

Reliability of a laryngo-pharyngeal esthesiometer and a method for measuring laryngo-pharyngeal mechano-sensitivity in a prospectively recruited cohort of patients

Luis F. Giraldo-Cadavid^{1,2,3}  · Javier Burguete⁴ · Felipe Rueda² · Ana M. Galvis⁵ · Natalia Castaneda² · Luis M. Agudelo-Otalora⁶ · William D. Moscoso⁶ · Nelson Paez³ · Secundino Fernandez⁷

Received: 18 November 2016 / Accepted: 14 March 2017 / Published online: 24 March 2017
© Springer-Verlag Berlin Heidelberg 2017

Abstract

Background There are not reliable methods for measuring laryngo-pharyngeal mechano-sensitivity (LPMS). We aimed to determine the reliability of a new method for measuring LPMS using a new laryngo-pharyngeal esthesiometer (LPEER) in a prospective cohort of dysphagic stroke and non-dysphagic patients. The patients underwent clinical and endoscopic evaluations of swallowing (FESSST). The LPMS assessments consisted of measurements by an expert and a novel rater of the

laryngeal-adductor reflex threshold (LART), cough reflex threshold (CRT) and gag reflex threshold (GRT) using the LPEER. We assessed the Bland–Altman limits of agreement, the intraclass correlation coefficients (ICCs) and Spearman correlation coefficients (SCCs). For the inter-rater comparisons, we contrasted the expert and novel raters. A total of 1608 measurements were obtained from 34 dysphagic stroke patients and 33 non-dysphagic patients. The intra-rater ICCs for all reflex thresholds were >0.90. The inter-rater ICCs were 0.87 for the LART, 0.79 for the CRT and 0.70 for the GRT. The intra-rater SCCs for all reflex thresholds were above 0.88 ($P < 0.0001$). The inter-rater SCC were 0.80 for the LART, 0.79 for the CRT and 0.70 for the GRT (all $P < 0.0001$). The Bland–Altman plots revealed good agreement for the LART and CRT and moderate agreement for the GRT. The median normal value was 0.14 mN for the LART, 4.4 mN for the CRT and 11.9 mN for the GRT. The median thresholds values in patients with aspiration were LART: 1.31 mN; CRT: 32.9 mN and GRT: 32.9 mN (all $P < 0.006$ vs normal thresholds). The LPEER exhibited substantial to excellent intra- and inter-rater reliability.

Presentation at meetings: A preliminary version of this study was presented at the 2015 ATS International Conference, at Denver, Colorado, USA, on May 2015.

Institutional Review Boards: This study was approved by the Institutional Review Board of University of La Sabana (meeting record number 29, of May 25, 2012) and the Institutional Review Board of Fundación Neumológica Colombiana (meeting record number 198, of January 31, 2014).

Electronic supplementary material The online version of this article (doi:10.1007/s00405-017-4536-5) contains supplementary material, which is available to authorized users.

✉ Luis F. Giraldo-Cadavid
lgiraldo@alumni.unav.es; luisgc@unisabana.edu.co

Javier Burguete
javier@unav.es

Felipe Rueda
anfeliipe_ra@hotmail.com

Ana M. Galvis
anagg@clinicaunisabana.edu.co

Natalia Castaneda
natacasv@gmail.com

Luis M. Agudelo-Otalora
mauricio.agudelo@unisabana.edu.co

William D. Moscoso
william.moscoso@unisabana.edu.co

Nelson Paez
npaez@neumologica.org

Secundino Fernandez
sfgonzalez@unav.es

¹ Clinical Formation Unit, School of Medicine, University of Navarra, Irunlarrea 1, 31080 Pamplona, Navarra, Spain

² Research Department and Clinica Universidad de La Sabana, University of La Sabana School of Medicine, Autonorte de Bogota Km 7, Chia 25001, Cundinamarca, Colombia

Keywords Bronchoscopy · Endoscopes · Larynx · Reliability · Sensory thresholds

Abbreviations

95% CI	95% confidence interval
CRT	Cough reflex threshold
DSS	Dysphagia severity scale
FEESST	Fibre-optic endoscopic evaluation of swallowing with sensory test
FEES	Fibre-optic endoscopic evaluation of swallowing
GRT	Gag reflex threshold
ICC	Intraclass correlation coefficient
IQR	Inter-quartile range (percentile 25 to percentile 75)
LART	Laryngeal-adductor reflex threshold
LPEER	Laryngo-pharyngeal esthesiometer
LPMS	Laryngo-pharyngeal mechano-sensitivity
NIHSS	National Institutes of Health Stroke Scale
OSA	Obstructive sleep apnoea
SCC	Spearman correlation coefficient
SLP	Speech language pathologist

Introduction

Alterations in laryngo-pharyngeal mechano-sensitivity (LPMS) induced either by hypo- or hypersensitive states are common disorders of the upper airway. Such alterations have been reported in dysphagia [1], obstructive sleep apnoea (OSA) [2], and cough due to laryngeal hypersensitivity [3, 4].

Dysphagia affects approximately 8.4% of the general population [5] and is a risk factor for aspiration pneumonia [6], which has a high burden on morbidity and mortality [7, 8]. OSA is a major cardiovascular risk factor having a prevalence above 10% [9] and a high impact on mortality [10]. On the other hand, cough due to laryngeal hypersensitivity may have a prevalence of 6% [11] and causes a significant impairment in health-related quality of life [12].

Moreover, there are experimental interventions that appear to improve hypo- and hypersensitivity problems in

the upper airway [3, 13–15], but objective and reproducible methods of assessing the true efficacies of these interventions are not currently available.

Aviv developed a trans-laryngeal air-pulse stimulator [16, 17]. Unfortunately, the positive intra-rater reproducibility of his test has not been reproduced among less expert raters (i.e. otolaryngology junior resident [18]), and the inter-rater reproducibility is poor [18]. Hammer developed an air-pulse stimulator with improved air-pulse pressure and duration reliabilities [19]. However, reliability studies of Hammer's device have been scarce.

