Eur Arch Otorhinolaryngol (2007) 264:1059–1064
DOI 10.1007/s00405-007-0299-8

HEAD AND NECK

# The accuracy of the modified Evans blue dye test in detecting aspiration in head and neck cancer patients

**U. Winklmaier · K. Wüst · P. K. Plinkert · F. Wallner**

Received: 29 November 2006 / Accepted: 12 March 2007 / Published online: 13 April 2007


**Abstract** The purpose of this study was to determine sensitivity and specificity of the modified Evans blue dye test (MEBDT) in tracheostomised patients after treatment of head and neck squamous cell carcinoma (HNSCC). This was a prospective study with 30 consecutive patients. All individuals underwent a MEBDT and a subsequent fiberoptic endoscopic evaluation of swallowing (FEES) immediately after the MEBDT for reconsidering the validity of the MEBDT. Aspiration was present in 20 patients documented by MEBDT and FEES. One patient was judged to aspirate by FEES but not by MEBDT (1 false-negative result). Nine patients showed no aspiration either by MEBDT or by FEES. The sensitivity of the MEBDT protocol in predicting aspiration among individuals in our cohort was 95.24%, the specificity 100%, respectively. The results of the current investigation suggest that the MEBDT is much more sensitive in tracheostomised HNSCC patients than in tracheostomised neurological patients. The MEBDT for tracheostomised HNSCC patients offers a quick and reliable method to identify aspiration risk in cases of severe dysphagia.

**Keywords** Aspiration · Dysphagia · Clinical swallowing examination · Tracheostomy · Head and neck cancer

U. Winklmaier (✉) · P. K. Plinkert · F. Wallner
Department of Otorhinolaryngology, Head and Neck Surgery, University of Heidelberg, Im Neuenheimer Feld 400, 69120 Heidelberg, Germany
e-mail: ursula.winklmaier@med.uni-heidelberg.de; ursula.winklmaier@gmx.de

K. Wüst
University of Applied Sciences Pforzheim, Pforzheim, Germany

## Introduction

Aspiration is most common in persons with neurological diseases and in those who have been treated for head and neck cancer. Tracheostomy may be performed in patients who show chronic airway obstruction, patients requiring prolonged ventilatory support but also in patients with chronic aspiration and recurrent aspiration pneumonia. The incidence of aspiration in patients with tracheostomy is 50–87% [1, 2]. Physiological factors may contribute to the development of dysphagia in tracheostomised patients, including reduced laryngeal elevation, reduced pharyngeal sensation, reduced cough response, and atrophy of the laryngeal musculature [3–5]. A tracheostomy tube with cuff usually is placed to prevent aspiration of secretions, aspiration of liquids and food, and aspiration of gastric contents. Because of longitudinal channels caused by folds in the cuff wall material an inflated cuff is not entirely sufficient in preventing aspirated material from entering the lower airway [6]. Differing viscosity in water versus artificially produced saliva accounts for the fact that the leakage of saliva was significantly reduced compared to water [6].

Evans blue dye is a diazo dye that has been used for determining blood volume in humans and animals. The dye was named after Herbert McLean Evans, an American anatomist and physiologist at the University of California, who published blood volume studies in 1920 [7]. The Evans blue dye test for aspiration in tracheostomised individuals was introduced by Cameron [8] in 1973. In the original Evans blue dye test (EBDT), drops of Evans blue dye are placed on the patient's tongue every 4 h and the trachea is suctioned at set intervals. The observation period is 48 h; the finding of suctioned coloured secretion suggests aspiration.

The modified Evans blue dye test (MEBDT) is based on the original EBDT but introduces a slight variation on the original examination as described by Cameron et al. [8]. In the MEBDT, the patient is given blue dyed food and liquids. The test is considered to be positive for aspiration when blue coloured material is suctioned through the tracheostomy. The MEBDT has become a standard clinical tool in the evaluation of patients with tracheostomy and suspected dysphagia because it offers the advantages of economy, simplicity and availability [9].

