating	0.793***	<.001	0.711***	<.001	0.808***	<.001

BMI body mass index, *FOIS* Functional Oral Intake Scale, *H&Y* Hoehn & Yahr, *PCF* Peak Cough Flow, *SDQ* Swallowing Disturbance Questionnaire, *TWST* Timed Water Swallow Test, *VAS* visual analog scale

*Significant at $p < .10$

**Significant at $p < .05$

***Significant at $p < .001$

PD [28] and used the DIGEST to score fiberoptic endoscopic evaluation of swallowing (FEES) studies performed before and after treatment. Improvements in DIGEST scores mirrored other improvements in swallowing, respiration, and cough-related measures [26, 28]. Plowman and colleagues [25] used the DIGEST as an outcome measure in a randomized controlled trial of expiratory muscle strength training in 48 individuals with ALS. There was a significant improvement in the proportion of individuals with abnormal Total and Efficiency scores on the DIGEST but not on the Safety subscale. These improvements were mirrored by a significant improvement on the FOIS, although no significant difference in Eating Assessment Tool scores was found. The current study adds to this body of literature and shows additional preliminary support for the use of the DIGEST in other populations. In terms of background characteristics of the participants, age was associated with significantly reduced safety scores and there was also a significant association between efficiency scores and diet. These findings are consistent with previous research [17, 25]. The raters' VAS scores of overall dysphagia severity on the VFSS were significantly associated with all three DIGEST scores. These findings provide additional support for the DIGEST as a potentially valid measure for individuals with PD.

In this study, we applied the DIGEST protocol to rate boluses that had been administered according to a previously described standardized VFSS protocol [17]. It is important to note that we did not utilize the DIGEST protocol as outlined in the original validation study. The standardized DIGEST protocol includes two trials each of 5 mL, 10 mL, and self-administered cup sip volumes of thin liquid barium, barium pudding, and a cracker coated in barium paste [23].

In contrast, our protocol included fewer thin liquid trials but a greater range of textures, including both honey and nectar-thick liquids, and only measured boluses. According to the developers, the protocol should ideally include 5–6 trials of a thin liquid bolus, and work on validating other bolus protocols is a future topic for investigation [34]. It is unclear how much DIGEST scores might be affected by the use of a different protocol and whether findings can be directly compared across studies, if different protocols are used [23]. Our rationale for using the DIGEST in this study was to generate summary scores for safety, efficiency, and overall severity and found the DIGEST to be well suited to this purpose. The authors have stated that deviation from the protocol is expected in “extreme cases” and that the DIGEST scores are robust even if there are trials are skipped or compensations are introduced early in the study for safety reasons [34] but this remains for further investigation.

The other aims for this study were to examine the predictive value of the SDQ as a screening tool. There is a considerable body of literature on dysphagia screening and the importance of screening is widely agreed [47, 48]. “Gold standard” instrumental exams are not only expensive but also not always readily available, which may delay preventative measures from being implemented [9]. In the acute stroke population, the value of dysphagia screening has been associated with reduced pneumonia rates, shortened patient length of stay in the hospital, and reduced hospital costs [49–51]. Even in the acute stroke population, however, issues with the accuracy of dysphagia screeners caused the Joint Commission to retire dysphagia screening as a performance indicator for Primary Stroke Center certification

in 2010 [52]. There is a similar need for outpatient dysphagia screening protocols in populations known to be at-risk for the negative consequences of dysphagia, such as those with PD. Accurate screening protocols would allow health-care providers to make appropriate and timely referrals for evaluation to prevent costly and potentially life-threatening complications. Despite considerable research, the optimal screening protocol for outpatient neurogenic populations remains to be defined. There are a large number of general and population-specific dysphagia surveys that have been validated in the literature [53]. Surveys and patient-report tools are a desirable method of screening as they are low cost, easy to administer, and (by definition) focus on the concerns of the individual patient.