Recently, a new laryngo-pharyngeal esthesiometer (LPEER) was developed, which includes an air pulse generator and an endoscopic laser rangefinder, with the goal of resolving the reliability problems of previous devices [20]. The study protocol also included measurements of laryngeal-adductor reflex threshold (LART), cough reflex threshold (CRT), and gag reflex threshold (GRT) [20, 21]. Preliminary tests revealed promising results, but we do not know whether the technological aids of the LPEER are sufficient to improve the reliability of the LPMS evaluations.

To validate the LPMS evaluations using the LPEER, we compared the results of expert and novel raters in a prospective cohort of patients. We determined the median values of the laryngo-pharyngeal reflex thresholds of aspiration patients and healthy patients and defined cut-off points at which measurement errors should be suspected when performing the test.

Materials and methods

Study population

We prospectively and consecutively recruited a cohort of patients to assess the reliability of the LPEER from two tertiary care university hospitals. The inclusion criteria were a patient age of 18 years or older and stroke with oropharyngeal dysphagia; volunteers without dysphagia served as controls. The exclusion criteria were respiratory failure, bleeding diathesis and anticoagulant therapy. The criteria for removing a recruited patient from the study were any degree of epistaxis or severe discomfort.

The institutional review board of each recruitment centre approved the protocol, all participants provided written informed consent, and the study adhered to good clinical practices. Patient enrolment was performed from December 30, 2013 through September 19, 2014.

Tests

The patients underwent a standard clinical evaluation by a Speech Language Pathologist (SLP) with 7 years of

³ Fundacion Neumologica Colombiana, Bogota, DC, Colombia

⁴ School of Sciences, University of Navarra, Pamplona, Navarra, Spain

⁵ Rehabilitation Department, Clinica Universidad de La Sabana, La Sabana University School of Medicine, Chia, Cundinamarca, Colombia

⁶ University of La Sabana School of Engineering, Chia, Cundinamarca, Colombia

⁷ Otorhinolaryngology Department, University Clinic of Navarra, Pamplona, Navarra, Spain

experience in dysphagia [21]. The clinical evaluation included a validated Spanish version of the EAT-10 [22], the determination of the Rankin Scale score [23] for stroke patients and administration of the Glasgow Coma Scale for patients with any abnormal level of consciousness [24].

The LPMS evaluation consisted of measurements of the LART, CRT and GRT by an endoscopist (rater) [21, 25] using the LPEER, which was connected to a conventional fibre bronchoscope (Pentax FB-10V, Pentax of America, Montvale, NJ, USA) with a working channel of 1.2 mm internal diameter. The bronchoscope was also connected to a video system (Pentax PSV-4000, Pentax of America, Montvale, NJ, USA), a light source (Pentax LH-150PC, Pentax of America, Montvale, NJ, USA) and a computer (Samsung RV420 Core-i5, Samsung Electronics, Suwon, South Korea) for image processing and recording of the entire test [20]. The bronchoscope was lubricated with a water-soluble gel (without any anaesthetics), introduced through the nasal cavity of the patient placed in a seated or semi-recumbent position and advanced into the pharynx. Each rater determined the sensory thresholds at predetermined points of the laryngo-pharyngeal tract: the LART and CRT were measured at the aryepiglottic fold at a point between the corniculate and cuneiform cartilages, while the GRT was explored at the lateral wall of the pharynx at a point lateral to the epiglottis. In preliminary observations [20], it was noted that these particular sites elicit more consistent reflexes. The details of the reflex threshold determinations are published elsewhere [21].

The LART was measured via a series of air pulses of 100 ms duration that decreased in intensity from 0.7 to 0.04 mN [16, 17, 19, 21]. The CRT and GRT were explored via a series of air pulses of 1000 ms duration that increased in intensity from 0.8 to 16.5 mN. Each reflex threshold was defined as the minimum air-pulse intensity that elicited the corresponding reflex. When the repeated reflex threshold measurements were different, the true threshold was the lowest air-pulse stimulus that elicited such a reflex.

We measured air-pulse intensity in mN rather than mmHg due to the geometric characteristics of air pulses, and to compare LPMS with the esthesiometry of other organs [20, 26, 27].

For the LART measurement, the air pulses were set to decrease in intensity because in a preliminary group of subjects [20], it was found that starting with a supra-threshold stimulus helped identify the normal reflex, which otherwise might go unnoticed by a non-experienced rater. The range of stimulus intensities used for the LART measurement was well tolerated and did not induce patient discomfort, except in those with laryngeal hypersensitivity. For members of the latter group who coughed or gagged when stimulated with air pulses of 0.7 mN, we started the air pulse series for the LART measurement at 0.4 mN.

The GRT and CRT were measured by administering air pulses that were set to increase in intensity. These reflexes are very clear and easily detected by any rater; setting the air pulse series to measure them at increasing intensities allowed the rater to stop administering the air pulses once the gag or cough reflex was elicited, to decrease patient discomfort produced by repetition of these reflexes.

An expert and a novel rater determined the LART, CRT and GRT during the same endoscopic procedure but at different moments, sequentially and randomly. We used a table of random numbers to establish this order such that each rater was first and second for similar numbers of evaluations. Raters performed two measurements of each reflex threshold on the right and left side of the laryngo-pharyngeal tract in each patient in the following order: GRT, CRT, and LART. The expert rater was a pulmonologist with 9 years' experience in the fibre-optic endoscopic evaluation of swallowing with sensory test (FEESST); this rater performed the LPMS evaluation with the 67 patients who were finally included in this study (34 with dysphagia due to stroke and 33 without dysphagia; see Table 1). These 67 patients also received the LPMS evaluation from a novel rater, who was a physician with less than 6 months of experience with the FEESST. We had several physicians serving as novel raters: a pulmonologist or pulmonary fellow for seven patients, a general practitioner for six patients or a medical intern undertaking a 6-month rotation in pulmonary medicine for 54 patients. The novel rater received a 1 month training including theoretical and practical sessions (using manikins and humans), upper airway endoscopy and FEESST training. During measurement, the air pulses were identified by a number instead of the air-pulse force (for blinding purposes), and the numbers were blindly replaced by the air-pulse force at the end of the study.

