CrossMark

Validation of the secretion severity rating scale

Petra Pluschinski[1] · Eugen Zaretsky[1] · Timo Stöver[2] · Joseph Murray[3] ·
Robert Sader[4] · Christiane Hey[1]

Abstract Accumulation of secretions within the hypopharynx, aditus laryngis, and trachea is one characteristic of severe dysphagia and is of high clinical and therapeutic relevance. For the graduation of the secretion severity level, a secretion scale was provided by Murray et al. in 1996. The purpose of the study presented here is the validation of this scale by analyzing the intra-rater and inter-rater reliability as well as concurrent validity. For examination of reliability and validity, a reference standard was defined by two expert clinicians who reviewed 40 video recordings of fiberendoscopic swallowing evaluations, with 10 videos for each severity grade. These videos were rated and rerated independently and blinded by 4 ENT-residents with an interval of 4 weeks. Both the intra-rater (Kendall's $\tau > 0.847$***) and inter-rater reliability (Kendall's $W > 0.951$***) were highly significant and can be considered good or very good. Correlation of the median of all ratings with the reference standard was close to the highest possible value 1 ($\tau = 0.984$***). The scale was proved to be a reliable and valid instrument for graduation of one of the principal symptoms of oropharyngeal dysphagia and is recommended as an evidence-based instrument for standardized fiberoptic endoscopic evaluation of swallowing.

Keywords Secretion scale · Validity · FEES® · Swallowing disorder · Dysphagia

P. Pluschinski and E. Zaretsky have contributed equally.

✉ Christiane Hey
Christiane.Hey@med.uni-marburg.de

1 Department of Phoniatrics and Pediatric Audiology,
University Hospital of Marburg, Baldingerstr. 1,
35032 Marburg, Germany

2 Department of Oto-Rhino-Laryngology, Head and Neck
Surgery, University Hospital of Frankfurt/Main, Theodor-
Stern-Kai 7, 60590 Frankfurt/Main, Germany

3 Audiology and Speech Pathology Service, VA Ann Arbor
Healthcare System, Fuller Road 2215, 48105 Ann Arbor, MI,
USA

4 Department for Oral, Cranio-Maxillofacial and Facial Plastic
Surgery, University Hospital of Frankfurt/Main, Theodor-
Stern-Kai 7, 60590 Frankfurt/Main, Germany

Introduction

Accumulated secretions within the pharyngolarynx and trachea are typical in patients with severe oropharyngeal dysphagia. Recent studies revealed that colonization by microbial pathogens in patients suffering from dental plaques is common [1, 2] and could lead to pneumonia due to secretion aspiration, especially in long-term ventilated patients. Accordingly, improved oral hygiene has been proved to decrease pneumonia in about 40 % of long-term ventilated patients [3].

Furthermore, the presence of secretions in pharyngolarynx and trachea indicates impaired swallowing [4] with reduced ability of secretion clearance as a possible consequence of diminished swallowing frequency or strength [5, 6]. Hence, such secretions are of a high relevance for clinical and scientific use with the need for unified and standardized documentation.

The fundamental idea of using accumulated secretions as an indicator for increased risk of aspiration [7] inspired further research. Several studies focused on secretion observation as a sign for impaired swallowing in different underlying pathologies, such as acute stroke [8], Parkinson's disease [9], and amyotrophic lateral sclerosis [5]. Ota

Springer

et al. [8] showed that the higher the volume of secretions, the higher the severity level of the swallowing disorder in stroke patients quantified by the frequency of aspiration. Moreover, they demonstrated a significantly increased risk of aspiration pneumonia with increased accumulation of secretions. Comparable findings were made by Kang et al. [10] in patients with brain injury and tracheostomy. Especially for the latter, the correct determination of secretion is important in respect of aspiration risk and decannulation, as Warnecke et al. [11] demonstrated for decannulation time and safety using fiberoptic endoscopic evaluation of swallowing. Beyond the correlation of secretion accumulation with aspiration and aspiration pneumonia, Takahashi et al. [12] determined an increased risk of malnutrition, at least in elderly institutionalized nursing home residents in a three months follow-up, irrespective of patients' volitional secretion clearing ability.