At our institution, we have used the SDQ as a screener for PD patients but have questioned its accuracy. In a recent investigation, we compared the SDQ to two other screening tools in a cohort of PD patients seen at our institution and found that none of the three was predictive of reduced airway protection on VFSS [17]. The previous study had a number of limitations, however, including the fact that airway protection was the only measure of swallowing function that was assessed. By comparing more global measures of safety, efficiency, and overall swallowing impairment in a sub-selected cohort and by using the two subscales of the SDQ, we hoped to find that the SDQ was more predictive of objective swallowing status. Unfortunately, we found no significant group differences for any of the DIGEST scores when individuals were grouped by the SDQ cut-off score of 11. When we examined the SDQ scores in a continuous fashion, there was only a correlation between the SDQ total and one of the VFSS measures. Higher SDQ scores were associated with worse swallowing efficiency. As such, it would appear that the SDQ is more predictive of abnormal swallowing strength on VFSS rather than swallowing safety. In further examining correlations between the subscales, the total SDQ score had a weak correlation with PAS scores regardless of whether the worst PAS score across all trials ($r_s = 0.172$, $p = 0.044$) or the average PAS score from all boluses ($r_s = 0.214$, $p = 0.012$) was used. In contrast, the SDQ was more strongly associated with efficiency measures such as diet scores on the FOIS ($r_s = -388$, $p < 0.001$) and swallowing efficiency on the TWST ($r = -351$, $p < 0.001$). Given the concerns about swallowing safety in the PD population, the accuracy of the SDQ in fulfilling its originally stated purpose of identifying those in need of an objective swallowing evaluation remains in doubt. This finding is consistent with other studies that have demonstrated poor associations between subjective complaints and objective dysphagia characteristics in this population [5, 10, 14, 54].

It is unclear why our findings differed from those of the original validation study for the SDQ. Our sample was similar in terms of age, gender, and disease severity to that of the

original sample and actually had a longer disease duration (9.1 vs. 6.7 years) [18]. One key difference is that the original validation study for the SDQ used clinical swallowing evaluations and FEES to assess dysphagia rather than VFSS. The only significant between-group difference in our study was BMI. This suggests that the SDQ may be sensitive to self-perceptions of mealtime difficulties and reduced eating efficiency that may place the individual at-risk for inadequate nutrition and may be most appropriate as a screening measure for this purpose. It does appear, however, that screening measures based on self-report cannot be considered trustworthy for predicting abnormal airway protection on VFSS in the PD population. It has been argued that VFSS itself may not be sufficiently sensitive for identifying early changes in swallowing function in individuals with PD. Jones and Ciucci [55] compared and found that high-resolution manometry (HRM) and a swallowing questionnaire item relating to problems with saliva were better than VFSS in differentiating individuals with PD from controls. Like ours, their study included individuals with early and mid-stage PD and their findings also suggest that changes in swallowing pressure and efficiency are most noticeable at this stage in the disease's development. One alternative explanation for the lack of association between patient-report measures and objective measures of swallowing function is that the objective measures themselves may not be sufficiently sensitive to subtle changes in swallowing physiology that occur early in the disease process. The structured nature of tasks on VFSS including a bolus hold and a cued swallow is well known to affect the timing of swallowing physiology [56, 57], and this may be particularly facilitative for individuals with PD [58]. In addition, the mealtime difficulties that many individuals experience are more than those that are just related to swallowing [59] and it is possible that survey measures may more accurately capture some of these global difficulties. At our institution, therefore, we have changed how we utilize the SDQ. Higher scores on the SDQ indicate that an individual is experiencing dysphagia symptoms that need to be assessed but not necessarily with an objective evaluation. A clinical swallow evaluation is used to better assess the nature of the problem and determine whether strategies are effective. If not, an objective evaluation using either VFSS or FEES is then performed.

Limitations

This study has a number of limitations that should be acknowledged. For this secondary analysis from the larger dataset, we analyzed the data from a small number of randomly selected individuals. Most of the participants demonstrated mild dysphagia as demonstrated by both VAS and DIGEST scores and this may have affected the sensitivity of our findings to those with more severe deficits.

Our study did not utilize the original VFSS protocol outlined in the original DIGEST validation study, and as such, the scores reported here may not be directly comparable to those that used this protocol particularly with regard to the safety subscale. Future work should examine the clinical and research utility of the DIGEST in larger and more diverse samples of participants. In addition, the consideration of VFSS as compared to other objective methods of swallowing assessment including FEES and HRM which might be more sensitive to early changes and whether more subtle changes are accurately captured in patient questionnaires are questions for future investigation.

Conclusion

The results of this study demonstrate that the SDQ did not accurately predict abnormal swallowing function on VFSS in a cohort of 20 adults with PD. Higher SDQ scores were associated with poorer swallowing efficiency but not swallowing safety or overall abnormality. This finding is consistent with previous research that has shown screening methods based on self-report to be poorly predictive of VFSS findings in the PD population. The DIGEST may be a useful tool to characterize global swallowing function in PD for research and clinical practice. In addition, the optimal objective method of screening and assessment in the PD population, particularly those in the early stages of the disease, remains to be defined.