After the sensory evaluation the expert rater with the assistance of a SLP performed a fibre-optic endoscopic evaluation of swallowing (FEES) in all patients, which was considered the reference standard for oropharyngeal dysphagia (the study population was divided according to FEES in patients with stroke and dysphagia and patients without dysphagia). During FEES, patients were tested with four food consistencies (i.e. pure, thick liquid, solid, and thin liquid) according to the standard FEES protocol [21, 25]. All food was green coloured for contrast.

The safety of swallowing was evaluated during the FEES by monitoring for penetration (i.e. the entrance of material into the laryngeal vestibule), residues (i.e. the presence of material on the pharynx after swallowing), aspiration (i.e. the entrance of material below the vocal cords) and premature spillage (i.e. the premature passage of food from the oral to the pharyngeal cavity). The severity of alteration in swallowing was rated based on the consensus of the expert rater and the SLP according to an 8-point

Table 1 General characteristics of the cohort

	Patients without dysphagia	Stroke with dysphagia	Total
No. (%)	33 (49%) ^a	34 (51%) ^a	67
Female, no. (%)	19 (58%) ^b	14 (41%) ^b	33 (49%) ^b
Age, mean (SD)	46 (15.2)	60.8 (17.6)	53.3 (18.2)
Body mass index (kg/m ²), mean (SD)	25.1 (4.10)	25.60 (4.50)	25.30 (4.30)
Comorbidities			
Smoking, no. (%)	12 (36%) ^b	13 (38%) ^b	25 (37%) ^b
COPD, no. (%)	4 (12%) ^b	4 (12%) ^b	8 (12%) ^b
Chronic renal failure, no. (%)	0 (0%) ^b	2 (6%) ^b	2 (3%) ^b
Diabetes, no. (%)	0 (0%) ^b	5 (15%) ^b	5 (7%) ^b
Heart failure, no. (%)	0 (0%) ^b	5 (15%) ^b	5 (7%) ^b
Collagen vascular diseases, no. (%)	1 (3%) ^b	1 (3%) ^b	2 (3%) ^b
GERD, no. (%)	13 (39%) ^b	12 (35%) ^b	25 (37%) ^b
Severity scales			
Glasgow Coma Scale: median (IQR)	N/A	15 (13–15)	
NIHSS at admission: median (IQR)	N/A	10 (2–18)	
Swallowing safety alterations			
Penetration (any food consistency), no. (%)	0 (0%) ^b	19 (56%) ^b	19 (28%) ^b
Aspiration (any food consistency), no. (%)	0 (0%) ^b	12 (35%) ^b	12 (18%) ^b
Penetration-aspiration (Rosenbek Scale): median (IQR)	1.0 (1.0–1.0)	3.0 (1.0–6.5)	1.0 (1.0–3.0)

GERD gastro-oesophageal reflux disease, IQR inter-quartile range, no. number, NIHSS National Institutes of Health Stroke Scale, N/A not applicable

^aPercentage over the entire row

^bPercentage over the entire column

penetration-aspiration scale [28, 29] and the dysphagia severity scale (DSS) [21, 30, 31].

Patients rated the pain, nausea, headache and discomfort experienced during the FEESST on a scale from 0 to 10, with 0 corresponding to the absence of symptoms and 10 corresponding to the maximum intensity of the symptom that the patient had ever experienced.

Statistical analysis

We performed Kolmogorov–Smirnov tests to determine whether the quantitative variables were normally distributed. We compared the normally distributed variables with *t*-tests and compared the non-normally distributed variables with Mann–Whitney U tests. For all of the statistical analyses, differences were considered as statistically significant at $P < 0.05$ (two-tailed).

To assess the LART, CRT and GRT reliabilities, we calculated the intra- and inter-rater intraclass correlation coefficients (ICCs) and Spearman correlation coefficients (SCCs) with their 95% confidence intervals (95% CIs). We evaluated intra- and inter-rater agreement using Bland–Altman plots of the limits of agreement with their 95% CIs.

The inter-rater comparisons were performed by contrasting the measurements of the expert and novel raters.

For the sample size calculation, we used the equation proposed by Bonnet for the ICC [32]. Based on this equation, for an ICC ≥ 0.7 with a 95% CI width ≤ 0.25 , we required 66 patients.

The statistical analyses were performed using IBM SPSS statistical software, version 20 (Armonk, NY, USA), MedCalc version 14.12.0 (MedCalc Software bvba, Ostend, Belgium), and Microsoft Excel 2007 (Microsoft Corporation, Redmond, WA, USA).

