Original Article

A Composite Thyromental-Modified Mallampati-Sternomental Score Predicts Difficult Laryngoscopy Better Than Individual Airway Examinations

Tribuddharat S, MD, PhD[1], Sathitkarnmanee T, MD[1], Sangsungnern P, MD[1]

[1] Department of Anesthesiology, Faculty of Medicine, Khon Kaen University, Khon Kaen, Thailand

Background: Current individual and combined bedside airway tests can predict difficult laryngoscopy with low sensitivity.

Objective: To develop a composite scoring model to predict difficult laryngoscopy with higher sensitivity.

Materials and Methods: The retrospectively recruited samples were randomly divided into two groups, 80% for the Training group, and 20% for the Validation group. In the Training group, the extracted data included 1) the body mass index (BMI), 2) pre-operative bedside airway tests, i.e., inter-incisor gap (IIG), thyromental distance (TMD), sternomental distance (SMD), and modified Mallampati test (MMT), and 3) laryngoscopic view according to Cormack-Lehane (C-L) classification. The authors classified grade 3 and 4, according to the C-L classification, as difficult laryngoscopy. A predictive scoring model was developed using multivariate logistic regression analyses. The model was then validated in the Validation group.

Results: Seven thousand eight hundred five patients with a prevalence of difficult laryngoscopy of 3.6%, were divided as 6,251 and 1,554 patients in the Training and Validation group, respectively. In the Training group, the developed thyromental-modified Mallampati-sternomental (TMS) score comprised of three tests, the TMD, MMT, and SMD. The TMS had a total score of 6 and a score of 2 or more can predict difficult laryngoscopy with a respective sensitivity and specificity of 66.1% (95% CI 59.4 to 72.3) and 92.5% (95% CI 91.8 to 93.2). The performance was comparable when validated in the Validation group.

Conclusion: The TMS score can predict difficult laryngoscopy with a higher AUC and sensitivity albeit slightly decreased specificity than the TMD, SMD, MMT, or IIG.

Keywords: Airway test, Predictive model, Validation study, Difficult laryngoscopy

Received 23 Sep 2019 | Revised 19 Nov 2019 | Accepted 26 Nov 2019

J Med Assoc Thai 2020;103(2):148-54
Website: http://www.jmatonline.com

The median (interquartile range [IQR]) of prevalence for difficult laryngoscopy from a recent systematic review with meta-analysis was 11% (6% to 19%)(1). According to the American Society of Anesthesiologists (ASA) Closed Claims and the Fourth National Audit Project in the United Kingdom, serious adverse outcomes including brain injury, death, emergency surgical airway, or unintended intensive care unit admission resulted from airway management(2). To avoid unanticipated difficult intubation, many bedside airway tests such as the modified Mallampati test (MMT), upper lip bite test (ULBT), measurement of inter-incisor gap (IIG), thyromental distance (TMD), and sternomental distance (SMD), were implemented in the pre-operative evaluation. The current individual and combined bedside tests, however, have inconsistent capacity, sensitivity, specificity, and positive likelihood ratio for discriminating between patients with difficult and easy airway(3). The objective of the current study was to develop and validate a composite scoring model to predict difficult laryngoscopy with a higher sensitivity.

Correspondence to:

Sathitkarnmanee T.

Department of Anesthesiology, Faculty of Medicine, Khon Kaen University, 123, Mittraphap Road, Khon Kaen 40002, Thailand.

Phone: +66-81-9547622

Email: thepakorns@gmail.com

How to cite this article: Tribuddharat S, Sathitkarnmanee T, Sangsungnern P. A Composite Thyromental-Modified Mallampati-Sternomental Score Predicts Difficult Laryngoscopy Better Than Individual Airway Examinations. J Med Assoc Thai 2020;103:148-54.

Materials and Methods

This was a retrospective and analytical study. It was approved by the Khon Kaen University Ethics Committee in Human Research (HE581149) and was registered with the ClinicalTrials.gov (NCT02944305). Since patient identification data were concealed, informed consent was waived. The data extracting sheet did not contain the name and hospital number of the patient, so a unique study number was used instead.