Even with this collection of studies that emphasize the significance of secretion evaluation and the persistent actuality of these clinical symptoms in dysphagic patients, the secretion severity rating scale, published by Murray et al. [7], has yet to be validated.

Therefore, the aim of this study presented here is proving the quality criteria intra-rater and inter-rater reliability as well as validity for the four-point secretion severity rating scale by Murray et al. [7].

Materials and methods

For the validation of the four-point secretion severity rating scale by Murray, 40 sequences of digitalized fiberoptic endoscopic evaluation of swallowing (FEES®) recordings, 10 for each grade, were chosen retrospectively from over 1000 FEES® examinations performed between 2009 and 2013 using the protocol developed by Langmore [13]. The representative videos were selected and defined as reference standard for validation analysis by two dysphagia experts from a pool of 35 patients. The selected video recordings represent the full view of the hypopharynx, including patients' spontaneous or cued clearing attempts with a maximum duration of 45 s.

Four blinded raters performed an initial rating of the 40 video sequences with a second rating 4 weeks later. The sequence viewing order was randomized for each session. All four raters were ENT-residents experienced in FEES® examination, two of them with more than 3 years of experience and two with less than 3 years of experiences. The videos were presented on a standard 17-inch PC monitor with the freeware player VirtualDub 1.9.11 (http://www.virtualdub.org). All raters had the opportunity to analyze the recordings in the default reproduction mode of real-time playback, frame-by-frame playback, or in slow-

motion. There was no limitation in viewing time for each individual sequence.

The scale used for the reliability and validity testing in this study is shown in Table 1. Bidirectionality of scale score "2" from the original scale was assumed and is stated as such in the table.

Statistical analysis

The intra-rater and inter-rater reliability and the validity according to the defined reference standard were analyzed. Due to the lack of normal distribution (Kolmogorov–Smirnov test: $ps < 0.05$), non-parametric tests were used. Both measures of association and difference were analyzed as even highly significant correlations cannot exclude significant differences. Therefore, Kendall's tau correlations and the Wilcoxon test were used to determine the intra-rater reliability and Kendall's W and Friedman test for inter-rater reliability.

The ratings from the first and second sessions were cross tabulated to evaluate differences in the distribution of the gradings with the Chi-square test. Moreover, it was of interest if any of the four grades introduced difficulty to the raters, resulting in more discrepancies between the two rating sessions. The heterogeneity of the ratings within each secretion scale grade was assessed by the Friedman test for both the first and second rating sessions.

To examine the concurrent validity, the median of all ratings was calculated over all four raters and both rating sessions. The agreement of this median with the reference standard was examined by means of Kendall's tau correlation. The differences between the median and the reference standard were analyzed by the Wilcoxon test for two paired samples.

With an ordinal regression, the association of the rating distribution with the following independent variables was determined: (1) four raters, (2) first and second rating sessions, (3) experience in FEES® examination, and (4) reference standard. The percentage of the explained variance was quantified by Nagelkerke's pseudo-R^2.

All calculations were performed using SPSS 21 (International Business Machines Corporation, Armonk, USA).

Results

The correlations of the intra-rater reliability were highly significant for all raters (rater 1: $\tau = 0.984$, rater 2: $\tau = 0.889$, rater 3: $\tau = 0.936$, rater 4: $\tau = 0.847$; $ps < 0.001$, $Ns = 40$). The Wilcoxon test identified no significant differences between the ratings of all four raters in the first and second sessions ($ps > 0.05$).

Table 1 Four-point secretion severity rating scale by Murray et al. [7]

0	Most normal rating. No visible secretions anywhere in the hypopharynx or some transient bubbles visible in the valleculae and pyriform sinuses. These secretions were not bilateral or deeply pooled
1	Any secretions evident upon entry or following a dry swallow in the channels surrounding the laryngeal vestibule that were bilaterally represented or deeply pooled. This rating would include cases where there is a transition in the accumulation of secretions during the observation segment. A subject could start with no visible secretions but accumulate secretions in an amount great enough to be bilaterally represented or deeply pooled. Likewise, a subject would be rated as a "1" if initially presenting with deeply pooled bilateral secretions and ending the observation segment with no visible secretions
2	Any secretions that changed from a "1" rating to a "3" rating, respectively, from a "3" rating to a "1" rating during the observation period
3	Most severe rating. Any secretions seen in the area defined as the laryngeal vestibule. Pulmonary secretions were included if they were not cleared by swallowing or coughing at the close of the segment

Similarly, the correlations of the inter-rater reliability were highly significant (rating 1: $W = 0.951$, rating 2: $W = 0.961$; $ps < 0.001$), without significant differences between the raters in the Friedman test ($ps > 0.05$).