Results

We evaluated 124 patients. Thirty-seven did not meet the inclusion criteria due to dysphagia secondary to conditions other than stroke or an age below 18 years. Two patients met the exclusion criteria: one due to anticoagulant therapy, and one due to respiratory insufficiency secondary to amyotrophic lateral sclerosis. Twelve patients declined consent to participate, and six were excluded during testing (four due to unobtainable thresholds due to continuous movements of the laryngo-pharyngeal tract, one due to severe

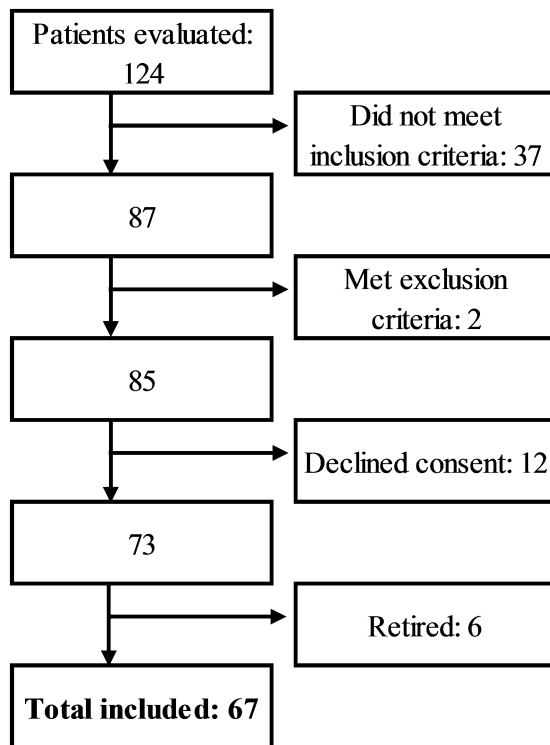


Fig. 1 Enrolment flowchart. Reasons for retirement: discomfort: 1, epistaxis: 1, continuous laryngo-pharyngeal movements impeding threshold exploration: 4

Table 2 Intra- and inter-rater intraclass correlation coefficients

	Intra-rater novel ICC (95% CI)	Intra-rater expert ICC (95% CI)	Inter-rater ICC (95% CI)
LART	0.97 (0.96–0.98)	0.98 (0.97–0.99)	0.87 (0.82–0.91)
CRT	0.96 (0.94–0.97)	0.97 (0.96–0.98)	0.79 (0.72–0.85)
GRT	0.90 (0.86–0.93)	0.95 (0.93–0.97)	0.70 (0.60–0.78)

LART laryngeal-adductor reflex threshold, CRT cough reflex threshold, GRT gag reflex threshold, ICC intraclass correlation coefficient, 95% CI 95% confidence interval

Inter-rater: expert vs. novel rater

Table 3 Intra- and inter-rater Spearman correlation coefficients

	Intra-rater novel SCC (95% CI)	<i>P</i>	Intra-rater expert SCC (95% CI)	<i>P</i>	Inter-rater SCC (95% CI)	<i>P</i>
LART	0.91 (0.87–0.94)	<0.0001	0.91 (0.88–0.94)	<0.0001	0.80 (0.73–0.86)	<0.0001 ^a
CRT	0.96 (0.94–0.97)	<0.0001	0.97 (0.95–0.98)	<0.0001	0.79 (0.72–0.85)	<0.0001 ^a
GRT	0.88 (0.83–0.91)	<0.0001	0.93 (0.91–0.95)	<0.0001	0.70 (0.60–0.78)	<0.0001 ^a

SCC Spearman correlation coefficient, LART laryngeal-adductor reflex threshold, CRT cough reflex threshold, GRT gag reflex threshold, 95% CI 95% confidence interval

^aStatistically significant; inter-rater: expert vs. novel rater

discomfort and one due to epistaxis). Thus, 67 patients were ultimately included in the analysis (Fig. 1).

The patients with stroke and dysphagia were older, were comprised a greater proportion of men and had more comorbidities (Table 1). The stroke patients had moderate strokes according to the NIHSS. The overall severity of dysphagia in the stroke patients according to the Penetration-Aspiration scale was moderate (ranging from mild to severe). Half of the stroke patients exhibited penetration, and one-third exhibited aspiration (Table 1).

Each rater performed four measurements per reflex in each patient (two measurements on each side of the laryngo-pharyngeal tract). We performed a total of 24 measurements per patient, yielding 1608 measurements over all 67 patients in the study.

The intra-rater ICCs for all of the reflex threshold determinations were above 0.90, and the inter-rater ICCs were 0.87 for the LART, 0.79 for the CRT and 0.70 for the GRT (Table 2).

The intra-rater SCCs for all of the reflex thresholds were above 0.88, and the inter-rater SCCs were 0.80 for the LART and CRT and 0.70 for the GRT (all *P* < 0.0001) (Table 3).

The Bland–Altman plots of the limits of agreement revealed mean intra- and inter-rater differences that were close to zero and free of increments in the variabilities at the extremes of the averages, with the exception of a mild trend toward lower values for the GRT in the expert rater measurements compared with those of the novel rater at the lower GRTs (Fig. 2). Ninety-five percent of the inter-rater differences in the LART and CRT were less than the differences in the thresholds between the aspirators and non-aspirators (Fig. 2; Table 4). Regarding the GRT, the intra-rater limits of agreement were lower than the differences in the thresholds between the aspirators and non-aspirators, but the inter-rater limits of agreement were greater (Fig. 2; Table 4).

The patients with aspiration showed higher reflex thresholds; this difference was clinically and statistically significant (Table 4).