The authors aimed to develop a predictive score for difficult laryngoscopy from five risk factors. The required sample size (n) to avoid overfitting of the logistic regression model, according to Tabachnick and Fidell[4], should ideally be 50+10(k). This meant a total of 110 patients were required based on the five relevant clinical risk factors (k) with a dropout margin of 10%. To increase the power and precision of the study, the authors included all adults that underwent general anesthesia with orotracheal intubation at Srinagarind Hospital between January 1, 2012 and December 31, 2014. The exclusion criterion was patients with incomplete medical records.

The following patient data were extracted from the medical records for analyses, 1) demographic data including the body mass index (BMI), 2) preoperative airway tests (viz., IIG, TMD, SMD, and MMT), and 3) laryngoscopic view according to the Cormack-Lehane (C-L) classification. The authors classified grade 3 and 4 according to the C-L classification as difficult laryngoscopy in accordance with other studies[1,5].

Statistical analysis

To assess model fit, the present study randomly divided the total recruited sample into two groups, 80% for the Training group to develop the predictive model, and 20% for the Validation group to validate the model. To evaluate the predictive value for difficult laryngoscopy in the Training group, each airway test, as well as BMI, was assessed for the area under receiver operating characteristic curve (AUC). The cut-off point of each factor was then identified by finding the point with the maximum AUC. Each factor at the cut-off point was assessed for AUC and the crude odds ratio using a univariate logistic regression analysis. All factors were then included in a multivariate logistic regression analysis to identify relevant factors. The coefficients of each relevant factor, derived from the multivariate analysis, were used to construct a predictive scoring model. The discriminating ability of the model was assessed by evaluating the AUC. The cut-off point of the model was identified as mentioned. The sensitivity, specificity, positive predictive value, and positive likelihood ratio of the model, as well as each airway test, were determined. The developed model was then validated in the Validation group. Statistical analyses were performed using SPSS, version 16.0 (SPSS Inc., Chicago, Ill, USA).

Results

The authors recruited 7,805 patients. The demographic and clinical data are presented in Table 1. Two hundred eighty-one patients (3.6%) had a C-L view grade of 3 or 4, referred to as difficult laryngoscopy.

Development of the predictive score

The Training group comprised of 6,251 patients with a prevalence of difficult laryngoscopy of 3.6%. The demographic and clinical data are presented in Table 1. The cut-off points for each factor, according to AUC analysis, were TMD of 6 cm or less, SMD of 12 cm or less, MMT greater than 2, IIG of less than 4 cm, and BMI of 24 kg per m² or more. Table 2 presents the respective AUC, crude odds ratio, and p-value of each risk factors for predicting difficult laryngoscopy.

Table 3 presents the results of the multivariate analysis. It identified only three factors, which were TMD of 6 cm or less, MMT greater than 2, and SMD of 12 cm or less. The coefficients of each factor from the multivariate analysis were rounded up and used to construct a predictive scoring model, which was the thyromental-modified Mallampati-sternomental (TMS) score with a total score of 6 (Table 3). The AUC of the TMS score for predicting difficult laryngoscopy was 0.826 (95% CI 0.790 to 0.863) (Figure 1). The cut-off point of TMS score was 2 or greater. The AUC, sensitivity, specificity, positive predictive value, and positive likelihood ratio of a TMS score was 2 or greater, as well as an IIG of less than 4 cm, an MMT of more than 2, a TMD of 6 cm or less, and a SMD of 12 cm or less are presented in Table 4. The TMS score had a higher AUC and sensitivity than all of the individual tests.

Validation of the predicting score

The Validation group comprised of 1,554 patients with a prevalence of difficult laryngoscopy of 3.9%. The demographic and clinical data are presented in Table 1. The AUC of the TMS score for predicting difficult laryngoscopy in the Validation group was

Table 1. Demographics and clinical data (n=7,805)