Blue dye test accuracy has been questioned since the 1980s [10]. In 1995, Thompson-Henry et al. [11] questioned the reliability of the MEBDT in five cases. In that retrospective study, the patients underwent both the MEBDT and a videofluoroscopic swallowing study (VFSS). They reported that in all five cases the MEBDT failed to detect aspiration in tracheostomised patients. Critical evaluation of this study included non-specific subject selection, the fact that the interval time between MEBDT and the VFSS was 4–22 days, and the lack of documentation of the severity of aspiration observed during the VFSS [12].

Brady et al. [9] passed through a prospective study to determine if modified barium swallow (MBS) aspiration corresponded with simultaneous MEBDT aspiration across 20 patients with neurological etiology and three per os consistencies reported that the MEBDT had a 50% false-negative error rate.

Donzelli et al. [13] also reported a 50% false-negative error rate for the detection of trace aspiration amounts. They underwent the investigation with 14 neurological patients. Their results suggest that at best the MEBDT should be viewed as a screening tool for the presence of gross amounts of aspiration in patients with a tracheostomy.

Belafsky et al. [14] evaluated a prospective study with 30 patients of an acute care hospital, etiology unknown. All individuals underwent a MEBDT and a subsequent fiberoptic endoscopic evaluation of swallowing (FEES) using a standard protocol. The MEBDT was carried out three times (30 and 60 min after the first testing). The sensitivity of the MEBDT was 82%, the specificity 38%. In patients with mechanical ventilation sensitivity was even higher (100%). The results support the use of the MEBDT as a screening tool for persons with a tracheostomy tube. The high temporal expenditure required by the investigation should be mentioned as a disadvantage.

Peruzzi et al. [15] assessed the reliability of a bedside coloured dye test in 20 tracheostomised patients with different etiology (11 spinal cord injury, 1 myocardial infarction, 1 respiratory failure, 1 cerebrovascular accident, 2 chronic obstructive pulmonary disease, 1 meningitis, 1 encephalitis, 1 closed head injury, 1 laryngeal cancer). Their data indicate that the coloured dye test for aspiration carries a low sensitivity of 38%, but a high specificity of 100%.

In summary, clinicians administer blue dye tests to detect aspiration without strong evidence of their accuracy in patients with neurological etiology. The primary purpose of the present study was to evaluate the accuracy of our technique of performing the MEBDT in predicting the presence of aspiration in head and neck cancer patients with a tracheostomy tube.

The global null hypothesis was that aspiration results of MEBDT and fiberoptic endoscopic examination of swallowing (FEES) would be unrelated. The experimental hypothesis was that aspiration results of the MEBDT and fiberoptic endoscopic examination of swallowing (FEES) would be related. That means that identification of aspiration during each MEBDT would correspond to identification of aspiration during FEES and that absence of aspiration during MEBDT would correspond to absence of aspiration during FEES.

## Patients and methods

### Subjects

Thirty individuals with HNSCC history participated in this study. They were patients in the Department of Otorhinolaryngology, Head and Neck Surgery, University of Heidelberg, and had known or suspected dysphagia. The data was collected from April 2005 to March 2006. Informed consent was obtained from each subject (or legal guardian) consistent with the ethics committee, University of Heidelberg, concerning research on human subjects. The age range for subjects was 43–78 years and the mean age was 59.6 years. Eight subjects were female and 22 were male.

Sixteen patients had a diagnosis of oropharyngeal carcinoma, 11 had a diagnosis of hypopharyngeal carcinoma, 2 had a diagnosis of oro-/hypopharyngeal carcinoma and 1 had a diagnosis of subglottical chondrosarcoma. All patients were cognitively able to perform the MEBDT and the FEES procedures and to take food orally. Fourteen patients had tracheostomy tubes with cuff, 14 had tracheostomy tubes without cuff. The types of tracheostomy tubes—including fenestration status (fenestrated/non fenestrated), cuff status (with/without cuff), and occlusion status (finger occlusion/speaking valve/without occlusion) during the procedure—were documented. Data related to the patients' demographics, diagnosis, TNM-classification, radiotherapy yes/no, status of tracheostomy tube, BMI and nourishment oral/feeding tube also were documented.