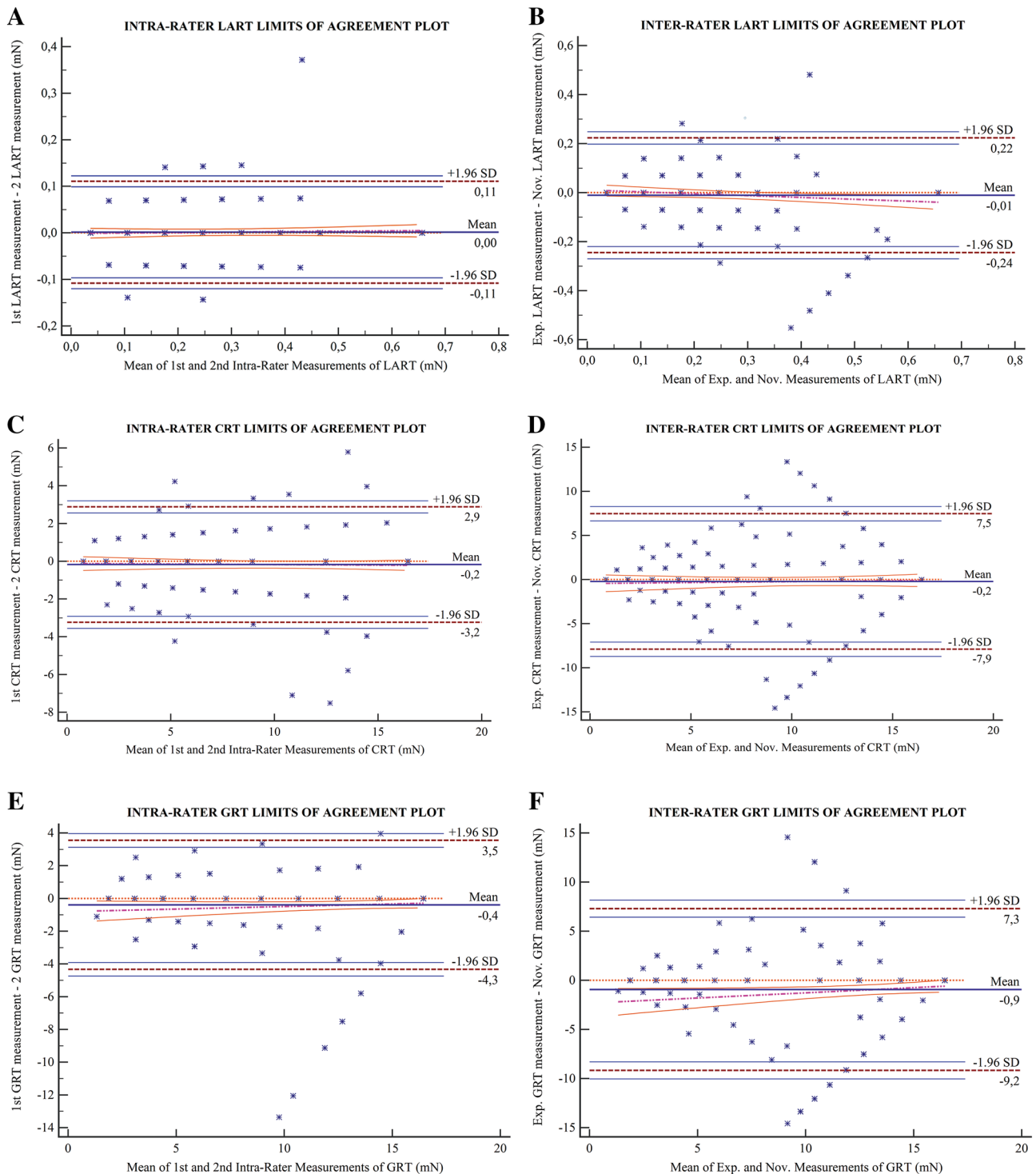


Fig. 2 **a** Intra-rater LART limits of agreement plot; **b** inter-rater LART limits of agreement plot; **c** intra-rater CRT limits of agreement plot; **d** inter-rater CRT limits of agreement plot; **e** intra-rater GRT limits of agreement plot; **f** inter-rater GRT limits of agreement plot.

The median normal value was 0.14 mN for the LART (IQR 0.11–0.24), 4.44 mN for the CRT (IQR 2.63–7.93) and 11.88 mN for the GRT (IQR 4.44–16.44).

LART laryngeal-adductor reflex threshold, CRT cough reflex threshold, GRT gag reflex threshold, mN millinewtons, SD standard deviation, Exp expert, Nov novel

We calculated the 95% CIs of the intra-rater differences in the limits of agreement to establish the upper limits as cut-offs for the identification of outliers; the resulting

Table 4 Reflex thresholds according to aspiration status

	Patients without dysphagia	Aspiration	<i>P</i>
Sum of LART (mN), median (IQR)	0.28 (0.22–0.48)	1.31 (0.49–1.31)	0.005 ^a
Sum CRT (mN), median (IQR)	8.88 (5.26–15.86)	32.88 (30.90–32.88)	<0.001 ^a
Sum GRT (mN), median (IQR)	23.75 (8.88–32.88)	32.88 (32.88–32.88)	<0.001 ^a

LART laryngeal-adductor reflex threshold, *CRT* cough reflex threshold, *GRT* gag reflex threshold, *mN* millinewtons, *IQR* inter-quartile range

The sum of the reflexes corresponds to the sum of the right and left minimum thresholds of each reflex as measured by the expert rater (to determine the median value of each reflex threshold, the sum must be divided by 2; e.g. LART in controls = 0.28/2 = 0.14 mN)

^aStatistically significant difference by Mann–Whitney *U* test (two-tailed)

cut-offs were 0.12 mN for the LART, 3 mN for the CRT and 4 mN for the GRT.

The patients did not report pain during the exam, and the majority rated their discomfort as mild to moderate (median discomfort: 4/10; IQR 3–6). We observed no cases of syncope, pre-syncope or laryngospasm and no requirements for emergency room care or hospitalization due to adverse events.

Discussion

In the present study of LPEER reliability, we observed ICCs that were within the range of excellent reliabilities for all of the reflex thresholds measurements with the exception of the inter-rater GRT; however, this latter measure also achieved substantial agreement [33, 34]. The SCCs for all of reflex threshold measurements were within the range of strong reliabilities [35]. The Bland–Altman limits of agreement plots revealed that the intra- and inter-rater differences were close to zero and free of trends related to the averages and signs of bias with the exception of a mild trend toward lower GRT values in the expert rater measurements at lower GRT values. The widths of the limits of agreement of the LART and CRT were adequate for comparing patients with and without dysphagia.