Parameter	Total (n=7,805) n (%)	Training group (n=6,251) n (%)	Validation group (n=1,554) n (%)
Age (year); mean±SD	51.05±16.35	51.09±16.42	51.03±16.32
Male	3,524 (45.15)	2,816 (45.05)	708 (45.56)
Female	4,281 (54.85)	3,435 (54.95)	846 (54.44)
Weight (kg); mean±SD	59.88±12.83	59.84±12.91	60.07±12.52
Height (m); mean±SD	1.60±0.08	1.60 0.08	1.60±0.08
BMI (kg/m²); mean±SD	23.33±4.46	23.32±4.48	23.36±4.38
IIG (cm)			
Mean±SD	4.28±0.58	4.28±0.58	4.29±0.60
Median (min-max)	4 (2 to 6)	4 (2 to 6)	4 (2 to 6)
TMD (cm)			
Mean±SD	7.23±0.63	7.23±0.63	7.25±0.65
Median (min-max)	7 (4 to 12)	7 (4 to 10)	7 (5 to 12)
SMD (cm)			
Mean±SD	14.75±1.42	14.74±1.40	14.81±1.50
Median (min-max)	15 (10 to 21)	15 (10 to 21)	15 (10 to 21)
MMT			
Class 1	4,329 (55.5)	3,451 (55.2)	878 (56.5)
Class 2	3,102 (39.7)	2,501 (40.0)	601 (38.7)
Class 3	342 (4.4)	276 (4.4)	66 (4.2)
Class 4	32 (0.4)	23 (0.4)	9 (0.6)
Median (min-max)	1 (1 to 4)	1 (1 to 4)	1 (1 to 4)
Cormack-Lehane classification			
Grade 1	6,627 (84.9)	5,304 (84.8)	1,323 (85.1)
Grade 2	897 (11.5)	726 (11.6)	171 (11.0)
Grade 3	257 (3.3)	205 (3.3)	52 (3.4)
Grade 4	24 (0.3)	16 (0.3)	8 (0.5)
Median (min-max)	1 (1 to 4)	1 (1 to 4)	1 (1 to 4)

BMI=body mass index; IIG=inter-incisor gap; TMD=thyromental distance; SMD=sternomental distance; MMT=modified Mallampati test; SD=standard deviation

Table 2. Results of univariate analysis in Training group (n=6,251)

Parameter	AUC	Crude odds ratio	95% CI	p-value
TMD ≤6 cm	0.770	33.49	27.02 to 45.90	<0.001
SMD ≤12 cm	0.661	19.74	14.93 to 26.30	<0.001
MMT >2	0.616	9.03	7.05 to 12.49	<0.001
IIG <4 cm	0.558	3.94	3.04 to 5.87	<0.001
BMI ≥24 kg/m²	0.554	1.55	1.22 to 1.97	<0.001

AUC=area under receiver operating characteristic curve; TMD=thyromental distance; SMD=sternomental distance; MMT=modified Mallampati test; IIG=inter-incisor gap; BMI=body mass index; CI=confidence interval

0.818 (95% CI 0.745 to 0.891) (Figure 2). The performance of a TMS score of 2 or more, an IIG of less than 4 cm, an MMT of 2 or more, a TMD of 6 cm or less, and a SMD of 12 cm or less to predict difficult laryngoscopy in the Validation group are presented in Table 5. The TMS score outperformed all other

Table 3. Results of multivariate analysis and TMS score in the Training group (n=6,251)

Parameter	Coefficient	Adjusted odds ratio	95% CI	p-value	TMS score
TMD ≤6 cm	2.92	18.54	14.38 to 27.80	<0.001	3
MMT >2	1.60	4.94	3.32 to 7.01	<0.001	2
SMD ≤12 cm	0.92	2.52	1.50 to 3.24	<0.001	1
Total					6

TMS=thyromental-modified Mallampati-sternomental; TMD=thyromental distance; MMT=modified Mallampati test; SMD=sternomental distance; CI=confidence interval

Table 4. Performance of different tests in the Training group (n=6,251)

Test	AUC (95% CI)	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	Positive predictive value (%) (95% CI)	Positive likelihood ratio (95% CI)
TMS score ≥2	0.793 (0.755 to 0.831)	66.1 (59.4 to 72.3)	92.5 (91.8 to 93.2)	24.5 (22.1 to 26.9)	8.8 (7.8 to 10.1)
TMD ≤6 cm	0.770 (0.729 to 0.810)	57.9 (51.1 to 64.5)	96.1 (95.5 to 96.5)	35.0 (31.3 to 38.9)	14.7 (12.4 to 17.4)
SMD ≤12 cm	0.661 (0.617 to 0.705)	34.8 (28.6 to 41.5)	97.4 (96.9 to 97.8)	32.6 (27.7 to 38.0)	13.2 (10.4 to 16.7)
MMT >2	0.616 (0.573 to 0.659)	27.2 (21.4 to 33.5)	96.0 (95.5 to 96.5)	20.1 (16.4 to 24.4)	6.9 (5.3 to 8.8)
IIG <4 cm	0.558 (0.516 to 0.600)	16.3 (11.7 to 21.8)	95.3 (94.7 to 95.8)	11.3 (8.4 to 14.9)	3.5 (2.5 to 4.8)