### Procedures

All MEBDT and FEES studies were conducted by a speech and language therapist (SLT) and an otorhinolaryngologist,

each with over 7 years of clinical and instrumental swallowing diagnostic experience. Tracheostomy tubes with cuff were deflated during the examination. All tracheostomy tubes were closed with a speaking valve or by finger occlusion to facilitate subglottic pressure during swallowing. The status of occlusion was documented, even if occlusion was not possible, because of laryngeal edema.

We did six swallowing tests with two different consistencies and two different amounts of test material. Independent of dysphagia severity we carried out two attempts. The clinical examination was determined when the patient aspirated during the first two swallowing trials, i.e. coughing/suctioning coloured secretions out of the tracheostomy tube. First of all, 2 × 5 ml coloured, artificial saliva was given. Draught three consisted of 1 × 15 ml coloured, artificial saliva. The swallowing-trials 4–6 were performed with coloured water: 2 × 5 ml (swallow 4 and 5), and 1 × 15 ml (swallow 6) coloured water. After every draught saliva or water the patient was requested to cough, besides we inserted the endoscope into the tracheostomy tube and observed the bronchial tree for coloured secretions or coloured mucosa. We suctioned at least three times: after the swallowing trials with saliva, after the swallowing examination with water, and 5 min after the examination was finished. If the patient's cough was insufficient or if the patient demonstrated a "wet voice" we suctioned more than three times but documented the exact numbers of suctioning. The specimen was examined for blue discolouration under full room lighting. The presence of blue dye in any tracheal secretions signified aspiration, i.e. a positive MEBDT.

After completion, we examined the subglottic structures up to the trachea and including the bronchial tree, to see if there were any coloured secretions or coloured mucosa. Then the fiberoptic endoscopic examination of swallowing (FEES) was administered. Our technique of FEES has been described elsewhere and is not reiterated [16]. We used the penetration–aspiration scale (P–A), an 8-point ordinal scale, which describes increasing severities of penetration (points 2–5) and aspiration events (points 6–8), no airway entry (point 1) [17]. The endoscopic presence of any food material below the level of the true vocal folds signified aspiration and a positive finding on FEES. Each FEES was performed by a speech language therapist (SLT) in conjunction with an otolaryngologist immediately after the MEBDT.

The prevalence of aspiration in the MEBDT was compared with the prevalence of aspiration on FEES. Using the FEES as the gold standard, the sensitivity and specificity of the MEBDT were calculated.

With the FEES as the objective measure of aspiration, the four possible MEBDT–FEES aspiration agreement outcomes for both types of analysis were as follows: (1) FEES and MEBDT positive for aspiration, (2) FEES and MEBDT negative for aspiration, (3) FEES positive for aspiration and MEBDT negative for aspiration, (4) FEES negative for aspiration and MEBDT positive for aspiration. Correlation between time of aspiration and tumour size (T1–T4), per os consistencies and aspiration severity were investigated descriptively.

### Data analysis

All statistics and figures were computed with the statistical software SPSS (SPSS Inc., Release 14.0.1, Chicago, IL, USA). We report two-tailed statistics throughout. The accepted type I error rate was chosen to be $\alpha = 0.05$, a $P$ value smaller than or equal to 0.05 ($P \leq 0.05$) was considered significant.

The main focus of the study was to calculate sensitivity and specificity of the MEBDT. Furthermore, a confirmatory analysis of the correlation between the time of aspiration and the size of the tumour was carried out. All further questions were analysed in an explorative manner.