Acknowledgements The authors would like to thank the study participants themselves as well as staff at the NW Clinic for Voice and Swallowing and those in the departments of Neurology and Radiology who provided assistance during the performance of this study. In particular, the authors would like to thank all of their co-investigators on the original study from which the current data were derived, namely Shannon Anderson, Deanna Britton, Linda Bryans, Sara Charney, and Jessica Pietrowski.

Funding This is unfunded, investigator-initiated research.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose. The authors have no conflicts of interest to declare that are relevant to the content of this article. All authors certify that they have no affiliations with or involvement in any organization or entity with any financial or non-financial interest in the subject matter or materials discussed in this manuscript. The authors have no financial or proprietary interests in any material discussed in this article.

References

- Raza C, Anjum R, Shakeel N. Parkinson's disease: mechanisms, translational models and management strategies. *Life Sci.* 2019;226:77–90.
- Suttrup I, Warnecke T. Dysphagia in Parkinson's disease. *Dysphagia.* 2016;31:24–32.
- Plowman-Prine EK, Sapienza CM, Okun MS, Pollock SL, Jacobson C, Wu SS, et al. The relationship between quality of life and swallowing in Parkinson's disease. *Mov Disord.* 2009;24:1352–8.
- Mehanna R, Jankovic J. Respiratory problems in neurologic movement disorders. *Park Relat Disord.* 2010;16:628–38.
- Kalf JG, de Swart BJM, Bloem BR, Munneke M. Prevalence of oropharyngeal dysphagia in Parkinson's disease: a meta-analysis. *Park Relat Disord.* 2012;18:311–5.
- Pflug C, Bihler M, Emich K, Niessen A, Nienstedt JC, Flügel T, et al. Critical dysphagia is common in Parkinson disease and occurs even in early stages: a prospective cohort study. *Dysphagia.* 2018;33:41–50.
- Lam K, Lam FKY, Kwok KL, Yiu KC, Kan EYL, Woo J, et al. Simple clinical tests may predict severe oropharyngeal dysphagia in Parkinson's disease. *Mov Disord.* 2007;22:640–4.
- Martinez-Ramirez D, Almeida L, Giugni JC, Ahmed B, Higuchi MA, Little CS, Chapman JP, Mignacca C, Shukla AW, Hess CW, Hegland KW. Rate of aspiration pneumonia in hospitalized Parkinson's disease patients: a cross-sectional study. *BMC Neurol.* 2015;15:1–6.
- Suiter D. Dysphagia screening: challenges and controversies. *Perspect ASHA Spec Interes Groups SIG.* 2018;13(3):82–8.
- Buhmann C, Flügel T, Bihler M, Gerloff C, Niessen A, Hidding U, et al. Is the Munich dysphagia test–Parkinson's disease (MDT-PD) a valid screening tool for patients at risk for aspiration? *Park Relat Disord.* 2019;61:138–43.
- Cohen JT, Manor Y. Swallowing disturbance questionnaire for detecting dysphagia. *Laryngoscope.* 2011;121:1383–7.
- Belo LR, Gomes NAC, Coriolano MDGWDS, de Souza ES, Moura DAA, Asano AG, Lins OG. The relationship between limit of dysphagia and average volume per swallow in patients with Parkinson's disease. *Dysphagia.* 2014;29:419–24.
- Bours GJJW, Speyer R, Lemmens J, Limburg M, De Wit R. Bedside screening tests vs. videofluoroscopy or fiberoptic endoscopic evaluation of swallowing to detect dysphagia in patients with neurological disorders: systematic review. *J Adv Nurs.* 2009;65:477–93.
- Volonté MA, Porta M, Comi G. Clinical assessment of dysphagia in early phases of Parkinson's disease. *Neurol Sci.* 2002;23:S121–2.
- Pitts T, Bolser D, Rosenbek J, Troche M, Sapienza C. Voluntary cough production and swallow dysfunction in Parkinson's disease. *Dysphagia.* 2008;23:297–301.
- Clarke CE, Gullaksen E, Macdonald S, Lowe F. Referral criteria for speech and language therapy assessment of dysphagia caused by idiopathic Parkinson's disease. *Acta Neurol Scand.* 2009;97:27–35.
- Palmer AD, Charney S, Pietrowski J, Anderson S, Britton D, Bryans L, et al. Dysphagia in Parkinson's disease patients prior to deep brain stimulation: is screening accurate? *Clin Neurol Neurosurg.* 2021;203:106587.
- Manor Y, Giladi N, Cohen A, Fliss DM, Cohen JT. Validation of a swallowing disturbance questionnaire for detecting dysphagia in patients with Parkinson's disease. *Mov Disord.* 2007;22:1917–21.
- Rosenbek JC, Robbins JA, Roecker EB, Coyle JL, Wood JL. A penetration-aspiration scale. *Dysphagia.* 1996;11:93–8.
- Borders JC, Brates D. Use of the penetration-aspiration scale in dysphagia research: a systematic review. *Dysphagia.* 2020;35:583–97.
- McCullough GH, Rosenbek JC, Robbins JA, Coyle JL, Wood JL. Ordinality and intervality of a penetration-aspiration scale. *J Med Speech Lang Pathol.* 1998;6:65–72.