AUC=area under receiver operating characteristic curve; IIG=inter-incisor gap; MMT=modified Mallampati test; TMD=thyromental distance; SMD=sternomental distance; TMS=thyromental-modified Mallampati-sternomental; CI=confidence interval

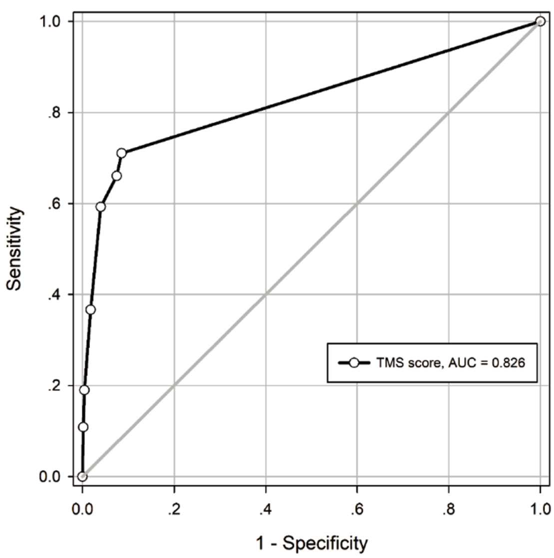

Figure 1. AUC curve of TMS score in the Training group.

The area under the receiver operating characteristic curve is 0.826 (95% CI 0.790 to 0.863)

TMS=thyromental-modified Mallampati-sternomental; AUC=area under receiver operating characteristic curve

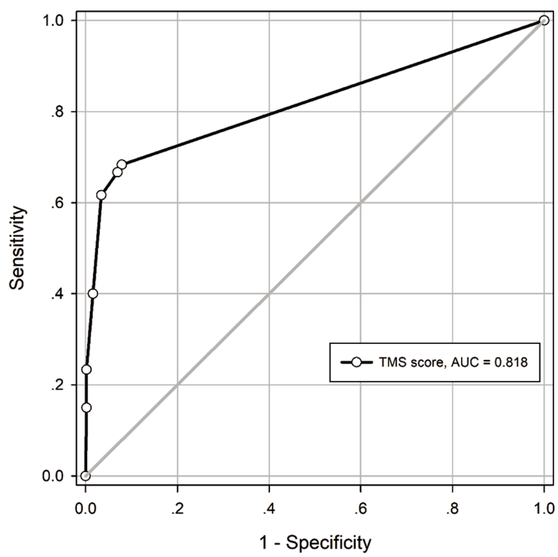

Figure 2. AUC curve of TMS score in the Validation group.

The area under the receiver operating characteristic curve is 0.818 (95% CI 0.745 to 0.891)

TMS=thyromental-modified Mallampati-sternomental; AUC=area under receiver operating characteristic curve

individual tests. The performance of each test was comparable in both the Training and Validation group.

Discussion

The prevalence of difficult laryngoscopy in the present study was 3.6%, which is lower than that reported in a recent systematic review with a median (IQR) of 11% (6% to 19%)[1]. The results of the current study showed that the TMS score, developed from the Training group, could predict difficult laryngoscopy

Table 5. Performance of different tests in the Validation group (n=1,554)

Test	AUC (95% CI)	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	Positive predictive value (%) (95% CI)	Positive likelihood ratio (95% CI)
TMS score ≥2	0.799 (0.727 to 0.871)	66.7 (53.3 to 78.3)	93.0 (91.6 to 94.3)	27.8 (22.9 to 33.2)	9.6 (7.4 to 12.4)
TMD ≤6 cm	0.783 (0.706 to 0.860)	60.0 (46.5 to 72.4)	96.7 (95.6 to 97.5)	41.9 (33.8 to 50.3)	17.9 (12.7 to 25.2)
SMD ≤12 cm	0.662 (0.578 to 0.746)	35.0 (23.1 to 48.4)	97.4 (96.5 to 98.1)	35.0 (25.3 to 46.1)	13.4 (8.4 to 21.3)
MMT >2	0.631 (0.548 to 0.714)	30.0 (18.9 to 43.2)	96.2 (95.1 to 97.1)	24.0 (16.6 to 33.4)	7.9 (5.0 to 12.5)
IIG <4 cm	0.578 (0.497 to 0.659)	20.0 (10.8 to 32.3)	95.6 (94.4 to 96.6)	15.4 (9.4 to 24.1)	4.5 (2.6 to 7.9)