The correlation between the time of aspiration and the size of the tumour will first be illustrated descriptively by a contingency table. Because for the test of the null hypothesis "time of aspiration and tumour size are independent" the number of patients per tumour size and therefore the expected frequencies in the corresponding contingency table were small, we built classes of tumour sizes. The tumour sizes T1 and T2 as well as T3 and T4 were pooled. The test was only carried out for patients that had aspired. The null hypothesis was tested with Fisher's exact test.

The body mass index (BMI) is reported with the mean ± standard deviation (minimum, median and maximum). Because the null hypothesis "the BMI of the included patients is normally distributed" did not have to be rejected ($P = 0.401$), the BMI could subsequently be regarded as normally distributed. For the null hypotheses "the BMI of patients with oral and non-oral nutrition does not differ" a parametric $t$ test was carried out. Also, we tested with the $t$ test for dependent samples, whether there was a significant change in the BMI during our study period. The BMI was measured at the time of clinical examination and 6 months later.

## Results

The primary purpose of this study was to determine the accuracy of the MEBDT across 30 tracheostomised head and neck cancer patients.

Aspiration was present during 21 of the 30 FEES studies (70%); aspiration was absent during the other 9 FEES studies (30%). Aspiration was present during 20 of the 30

MEBDT studies (66.7%); aspiration was absent during the other 10 MEBDT studies (33.3%).

These figures indicate that 70% of our cohort aspirated, MEBDT and FEES agreed in 20 patients (67.7%), in one patient (3.3%) the MEBDT and the FEES result disagreed, i.e. the MEBDT showed a false-negative result (Fig. 1).

During 18 of the 20 procedures FEES and MEBDT results revealed aspiration the penetration–aspiration scale score [17] was 6 the other two subjects received a P–A scale score of 7. The one subject, which was false negative in the MEBDT, revealed a P–A scale score of 8. The nine subjects who did not aspirate in the FEES received a P–A scale scores of 1–4 (Fig. 2).

With the FEES as an objective test of presence/absence of aspiration, MEBDT sensitivity and specificity identifying presence/absence of aspiration were 95.24 and 100%, respectively. The MEBDT correctly identified approximately 95.24% of tracheostomised patients who showed signs of aspiration.

Inspecting the time of aspiration in 21 study cases, we detected that 13 patients aspirated artificially produced saliva and eight patients aspirated water during swallowing. Of the 13 patients who aspirated saliva, 9 patients already aspirated when given 5 ml of saliva, 4 patients first aspirated at an administration of 15 ml. Eight patients aspirated after they had subsequently been given 5 ml of water.

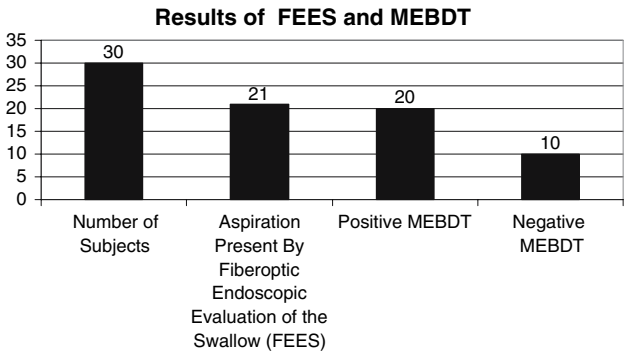


**Fig. 1** Results of FEES and MEBDT

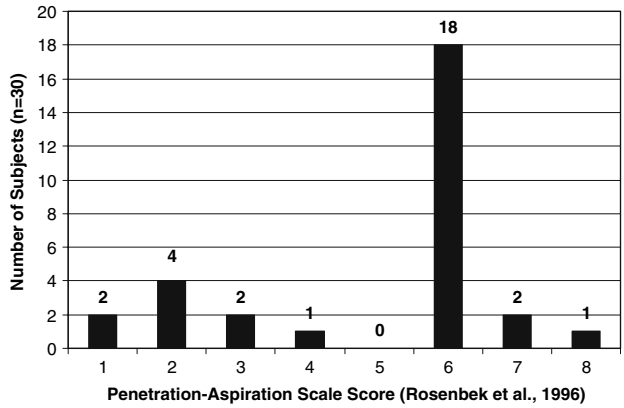


**Fig. 2** Penetration–aspiration scale score, [17]

We examined if there was a correlation between time of aspiration and tumour size. Fisher's exact test shows a non-significant result ($P = 0.656$), the independence between time of aspiration and tumour size cannot be rejected.