22. Steele CM, Grace-Martin K. Reflections on clinical and statistical use of the penetration-aspiration Scale. *Dysphagia*. 2017;32:602–16.
23. Hutcheson KA, Barrow MP, Barringer DA, Knott JK, Lin HY, Weber RS, et al. Dynamic imaging grade of swallowing toxicity (DIGEST): scale development and validation. *Cancer*. 2017;123:62–70.
24. Chapin JL, Gray LT, Vasilopoulos T, Anderson A, DiBiase L, York JD, et al. Diagnostic utility of the amyotrophic lateral sclerosis functional rating scale-revised to detect pharyngeal dysphagia in individuals with amyotrophic lateral sclerosis. *PLoS ONE*. 2020;15:e0236804.
25. Plowman EK, Tabor-Gray L, Rosado KM, Vasilopoulos T, Robison R, Chapin JL, et al. Impact of expiratory strength training in amyotrophic lateral sclerosis: results of a randomized, sham-controlled trial. *Muscle Nerve*. 2019;59:40–6.
26. Curtis JA, Seikaly ZN, Troche MS. Respiratory–swallow coordination training improves swallowing safety and efficiency in a person with anoxic brain injury. *Am J Speech-Language Pathol*. 2020;29:1965–75.
27. Tabor LC, Plowman EK, Romero-Clark C, Youssof S. Oropharyngeal dysphagia profiles in individuals with oculopharyngeal muscular dystrophy. *Neurogastroenterol Motil*. 2018;30:1–6.
28. Curtis JA, Dakin AE, Troche MS. Respiratory–swallow coordination training and voluntary cough skill training: a single-subject treatment study in a person with Parkinson's disease. *J Speech, Lang Hear Res*. 2020;63:472–86.
29. Crary MA, Carnaby Mann GD, Groher ME. Initial psychometric assessment of a functional oral intake scale for dysphagia in stroke patients. *Arch Phys Med Rehabil*. 2005;86:1516–20.
30. Hoehn MM, Yahr MD. Parkinsonism: onset, progression, and mortality. *Neurology*. 1967;17:427–42.
31. Nicklin J, Nathadwarawala KM, Wiles CM. A timed test of swallowing capacity for neurological patients. *J Neurol Neurosurg Psychiatry*. 1992;55:822–5.
32. Lee EC, Whitehead AL, Jacques RM, Julious SA. The statistical interpretation of pilot trials: should significance thresholds be reconsidered? *BMC Med Res Methodol*. 2014;14:1–8.
33. Robbins J, Coyle J, Rosenbek J, Roecker E, Wood J. Differentiation of normal and abnormal airway protection during swallowing using the penetration–aspiration scale. *Dysphagia*. 1999;14:228–32.
34. Hutcheson KA, Barbon CEA, Alvarez CP, Warneke CL. Refining measurement of swallowing safety in the dynamic imaging grade of swallowing toxicity (DIGEST) criteria: validation of DIGEST version 2. *Cancer*. 2022;128:58–66.
35. Hutcheson KA, Barbon C, Alvarez C, Lewin JS. 2020. Refining Measurement of Mild Safety Impairment (Safety Grade 1) in the DIGEST Criteria: Validation of DIGEST v2. Poster presentation at the Dysphagia Research Society Annual Meeting.
36. Kempster GB, Gerratt BR, Abbott KV, Barkmeier-Kraemer J, Hillman RE. Consensus auditory-perceptual evaluation of voice: development of a standardized clinical protocol. *Am J Speech-Language Pathol*. 2009;18:124–32.
37. Miller MD, Ferris DG. Measurement of subjective phenomena in primary care research: the visual analogue scale. *Fam Pract Res J*. 1993;13:15–24.
38. Pisegna JM, Kaneoka A, Leonard R, Langmore SE. Rethinking residue: determining the perceptual continuum of residue on FEES to enable better measurement. *Dysphagia*. 2018;33:100–8.
39. Wewers ME, Lowe NK. A critical review of visual analogue scales in the measurement of clinical phenomena. *Res Nurs Health*. 1990;13:227–36.
40. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20:37–46.