The distribution of rating values in the first and second rating sessions is shown in Table 2. The overall agreement between two rating sessions was 88.9 %, with a significant difference in the value distribution in two rating sessions for all four raters ($\chi^2_{(9)} = 350.34$, $p < 0.001$).

According to the Friedman test for paired samples, no statistically significant differences were identified within each secretion scale grade in either the first or second rating sessions ($ps > 0.05$).

For the examination of the concurrent validity, the median of all eight ratings (4 raters × 2 rating sessions) correlated highly significantly with the reference standard ($\tau = 0.984$, $p < 0.001$).

The ordinal regression with rating values as dependent variable explained 93 % of the variance (Nagelkerke's pseudo-$R^2 = 0.931$; Pearson $\chi^2_{(86)} = 2204.43$, $p < 0.001$). Only one independent variable appeared to be statistically significant, the "reference standard" ($p < 0.001$). The influence of the other independent variables "first and second rating sessions," "experience in FEES® examination (less than 3 years versus more than 3 years)," and "raters" did not reach statistical significance.

Table 2 Cross classification of the scores of the first and second rating sessions for all four raters

Murray	Second rating session					
	0	1	2	3	Total	%
First rating session						
0	33	2	1	0	36	22.5
1	6	36	1	0	43	26.9
2	0	2	34	3	39	24.4
3	0	1	2	39	42	26.3
Total	39	41	38	42	160	
%	24.4	25.6	23.8	26.3		100

Discussion

The four-point secretion severity rating scale by Murray et al. [7] was demonstrated to be reliable and valid in the evaluation and graduation of accumulated secretions within the pharyngolarynx and trachea.

The intra-rater reliability of all four raters was high without significant differences. The same was demonstrated for the inter-rater reliability in both rating sessions. Hence, the secretion severity rating scale by Murray et al. is reliable in its use.

No significant differences could be identified between the four ratings within each of the four grades in both sessions, which indicates the homogeneity of the ratings. However, when examining the cross tabulation of the first and second sessions, the raters tended to slightly prefer grades 1 and 3 and to choose grade 0 less frequently in their responses (see Table 2). Nevertheless, the raters presented a high percentage of agreement between their results in the first and second rating sessions (89 %).

The concurrent validity, comparing the median of all ratings with the defined reference standard, demonstrated a high correlation without significant difference. Hence, the secretion severity rating scale by Murray et al. [7] can also be considered valid in use.

The recently published validation of the German version of the penetration–aspiration scale developed by Rosenbek et al. [14] revealed significant differences in the accuracy of ratings between more and less experienced raters, which suggests the possibility that validity of that scale could be improved with training [15]. In this study, however, when assessing the reliability and validity of the secretion severity rating scale by Murray et al. [7], the ratings demonstrated a very homogenous distribution and no significant influence of rater experience. In addition, significant differences in the intra-rater reliability as a possible indicator for a learning effect from the first to second rating sessions were not identified. The only statistically significant factor that influenced the distribution in the ordinal regression was the reference standard defined by two dysphagia experts. Again, the scale scores are valid in use.

While the results regarding reliability and validity are impressive, the sample size in this study presented here is small, which limits the interpretation of the data, and a larger sample size could confirm or refute the findings presented here.

The scale is designed to predict the risk of aspiration of food and liquid during a FEES® examination by estimating the localization of accumulated secretions within the pharyngolarynx and trachea. Other attempts of scaling secretions have included an estimation of secretion volume in the five-point Marianjoy secretion scale published by Donzelli et al. [16]. The scale the authors offered did not present a grading for "normal" or the inclusion of the dynamics of secretion flow, and there were no operational guidelines for the determination of volume. Donzelli et al. [17] reduced the five-point scale to a three-point scale at the expense of the dimension of volume (amount grading) without losing the power of the reduced scale to predict aspiration risk of food and liquid.