At the beginning of the study the BMI of the patients was 22.1 ± 4.9 (14.0; 21.0; 40.0) [mean ± standard deviation (minimum; median; maximum)], at the backup examination 6 months later, we computed a BMI of 22.91 ± 4.79 (17.0; 22.0; 39.0).

Twenty-seven patients had feeding tubes (4 nasogastric tubes, 23 percutaneous gastrostomy tubes, PEG) and were nourished completely or principally via feeding tube, three patients were exclusively nourished orally. There is no significant difference between nutrition and BMI ($P = 0.861$). Furthermore, no significant difference in BMI from the beginning of the study to the end of treatment could be stated ($P = 0.815$). However, 6 months after the first investigation, 10 patients were exclusively nourished orally, 12 patients were completely or principally nourished via PEG, 8 patients died in consequence of the tumour disease. In comparison to the first investigation, a follow-up study 6 months later revealed that 12 patients had been decannulated, 10 patients still had tracheostomy tubes, 8 patients had died in consequence of the tumour disease.

## Discussion

Each individual in our investigation received six swallowing trials when possible, but at least two draughts when aspirating during the first two swallows. We gave two different consistencies and two different amounts of test material. When inspecting the time of aspiration in 21 study cases, we detected that 13 patients (62%) aspirated artificially produced saliva and 8 patients (38%) aspirated water during swallowing. Of the 13 patients who aspirated saliva, 9 (43%) patients already aspirated when given 5 ml of saliva, 4 patients (19%) first aspirated at an administration of 15 ml. Eight patients (38%) aspirated when given 5 ml of water subsequently. We used artificially produced saliva and water because we decided to evaluate whether patients were swallowing saliva much better than water. But the results showed that in severe swallowing disorders even saliva was aspirated. We decided to use this protocol with two different consistencies (saliva and water) and two different amounts of material (5 and 15 ml) because we were interested to evaluate the timing of aspiration in dependence of consistence and amount of the test material. We did not use food (semisolid pudding or pureed solids) as did Brady et al. [9] or Thompson-Henry et al. [11]. We wanted to use material which consistency and viscosity is

the same everywhere. The type and quantity of food that was administered were not specified in the examination of Donzelli et al. [13]. We think that for changing dietary measures FEES and VFSS should be used, but this screening tool determines important information whether the patient is aspirating and how the management of secretions is performed.

Feeding methods with nasogastric tube (NG) or percutaneous gastrostomy tube (PEG) were found to be equally effective at maintaining body weight and BMI [18]. There was no significant difference between NG versus PEG feeding and body weight and BMI ($P = 0.861$) in our cohort.

The primary purpose of the study was to determine the accuracy of the MEBDT across 30 tracheostomised head and neck cancer patients. With FEES as an objective test of the presence or absence of aspiration, the MEBDT sensitivity and specificity for identifying the presence or absence of aspiration were 95.24 and 100%, respectively. That is, the MEBDT did not correctly identify approximately 5% of the tracheostomised patients with swallowing disorders in consequence of tumor treatment (operation/radiation). In contrast, 100% of the tracheostomised patients showing no signs of aspiration were identified correctly by the MEBDT.