Our intra-rater ICC and SCC results from the expert rater are similar to those reported by Aviv [17] and better than those of Cunningham [18]; however, we also observed excellent reliability in the results from the novel raters. Furthermore, our results revealed excellent inter-rater reliabilities in the comparisons of the expert and novel raters, which have not been reported with previous devices. We chose two raters at opposite ends of the expertise spectrum to include the full range of variability existing in real clinical practice: the difference of measurement between raters with more closely aligned degrees of expertise would be lower than what we found, yielding even better agreement when using the Bland–Altman limit of agreement method. To gain worse results than ours when deploying said method, a greater difference in rater expertise than ours

would be needed, which is unlikely to be found in real clinical practice. These results demonstrate the efficacies of the technological aids that have been introduced to the LPEER [20] in improving the reliability of LPMS evaluation.

It was unclear the extent to which other clinical factors, e.g. the viscosities of secretions and continuous movements of the laryngo-pharyngeal tract, would affect the reliability of LPMS evaluation. However, the excellent reliabilities obtained for the tests that were performed on patients with a wide range of deglutition alterations (Penetration-Aspiration score ranging from 1 to 7) [28] suggest that these other factors were either not clinically relevant or were indirectly controlled for by the LPEER technological improvements [20]. Indeed, the LPEER enables control of the distance and angle of the stimulus delivery over the target surface, similarly to esthesiometers that have been designed for other organs [36].

To the best of our knowledge, this study provides one of the most comprehensive evaluations of the LPMS in terms of reliability, including assessments of the ICCs, SCCs and limits of agreement. We validated a device and method for exploring the LART, CRT and GRT. Our results support the utilization of the LART and CRT in normal patients and patients with dysphagia secondary to stroke. However, the limits of agreement for the inter-rater GRT were greater than the differences between the patients with and without dysphagia. These findings are consistent with those of previous studies reporting a lower reliability of this reflex than those of other reflexes in dysphagic patients [37, 38]. Although the GRT exhibited less utility for the dysphagic patients, our study does not allow us to rule out its utility for diseases involving lower GRTs, such as motor neuron diseases [39, 40] and other laryngo-pharyngeal hypersensitivity states.

All reflexes exhibited significantly higher thresholds in the dysphagic stroke patients, which highlights the sensory compromises that are typical of this condition [37, 41]. Similar findings have previously been reported for the LART [37, 41]. While there have been prior reports of cough reflex compromise in stroke patients assessed with chemical stimuli [42, 43], none have used standard

mechanical stimuli like those provided by the LPEER for such assessments. These sensory compromises of the airway-protecting reflexes most likely contribute to the development of dysphagia in stroke patients and to the increased risk of pneumonia in said individuals [44, 45]. Therefore, they could be used to identify patients at high risk of pneumonia who would benefit from interventions to reduce pneumonia incidence, including vaccination, oral care, diet modification and the better selection of patients for gastrostomy. In addition, there are experimental interventions that are undergoing evaluations in clinical trials, with the goal of improving the laryngo-pharyngeal sensory compromise observed in stroke patients to improve their deglutition alterations [13, 46, 47], and these trials could benefit from our quantitative method of LPMS measurement for the evaluation of these interventions' efficacy.

Our normal LART value was 0.14 mN, this value corresponds to 2.5 mmHg in Aviv's system of measurement [20, 21]. This normal value is consistent with Aviv's work [16, 48] and is also comparable to the results of the study by Grushka [27] in which a tactile sensory threshold of 14.9 mg (equivalent to 0.15 mN) was observed at the most sensitive point of the tongue.

To rule out any measurement errors in clinical explorations of the LPMS, each patient should undergo at least two measurements per reflex per side of the laryngo-pharyngeal tract. Whenever differences between the measurements that exceed the outlier limits that we have reported (e.g. 0.12 mN for the LART, 3 mN for the CRT and the GRT) are observed, the rater should suspect measurement error and perform additional measurements.

We did not observe clinically relevant adverse events, and our patients reported only mild to moderate discomfort, which is consistent with previous FEESST safety reports [49].

Conclusion

The explorations of the LARTs and CRTs with the LPEER showed excellent intra- and inter-rater reliability. The GRT exploration exhibited substantial reliability, but the larger width of the inter-rater limits of agreement of this test could limit its application in stroke patients.

Acknowledgements This study was funded by grants from Colciencias, Colombia and the University of La Sabana Chia, Colombia.

Author contributions LFG-C is the guarantor of the content of the manuscript, including the data and analysis. LFG-C contributed substantially to the study design, recruitment and evaluation of patients, data acquisition, analysis and interpretation, the manuscript writing and critical manuscript revision and provided final approval of the version to be published. FR, AMG, NC and NP contributed substantially

to the study design, recruitment and evaluation of patients, data acquisition, analysis and interpretation and the critical manuscript revision and provided final approval of the version to be published. JB, SF, LMA-O and WDM contributed substantially to the study design, data analysis and interpretation and the critical manuscript revision and provided final approval of the version to be published.

Compliance with ethical standards

Funding information This research project was funded by Colciencias, Colombia and by the University of La Sabana, Chia, Colombia.

Conflict of interest L.F. Giraldo-Cadavid, J. Burguete, L.M. Agudelo-Otalora and S. Fernandez have a patent granted by the Colombian's Superintendencia de Industria y Comercio to the Laryngo-pharyngeal Endoscopic Esthesiometer and Range Finder (LPEER) (Number: 13-121843-00003-0000), a PCT (International Application No.: PCT/IB2014/061063) and a US Patent in process. University of La Sabana (Colombia) is the owner of the financial rights for this patent, and the authors own the intellectual property over this patent. F. Rueda, A.M. Galvis, N. Castaneda, W.D. Moscoso and N. Paez have no conflict of interests related to the subject of this manuscript.