The four-point secretion severity rating scale by Murray et al. [7] used as a classification system for the graduation of accumulated secretions turned out to be reliable, valid and, due to the clear conception, unambiguous, and simple in use. Its systematic application offers a standardized and uniform documentation of accumulated secretions for clinical and scientific routine. Therefore, the presented validated version of the four-point secretion severity rating scale by Murray et al. [7] is highly recommended for the consistent implementation in FEES® examination.

Compliance with ethical standards

References

1. Fourrier F, Duvivier B, Boutigny H, Roussel-Delvallez M, Chopin C (1998) Colonization of dental plaque: a source of nosocomial infections in intensive care unit patients. Crit Care Med 26(2):301–308

2. Scannapieco FA, Stewart EM, Mylotte JM (1992) Colonization of dental plaque by respiratory pathogens in medical intensive care patients. Crit Care Med 20(6):740–745

3. Shi Z, Xie H, Wang P, Zhang Q, Wu Y, Chen E, Ng L, Worthington HV, Needleman I, Furness S (2013) Oral hygiene care for critically ill patients to prevent ventilator-associated pneumonia. Cochrane Database Syst Rev 13(8):CD008367

4. Sakamoto T, Horiuchi A, Nakayama Y (2013) Transnasal endoscopic evaluation of swallowing: a bedside technique to evaluate ability to swallow pureed diets in elderly patients with dysphagia. Can J Gastroenterol 27(8):459–462

5. Easterling CC, Antinoja J, Cashin S, Barkhaus PE (2013) Changes in tongue pressure, pulmonary function, and salivary flow in patients with amyotrophic lateral sclerosis. Dysphagia 28(2):217–225

6. Robbins J, Levine R, Wood J, Roecker EB, Luschei E (1995) Age effects on lingual pressure generation as a risk factor for dysphagia. J Gerontol A Biol Sci Med Sci 50A(5):M257–M262

7. Murray J, Langmore SE, Ginsberg S, Dostie A (1996) The significance of accumulated oropharyngeal secretions and swallowing frequency in predicting aspiration. Dysphagia 11(2):99–103

8. Ota K, Saitoh E, Baba M, Sonoda S (2011) The secretion severity rating scale: a potentially useful tool for management of acute-phase fasting stroke patients. J Stroke Cerebrovasc Dis 20(3):183–187

9. Rodrigues B, Nóbrega AC, Sampaio M, Argolo N, Melo A (2011) Silent saliva aspiration in Parkinson's disease. Mov Disord 26(1):138–141

10. Kang Y, Chun MH, Lee SJ (2013) Evaluation of salivary aspiration in brain injured patients with tracheostomy. Ann Rehab Med 37(1):96–102

11. Warnecke T, Suntrup S, Teismann IK, Hamacher C, Oelenberg S, Dziewas R (2013) Standardized endoscopic swallowing evaluation for tracheostomy decannulation in critically ill neurologic patients. Crit Care Med 41(7):1728–1732

12. Takahashi N, Kikutani T, Tamura F, Groher M, Kuboki T (2012) Videoendoscopic assessment of swallowing function to predict the future incidence of pneumonia of the elderly. J Oral Rehabil 39(6):429–437

13. Langmore S (2001) Endoscopic evaluation and treatment of swallowing disorders. Thieme Verlag, New York

14. Rosenbek JC, Robbins JA, Roecker EB, Coyle JC, Wood JL (1996) A penetration-aspiration scale. Dysphagia 11(2):93–98

15. Hey C, Pluschinski P, Zaretsky Y, Almahameed A, Hirth D, Vaerst B, Wagenblast J, Stöver T (2014) Penetrations-Aspirations-Skala nach Rosenbek. Validierung der deutschen Version für die endoskopische Dysphagiediagnostik. HNO 62(4):276–281

16. Donzelli J, Brady S, Wesling M, Craney M (2003) Predictive value of accumulated oropharyngeal secretions for aspiration during video nasal endoscopic evaluation of the swallow. Ann Otol Rhinol Laryngol 112(5):469–475

17. Donzelli J, Brady S, Wesling M, Theisen M (2006) Secretions, occlusion status, and swallowing in patients with a tracheotomy tube: a descriptive study. Ear Nose Throat J 85(12):831–834