41. Cohen J. Weighted kappa: nominal scale agreement provision for scaled disagreement or partial credit. *Psychol Bull*. 1968;70:213–20.
42. Kendall MG. A new measure of rank correlation. *Biometrika*. 1938;30:81–93.
43. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159–74.
44. Kirsh E, Naunheim M, Holman A, Kammer R, Varvares M, Goldsmith T. Patient-reported versus physiologic swallowing outcomes in patients with head and neck cancer after chemoradiation. *Laryngoscope*. 2019;129:2059–64.
45. Harpe SE. How to analyze Likert and other rating scale data. *Curr Pharm Teach Learn*. 2015;7:836–50.
46. Steele CM, Peladeau-Pigeon M, Nagy A, Waito AA. Measurement of pharyngeal residue from lateral view videofluoroscopic images. *J Speech, Lang Hear Res*. 2020;63:1404–15.
47. Brodsky MB, Suiter DM, González-Fernández M, Michtalik HJ, Frymark TB, Venediktov R, et al. Screening accuracy for aspiration using bedside water swallow tests: a systematic review and meta-analysis. *Chest*. 2016;150:148–63.
48. Donovan NJ, Daniels SK, Edmiaston J, Weinhardt J, Summers D, Mitchell PH. Dysphagia screening: state of the art. *Stroke*. 2013;44:e24–31.
49. Hinchey JA, Shephard T, Furie K, Smith D, Wang D, Tonn S. Formal dysphagia screening protocols prevent pneumonia. *Stroke*. 2005;36:1972–6.
50. Odderson IR, Keaton JC, McKenna BS. Swallow management in patients on an acute stroke pathway: quality is cost effective. *Arch Phys Med Rehabil*. 1995;76:1130–3.
51. Lee Titsworth W, Abram J, Fullerton A, Hester J, Guin P, Waters MF, et al. Prospective quality initiative to maximize dysphagia screening reduces hospital-acquired pneumonia prevalence in patients with stroke. *Stroke*. 2013;44:3154–60.
52. Lakshminarayan K, Tsai AW, Tong X, Vazquez G, Peacock JM, George MG, et al. Utility of dysphagia screening results in predicting poststroke pneumonia. *Stroke*. 2010;41:2849–54.
53. Patel DA, Sharda R, Hovis KL, Nichols EE, Sathe N, Penson DF, et al. Patient-reported outcome measures in dysphagia: a systematic review of instrument development and validation. *Dis Esophagus*. 2017;30:1–23.
54. Monteiro L, Souza-Machado A, Pinho P, Sampaio M, Nóbrega AC, Melo A. Swallowing impairment and pulmonary dysfunction in Parkinson's disease: the silent threats. *J Neurol Sci*. 2014;339:149–52.
55. Jones CA, Ciucci MR. Multimodal swallowing evaluation with high-resolution manometry reveals subtle swallowing changes in early and mid-stage Parkinson disease. *J Parkinsons Dis*. 2016;6:197–208.
56. Daniels SK, Schroeder MF, DeGeorge PC, Corey DM, Rosenbek JC. Effects of verbal cue on bolus flow during swallowing. *Am J Speech-Language Pathol*. 2007;16:140–7.
57. Nagy A, Leigh C, Hori SF, Molfenter SM, Shariff T, Steele CM. Timing differences between cued and noncued swallows in healthy young adults. *Dysphagia*. 2013;28:428–34.
58. Curtis JA, Huber JE, Dakin AE, Troche MS. Effects of bolus holding on respiratory–swallow coordination in Parkinson's Disease. *Am J Speech-Language Pathol*. 2022;31:705–21.
59. Westergren A, Hagell P, Wendin K, Sjö Dahl Hammarlund C. Conceptual relationships between the ICF and experiences of mealtimes and related tasks among persons with Parkinson's disease. *Nord J Nurs Res*. 2016;36:201–8.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Andrew D. Palmer PhD, CCC-SLP

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Mehak Noorani MS, CCC-SLP

Rachel K. Bolognone MS, CCC-SLP

Donna J. Graville PhD, CCC-SLP