References

1. Aviv JE, Kim T, Sacco RL, Kaplan S, Goodhart K, Diamond B (1998) FEESST: a new bedside endoscopic test of the motor and sensory components of swallowing. *Ann Otol Rhinol Laryngol* 107(5):378–387
2. Nguyen AT, Jobin V, Payne R, Beaugard J, Naor N, Kimoff RJ (2005) Laryngeal and velopharyngeal sensory impairment in obstructive sleep apnea. *Sleep* 28(5):585–593
3. Ludlow CL (2011) Central nervous system control of interactions between vocalization and respiration in mammals. *Head Neck* 33(Suppl 1):S21–S25
4. Morice AH, Millqvist E, Belvisi MG, Bieksiene K, Birring SS, Chung KF, Dal Negro RW, Dicpinigaitis P, Kantar A, McGarvey LP, Pacheco A, Sakalauskas R, Smith JA (2014) Expert opinion on the cough hypersensitivity syndrome in respiratory medicine. *Eur Respir J* 44(5):1132–1148
5. Kertscher B, Speyer R, Fong E, Georgiou AM, Smith M (2015) Prevalence of oropharyngeal dysphagia in the Netherlands: a telephone survey. *Dysphagia* 30(2):114–120
6. Torres A, Peetermans WE, Viegi G, Blasi F (2013) Risk factors for community-acquired pneumonia in adults in Europe: a literature review. *Thorax* 68(11):1057–1065
7. Komiya K, Ishii H, Umeki K, Mizunoe S, Okada F, Johkoh T, Kadota J (2013) Impact of aspiration pneumonia in patients with community-acquired pneumonia and healthcare-associated pneumonia: a multicenter retrospective cohort study. *Respirology (Carlton, Vic)* 18 (3):514–521
8. Lanspa MJ, Jones BE, Brown SM, Dean NC (2013) Mortality, morbidity, and disease severity of patients with aspiration pneumonia. *J Hosp Med Off Publ Soc Hosp Med* 8(2):83–90
9. Peppard PE, Young T, Barnet JH, Palta M, Hagen EW, Hla KM (2013) Increased prevalence of sleep-disordered breathing in adults. *Am J Epidemiol* 177(9):1006–1014
10. Wang X, Ouyang Y, Wang Z, Zhao G, Liu L, Bi Y (2013) Obstructive sleep apnea and risk of cardiovascular disease and all-cause mortality: a meta-analysis of prospective cohort studies. *Int J Cardiol* 169(3):207–214
11. Johansson A, Millqvist E, Nordin S, Bende M (2006) Relationship between self-reported odor intolerance and