AUC=area under receiver operating characteristic curve; IIG=inter-incisor gap; MMT=modified Mallampati test; TMD=thyromental distance; SMD=sternomental distance; TMS=thyromental-modified Mallampati-sternomental; CI=confidence interval

better than all of the individual tests. When the authors applied the TMS score in the Validation group, the performance was comparable to that of the Training group, indicating that there was no overfitting of the model.

Harms from airway management occur mostly as a consequence of unanticipated and unplanned management of difficult airway. To avoid unexpected difficult intubation, bedside airway tests should have high sensitivity to reduce false negatives. Low specificity of tests increases false positives which may increase the workload of the anesthesiologist, while not affecting patient safety. Most individual airway tests have low sensitivity and high specificity[1,3], so they are not good at predicting difficult laryngoscopy.

In the present study, IIG had a respective sensitivity and specificity of 16.3% (95% CI 11.7 to 21.8) and 95.3% (95% CI 94.7 to 95.8), which is similar to the respective summary sensitivity and specificity of 22% (95% CI 13 to 33) and 94% (95% CI 90 to 97)[1], hence the lowest discriminating ability. The respective sensitivity and specificity of MMT were 27.2% (95% CI 21.4 to 33.5) and 96.0% (95% CI 95.5 to 96.5), which is different from the respective summary sensitivity and specificity of 53% (95% CI 47 to 59) and 80% (95% CI 74 to 85)[1]. The respective sensitivity and specificity of SMD was 34.8% (95% CI 28.6 to 41.5) and 97.4% (95% CI 96.9 to 97.8), which is similar to the 33% (95% CI 16 to 56) and 92% (95% CI 86 to 96)[1] that indicated poor capability to differentiate. Among all individual tests, TMD had the best separating power with a respective sensitivity and specificity of 57.9% (95% CI 51.1 to 64.5) and 96.1% (95% CI 95.5 to 996.5), which is better than the respective summary sensitivity and specificity of 37% (95% CI 28 to 47) and 89% (95% CI 84 to 93)[1].

To increase the sensitivity to predict difficult laryngoscopy, many studies reported using a multifactor predictive score comprising many bedside airway tests with clinical history and symptoms and signs. Frerk studied 144 adults combining MMT with TMD to predict difficult laryngoscopy and found that the new method increased specificity but not sensitivity[6]. Pottecher et al also found that a combination of MMT with IIG in 663 women increased specificity but not sensitivity[7]. Descoins et al proposed a new multifactor screening score be established by allocating points (0, 3, 5, or 7) depending on the degree of presence of seven factors (viz., pathology known to be associated with difficult intubation, clinical signs of airway pathology, IIG and mandible luxation, submental mandibular-thyroid distance, normal or short and broad neck, head and neck movements, and MMT). A score of 11 or more allowed for the prediction of difficult intubations with a sensitivity of 96% and a specificity of 90%[8]. Since this score was constructed from only 295 patients, it tended to be overfitting. el-Ganzouri et al reported a composite airway risk index, in 10,507 patients, comprising of seven risk factors including three airway tests that were IIG, TMD, and MMT. Compared with individual airway tests, this composite index increased specificity but not sensitivity[9]. Arne et al reported a clinical multivariate risk index, in 1,200 patients, comprising the seven risk factors, including IIG, TMD, and MMT. They used a total score of 48 and a cut-off point of 11 or greater for predicting difficult tracheal intubation with a respective sensitivity and specificity of 92% and 93%[10]. The other risk factors in this score comprised results of history taking and other physical examinations, including previous knowledge of difficult intubation with a score of 10 from 48, which was nearly equal to the cut-off point. Generally, if a patient reports a history of previous difficult intubation, it is most likely that he will have difficult intubation in the next anesthesia. The very high sensitivity of this score arises from this reason, and not from the integration of various bedside airway tests per se.