Donzelli et al. [13] performed simultaneous FEES and MEBDT on 15 individuals with neurological diseases and tracheostomy tubes. The authors reported a 50% false-negative error rate for the MEBDT. These results agree with the findings of Brady et al. [9], who described a 50% false-negative error rate for simultaneous MEBDT and videofluoroscopic swallowing studies in patients with neurological history. Both of these investigations revealed increased sensitivity of the MEBDT in tracheostomised patients who aspirated more than trace amounts, suggesting that the quantity of the aspirated food bolus is associated with the accuracy of the test. In comparison to the 50% false-negative error rate reported by Donzelli et al. [13] and Brady et al. [9], the false-negative error rate of our investigation was only 5% (1 of 21). We chose not to differentiate between the aspiration of trace and gross amounts of material. Given to the high acuity of our population, we considered even trace amounts of aspiration to be significant. The overall sensitivity of the MEBDT in our cohort of 95.24% is very acceptable and supports the role of this examination as a screening tool in tracheostomised head and neck cancer patients. The high specificity of our MEBDT results of 100% conflicted with the poor specificity of 38%, which Belafsky et al. [14] presented. However, Peruzzi et al. [15] also showed a high specificity of 100%. The poor specificity in the examination by Belafsky et al. [14] may be a more direct consequence of the impact of gravity on normal secretion flow and of keeping the trachea moist when small amounts of blue-tinged secretions are suctioned from the patients tracheostomy tube for the first time several hours after stimulus presentation [3].

Several factors may account for the discrepancy between the sensitivity results of our study and those of Donzelli et al. [13] and Brady et al. [9]. The explanation, which seems most likely to account for the discrepancy exists in the different etiology of the patients and in the homogeneous sample population. All subjects examined in our study suffered from head and neck cancer treatment, the sample population can be qualified to be homogeneous. Brady et al. [9] assessed patients with swallowing disorders due to a diagnosis of neuromuscular weakness, cerebrovascular accident, traumatic brain injury or spinal cord injury. The patients of Donzelli et al. [13] had dysphagia because of cerebrovascular accidents, neuromuscular weakness and status after brain tumour surgery. The sample population of Brady et al. [9] and Donzelli et al. [13] all had dysphagia because of neurological etiology, but because of variable, heterogeneous neurological diseases. It seems that sensitivity is much better preserved in our homogeneous cohort than expected. The FEES swallowing examination showed only one patient aspirating without effort to eject the aspirated material (P–A scale score 8). The other 20 patients aspirating as shown by FEES and MEBDT tried to eject the aspirated material out of the trachea, 18 patients ejected the aspirated material into the larynx or out of the airway (P–A scale score 6), and two patients did not eject the aspirated material from the trachea despite effort (P–A scale score 7). But all 20 patients with P–A scale score 6 or 7 recognized the aspiration and reacted with powerful coughing. In comparison to the study of Eisbruch et al. [19] where 65% of the patients treated with chemoradiation showed "silent" aspiration. In our cohort 24 patients had already undergone radiotherapy in combination with chemotherapy. Six patients suffered from postoperative swallowing disorders without radiotherapy/chemotherapy. Only one individual showed "silent" aspiration and a false-negative result in MEBDT, but this person did not have chemoradiatio, but postoperative swallowing difficulties. While laryngeal closure provides good protection against aspiration, cough is another effective barrier to protect the lower airway from contamination [20].

In our study, we revealed that administering the examination on only one occasion, as was the case in the studies of Donzelli et al. [13] and Brady et al. [9], may not in principle limit the sensitivity as suggested by Belafsky et al. [14], who conducted three separate trials.

In summary, being able to determine if a tracheostomised patient aspirates is of great clinical relevance. Clinicians commonly administer a MEBDT at the patient's bedside to investigate aspiration risk. The sensitivity of our MEBDT protocol in predicting aspiration among tracheostomised head and neck cancer patients in our cohort was 95.24%. These results support the use of the MEBDT as a screening

tool detecting aspiration risk in HNSCC patients. The MEBDT offers the advantages of economy, simplicity and availability. Because the sample population was small, further studies might provide more robust results and support the validity of this clinical dysphagia evaluation.