- sensitivity to inhaled capsaicin: proposed definition of airway sensory hyperreactivity and estimation of its prevalence. *Chest* 129(6):1623–1628
12. Vertigan AE, Bone SL, Gibson PG (2014) Development and validation of the Newcastle laryngeal hypersensitivity questionnaire. *Cough* 10(1):1
 13. Rofes L, Arreola V, Lopez I, Martin A, Sebastian M, Ciurana A, Clave P (2013) Effect of surface sensory and motor electrical stimulation on chronic poststroke oropharyngeal dysfunction. *Neurogastroenterol Motil* 25(11):701–888
 14. Pauloski BR, Logemann JA, Rademaker AW, Lundy D, Sullivan PA, Newman LA, Lazarus C, Bacon M (2013) Effects of enhanced bolus flavors on oropharyngeal swallow in patients treated for head and neck cancer. *Head Neck* 35(8):1124–1131
 15. Gibson P, Wang G, McGarvey L, Vertigan AE, Altman KW, Birring SS (2016) Treatment of unexplained chronic cough: chest guideline and expert panel report. *Chest* 149(1):27–44
 16. Aviv JE, Martin JH, Kim T, Sacco RL, Thomson JE, Diamond B, Close LG (1999) Laryngopharyngeal sensory discrimination testing and the laryngeal adductor reflex. *Ann Otol Rhinol Laryngol* 108(8):725–730
 17. Aviv JE, Martin JH, Keen MS, Debell M, Blitzer A (1993) Air pulse quantification of supraglottic and pharyngeal sensation: a new technique. *Ann Otol Rhinol Laryngol* 102(10):777–780
 18. Cunningham JJ, Halum SL, Butler SG, Postma GN (2007) Intra-observer and interobserver reliability in laryngopharyngeal sensory discrimination thresholds: a pilot study. *Ann Otol Rhinol Laryngol* 116(8):582–588
 19. Hammer MJ (2009) Design of a new somatosensory stimulus delivery device for measuring laryngeal mechanosensory detection thresholds in humans. *IEEE Trans Biomed Eng* 56(4):1154–1159
 20. Giraldo-Cadavid LF, Agudelo-Otalora LM, Burguete J, Arbulu M, Moscoso WD, Martínez F, Ortiz AF, Diaz J, Pantoja JA, Rueda-Arango AF, Fernández S (2016) Design, development and validation of a new laryngo-pharyngeal endoscopic esthesiometer and range-finder based on the assessment of air-pulse variability determinants. *Biomed Eng Online* 15(1):1–23
 21. Giraldo-Cadavid LF (2013) Fiberoptic endoscopic evaluation of swallowing with sensory test (FEESST) protocol for the validation of a laryngo-pharyngeal endoscopic esthesiometer and rangefinder (LPEER). (Dec.):1–24. <https://upperairwaylfgiraldo.wordpress.com/about/>. Accessed: 22 Aug 2015
 22. Burgos R, Sarto B, Segurolo H, Romagosa A, Puiggrós C, Vazquez C, Cardenas G, Barcons N, Araujo K, Pérez C (2012) Traducción y validación de la versión en español de la escala EAT-10 (Eating Assessment Tool-10) para despistaje de la disfagia. *Nutrición Hospitalaria* 27(6):2048–2054
 23. van Swieten JC, Koudstaal PJ, Visser MC, Schouten HJ, van Gijn J (1988) Interobserver agreement for the assessment of handicap in stroke patients. *Stroke* 19(5):604–607
 24. Middleton PM (2012) Practical use of the Glasgow Coma Scale; a comprehensive narrative review of GCS methodology. *Australas Emerg Nurs J AENJ* 15(3):170–183
 25. Aviv JE (2000) Prospective, randomized outcome study of endoscopy versus modified barium swallow in patients with dysphagia. *Laryngoscope* 110(4):563–574
 26. Golebiowski B, Papas E, Stapleton F (2011) Assessing the sensory function of the ocular surface: implications of use of a non-contact air jet aesthesiometer versus the Cochet-Bonnet aesthesiometer. *Exp Eye Res* 92(5):408–413
 27. Grushka M, Sessle BJ, Howley TP (1987) Psychophysical assessment of tactile, pain and thermal sensory functions in burning mouth syndrome. *Pain* 28(2):169–184
 28. Rosenbek JC, Robbins JA, Roecker EB, Coyle JL, Wood JL (1996) A penetration-aspiration scale. *Dysphagia* 11(2):93–98
 29. Kelly AM, Drinnan MJ, Leslie P (2007) Assessing penetration and aspiration: how do videofluoroscopy and fiberoptic endoscopic evaluation of swallowing compare? *Laryngoscope* 117(10):1723–1727
 30. Gates J, Hartnell GG, Gramigna GD (2006) Videofluoroscopy and swallowing studies for neurologic disease: a Primer. *RadioGraphics* 26(1):e22
 31. Waxman MJ, Durfee D, Moore M, Morantz RA, Koller W (1990) Nutritional aspects and swallowing function of patients with Parkinson's disease. *Nutr Clin Pract* 5(5):196–199
 32. Bonett DG (2002) Sample size requirements for estimating intraclass correlations with desired precision. *Stat Med* 21(9):1331–1335
 33. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* 33(1):159–174
 34. Sharrack B, Hughes RAC, Soudain S, Dunn G (1999) The psychometric properties of clinical rating scales used in multiple sclerosis. *Brain* 122(1):141–159
 35. Sheskin DJ (2003) Measures of Association/Correlation. In: *Handbook of parametric and nonparametric statistical procedures*, 3rd edn. CRC Press, pp 943–1120
 36. Teson M, Calonge M, Fernandez I, Stern ME, Gonzalez-Garcia MJ (2012) Characterization by Belmonte's gas esthesiometer of mechanical, chemical, and thermal corneal sensitivity thresholds in a normal population. *Invest Ophthalmol Vis Sci* 53(6):3154–3160
 37. Aviv JE, Martin JH, Sacco RL, Zagar D, Diamond B, Keen MS, Blitzer A (1996) Supraglottic and pharyngeal sensory abnormalities in stroke patients with dysphagia. *Ann Otol Rhinol Laryngol* 105(2):92–97
 38. Davies AE, Kidd D, Stone SP, MacMahon J (1995) Pharyngeal sensation and gag reflex in healthy subjects. *Lancet* 345(8948):487–488
 39. Hughes TA, Wiles CM (1996) Palatal and pharyngeal reflexes in health and in motor neuron disease. *J Neurol Neurosurg Psych* 61(1):96–98
 40. Ross MA, Miller RG, Berchert L, Parry G, Barohn RJ, Armon C, Bryan WW, Petajan J, Stromatt S, Goodpasture J, McGuire D (1998) Toward earlier diagnosis of amyotrophic lateral sclerosis: revised criteria. *rhCNTF ALS Study Group. Neurology* 50(3):768–772
 41. Onofri SM, Cola PC, Berti LC, da Silva RG, Dantas RO (2014) Correlation between laryngeal sensitivity and penetration/aspiration after stroke. *Dysphagia* 29(2):256–261
 42. Stephens RE, Addington WR, Widdicombe JG (2003) Effect of acute unilateral middle cerebral artery infarcts on voluntary cough and the laryngeal cough reflex. *Am J Phys Med Rehabil* 82(5):379–383
 43. Ward K, Seymour J, Steier J, Jolley CJ, Polkey MI, Kalra L, Moxham J (2010) Acute ischaemic hemispheric stroke is associated with impairment of reflex in addition to voluntary cough. *Eur Respir J* 36(6):1383–1390
 44. Smith Hammond CA, Goldstein LB, Horner RD, Ying J, Gray L, Gonzalez-Rothi L, Bolser DC (2009) Predicting aspiration in patients with ischemic stroke: comparison of clinical signs and aerodynamic measures of voluntary cough. *Chest* 135(3):769–777
 45. Nakajoh K, Nakagawa T, Sekizawa K, Matsui T, Arai H, Sasaki H (2000) Relation between incidence of pneumonia and protective reflexes in post-stroke patients with oral or tube feeding. *J Intern Med* 247(1):39–42
 46. Poorjavad M, Talebian Moghadam S, Nakhostin Ansari N, Daemi M (2014) Surface electrical stimulation for treating swallowing disorders after stroke: a review of the stimulation intensity levels and the electrode placements. *Stroke Res Treat* 918057

47. Veldman MP, Maffioletti NA, Hallett M, Zijdwind I, Hortobágyi T (2014) Direct and crossed effects of somatosensory stimulation on neuronal excitability and motor performance in humans. *Neurosci Biobehav Rev* 47:22–35
48. Aviv JE, Kim T, Thomson JE, Sunshine S, Kaplan S, Close LG (1998) Fiberoptic endoscopic evaluation of swallowing with sensory testing (FEESST) in healthy controls. *Dysphagia* 13(2):87–92
49. Aviv JE, Murry T, Zschommler A, Cohen M, Gartner C (2005) Flexible endoscopic evaluation of swallowing with sensory testing: patient characteristics and analysis of safety in 1340 consecutive examinations. *Ann Otol Rhinol Laryngol* 114(3):173–176