Compared with the aforementioned multifactor predictive scores, the TMS score comprises only of bedside airway tests. The authors weighed the score of each test according to the coefficient from a multivariate logistic regression to reflect the predictive ability of each test. The TMS score can predict difficult laryngoscopy with a high AUC in both the Training and Validation group (0.826 and 0.818). The TMS score outperforms all individual tests with higher AUC and sensitivity at the expense of only a slight decrease in specificity in both the Training and Validation group. The low positive predictive values of these tests are the result of the low prevalence of difficult laryngoscopy. The low positive likelihood ratios arise from the high specificity of these tests.

Limitation

There are some limitations to the present study. Since this was a retrospective study, some of the data may not be accurate. The authors' hospital did not include UBT as a standard bedside airway test, so the UBT was not included. The examiners of bedside airway tests in the current study comprised of several skill level competences among residents, which may affect inter-rater reliability. Finally, the TMS was developed from data of a single university hospital, validation of this score in different context is recommended.

Conclusion

The TMS score comprises three bedside airway tests, the TMD, MMT, and SMD. The total score was 6 with a score of 2 or more being predictive of difficult laryngoscopy with a respective sensitivity and specificity of 66.1% (95% CI 59.4 to 72.3) and 92.5% (95% CI 91.8 to 93.2). The TMS score can predict difficult laryngoscopy better than the TMD, SMD, MMT, and IIG with a higher AUC and sensitivity at the expense of a slight decrease in specificity. Notwithstanding, the sensitivity of all bedside airway tests, both individual and composite, is still not high enough to minimize false negatives. A difficult intubation can occur unexpectedly, so the attending anesthesiologists should be both knowledgeable and skillful and have a multi-level plan. Additional special airway equipment and experienced senior personnel should be at hand.

What is already known on this topic?

Serious adverse outcomes, including brain injury, death, emergency surgical airway, or unintended intensive care unit admission, can result from airway management. Current individual and combined bedside airway tests can predict difficult laryngoscopy with low sensitivity.

What this study adds?

The TMS score can predict difficult laryngoscopy better than the TMD, SMD, MMT, and IIG with a higher AUC and sensitivity at the expense of a slight decrease in specificity.

Acknowledgement

The authors would like to thank Mr. Bryan Roderick Hamman for assistance with the English-language presentation of the manuscript under the aegis of the Khon Kaen University Publication Clinic.

Funding disclosure

The present study was supported by the Faculty of Medicine, Khon Kaen University.

Conflicts of interest

The authors declare no conflict of interest.

Data availability

The data used to support the findings of this study are available from the corresponding author upon request.

References

1. Roth D, Pace NL, Lee A, Hovhannisyan K, Warenits AM, Arrich J, et al. Airway physical examination tests for detection of difficult airway management in apparently normal adult patients. Cochrane Database Syst Rev 2018;5:CD008874.
2. Cooper RM. Preparation for and management of "Failed" laryngoscopy and/or intubation. Anesthesiology 2019;130:833-49.
3. Vannucci A, Cavallone LF. Bedside predictors of difficult intubation: a systematic review. Minerva Anestesiol 2016;82:69-83.
4. Tabachnick BG, Fidell LS. Using multivariate statistics. 5th ed. Upper Saddle River, NJ: Pearson Allyn &Bacon; 2007.
5. Cook TM. A new practical classification of laryngeal view. Anaesthesia 2000;55:274-9.
6. Frerk CM. Predicting difficult intubation. Anaesthesia 1991;46:1005-8.
7. Pottecher T, Velten M, Galani M, Forrler M. Comparative value of clinical signs of difficult tracheal intubation in women. Ann Fr Anesth Reanim 1991;10:430-5.
8. Descoins P, Arné J, Bresard D, Ariès J, Fusciardi J. Proposal for a new multifactor screening score

of difficult intubation in ORL and stomatognathic surgery: preliminary study. Ann Fr Anesth Reanim 1994;13:195-200.
9. el Ganzouri AR, McCarthy RJ, Tuman KJ, Tanck EN, Ivankovich AD. Preoperative airway assessment: predictive value of a multivariate risk index. Anesth Analg 1996;82:1197-204.
10. Arné J, Descoins P, Fusciardi J, Ingrand P, Ferrier B, Boudigues D, et al. Preoperative assessment for difficult intubation in general and ENT surgery: predictive value of a clinical multivariate risk index. Br J Anaesth 1998;80:140-6.