**Acknowledgments** U. Winklmaier is a Scholar of the Hanns-Seidel-Stiftung, Germany. We thank Dr. S. Schiller and J. Kentner for their contribution to our good work.

## References

1. Elpern E, Scott M, Ries P (1994) Pulmonary aspiration in mechanically ventilated patients with tracheostomies. Chest 105:563–566
2. Pannunzio T (1996) Aspiration of oral feedings in patients with tracheotomies. AACN Clin Issues 7:560–569
3. Logemann J (1998) Evaluation and treatment of swallowing disorders. PRO-ED, Austin
4. Dikeman K, Kazandjian M (1997) Communication and swallowing management of trachomized and ventilator-dependent adults. Singular, San Diego
5. Tolep K, Getch C, Criner G (1996) Swallowing dysfunction in patients receiving prolonged mechanical ventilation. Chest 109:167–172
6. Winklmaier U, Wüst K, Wallner F (2005) Evaluation des aspirationsschutzes blockbarer trachealkanülen. HNO 12:1057–1062
7. Dawson A, Evans H, Wipple G (1920) Blood volume studies: behaviour of large series of dyes introduced into circulating blood. Am J Physiol 51:232
8. Cameron J, Reynolds J, Zuidema G (1973) Aspiration in patients with tracheotomies. Surg Gynecol Obstet 136:68–70
9. Brady S, Hildner C, Hutchins B (1999) Simultaneous videofluoroscopic swallow study and modified Evans blue dye procedure: an evaluation of blue dye visualization in cases of known aspiration. Dysphagia 14:146–149
10. Elpern E, Jacobs E, Bone R (1987) Incidence of aspiration in tracheally intubated adults. Heart Lung 16:527–531
11. Thompson-Henry S, Braddock B (1995) The modified Evan's blue dye procedure fails to detect aspiration in the tracheostomized patient: five case reports. Dysphagia 10:172–174
12. Leder S (1996) Comment on Thompson-Henry and Braddock: the modified Evan's blue dye procedure fails to detect aspiration in tracheostomized patient: five case reports. Dysphagia 11:80–81
13. Donzelli J, Brady S, Wesling M, Craney M (2001) Simultaneous modified Evans blue dye procedure and video nasal endoscopic evaluation of the swallow. Laryngoscope 111:1746–1750
14. Belafsky P, Blumenfeld L, LePage A, Nahrstedt K (2003) The accuracy of the modified Evan's blue dye test in predicting aspiration. Laryngoscope 113:1969–1972
15. Peruzzi W, Logeman J, Currie D, Moen S (2001) Assessement of aspiration in patients with tracheostomies: comparison of the bedside colored dye assessment with videofluoroscopic examination. Respir Care 46:243–247
16. Langmore S (2001) Endoscopic evaluation and treatment of swallowing disorders. Thieme, New York
17. Rosenbek J, Robbins J, Roecker E, Coyle J, Wood J (1996) A penetration–aspiration scale. Dysphagia 11:93–98
18. Magné N, Marcy P, Foa C, Falewee M, Schneider M, Demard F, Bensadoun R (2001) Comparison between nasogastric tube feeding and percutaneous fluoroscopic gastrostomy in advanced head and neck cancer patients. Eur Arch Otolaryngol 258:89–92
19. Eisbruch A, Lyden T, Bradford C (2002) Objective assessment of swallowing disfunction and aspiration after radiation concurrent with chemotherapy for head-and-neck cancer. Int J Radiat Oncol Biol Phys 53:23–28
20. Flaksman H, Ron Y, Ben-Davis N, Cinamon U, Levy D, Russo E, Sokolov M, Avni Y, Roth Y (2006) Modified endoscopic swallowing test for improved diagnosis and prevention of aspiration. Eur Arch Otolaryngol 263:637–640