RESEARCH ARTICLE

Comparison of Four Methods for Predicting Difficult Laryngoscopy: A Prospective Study of Validity Indexes

Farhad Etezadi[1], Leila Saeedinia[1], Pejman Pourfakhr[1]*, Atabak Najafi[1], Mohammadreza Khajavi[1], Aylar Ahangari[1], Reza Shariat Moharari[1]

Background: Considering that a simple yet valid airway predictor is basically preferred to a multivariable test with the same level of validity, we designed a study to compare the validity indexes of the TMHT, as a single variable test, with that of Wilson, Arne, and Naguib tests.

Methods: 270 consecutive ASA physical statues I and II patients aged ≥16 years, candidate to receive general anesthesia for elective surgeries, were enrolled in the study. All patients preoperatively were evaluated using four airway predictor tests including Wilson, Arne, Naguib, and Thyromental height test (TMHT). Cut-off points to predict a difficult laryngoscopy were extracted from previous published studies for each test. Based on each predictor, all patients were classified into either easy or difficult. Based on Cormack-lehane scoring system, as the gold standard, all patients during laryngoscopy were categorized into two groups of "easy, grades I, II" and "difficult, grades III, IV". Finally, validity indexes for each test were calculated and compared to one another.

Results: The incidence of difficult laryngoscopy according to Cormack-lehane grading was 17.4% (47/270). The predicted rates of difficult laryngoscopy were 47 (17.4%), 2 (0.7%), 10 (3.7%) and 61 (22.6%) for TMHT, Wilson, Arne and Naguib respectively. Based on Chi-square test, the TMHT as well as the Naguib could significantly predict difficult laryngoscopy. However, the highest rate of sensitivity, positive predictive value, negative predictive value, and accuracy belonged to TMHT.

Conclusion: Both TMHT and Naguib are acceptable predictors of difficult laryngoscopy while the TMHT is a more accurate predictor of difficult laryngoscopy than the other multivariable models.

Keywords: difficult laryngoscopy; airway assessment; airway predictor test; cormack-lehane grading

Maintenance of a secure airway for proper oxygenation and ventilation is the main duty of an anesthesiologist. Failure to establish and keep a secure airway is reported to be the most important cause of mortality in this field [1-3]. The incidence of difficult laryngoscopy in various surveys were reported in the range of 1.2-20% [1-4]. Difficult laryngoscopy is almost always synonymous with difficult intubation in majority of patients [5]. Patients with a high probability of having a difficult airway requires specific protection and preparation to avoid exposing to hazardous consequences. Several investigations explained prediction schemes by applying either a single variable predictor or a multivariable test [6-15]. Many of multivariable models are focused on the best combination of clinical and anatomical dimensions including Wilson, Arne, and Naguib models. On the other hand, a newly introduced single variable predicting model, which is named thyromental height test (TMHT), showed promising result [10]. Considering that a univariate test with a higher or even the same degree of validity in comparison to its multivariable counterpart is rationally more applicable in real practice, we designed a study to compare the validity indexes of TMHT with that of Wilson, Arne, and Naguib models.

Methods

This prospective observational study was approved by Ethics committee of Tehran University of Medical Sciences. Ethics committee ordered that it should not be registered as a clinical trial, since the study was an observational one and no new intervention would apply on patients. Oral patient consent was obtained from all patients in the preoperative clinic because our evaluations did not put patients at a further risk or cost than usual. We studied 270 consecutive ASA physical statues I and II patients aged ≥16 years who were eligible to receive general anesthesia. All patients required endotracheal intubation for elective surgeries including orthopedic (trauma, spine), urologic, gastro-intestinal (including bariatric surgery), and vascular (central, peripheral). Patients with compromised airway, maxillofacial malformations, age less than 15-year-old, and those who needed to undertake emergency surgery and candidates for awake intubation were excluded. Preoperative airway assessments were performed for all patients in preoperative clinic by the same anesthesiologist who was blind to laryngoscopy evaluations in operating rooms.

In preoperative clinic, in addition to routine history taking and physical exams, all patients were evaluated using four

From the [1]Department of Anesthesiology and critical Care, Sina Hospital, Tehran University of Medical Sciences, Tehran, Iran.
Received: 1 March 2018, Revised: 23 March 2018, Accepted: 4 April 2018
The authors declare no conflicts of interest.
*Corresponding author: Pejman Pourfakhr, MD. Department of Anesthesiology and critical Care, Sina Hospital, Tehran University of Medical Sciences, Tehran, Iran. E-mail: pourfakhr@tums.ac.ir

predictive tests including Wilson, Arne, Naguib, and TMHT. Wilson score consists of weight, head and neck movement, receding mandible, jaw movement and presence of buck teeth in a way that each parameter gets the score of 0-2. Total score is derived by adding up the numbers and when it is 4 or more, the intubation is considered as difficult [4]. Arne score consists of previous knowledge of difficult intubation, any disease associated with difficult intubation, clinical symptom of airway pathology, incisor gap, mandible subluxation, Thyromental distance, maximum range of head and neck movement, and Mallampati score (classes 1-4). Maximum Arne score is 48 and more than 11 was considered to be a predictive of difficult intubation [8]. The last multifactorial score named Naguib, consists of Mallampati score (classes 1-4), Thyrosternal distance, Thyromental distance, and neck circumference which is calculated with the given formula: {4.9504 + (thyrosternal distance * 1.1003) - (Mallampati score * 2.6076) + (thyromental distance * 0.9684) - (neck circumference * 0.3966)}. A Naguib score less than zero was regarded as predictive of difficult intubation [9]. TMHT is the vertical distance between the upper border of the mentum and the thyroid (Adam's apple) with the patient lying on his back on a horizontal flat surface while the mouth is closed. The height was measured using a depth gauge (INSIZE® Electronic Depth Gage, INSIZE Co LTD., Suzhou New District, China). A height less than 50 mm is regarded as predictive of difficult intubation [10].

For all patients, standard monitoring was established before commencement of anesthesia. Induction of general anesthesia was started using injection of 0.05-0.15 mg/kg midazolam and 2-3mcg/kg fentanyl as premedication drugs and then complete induction performed with injecting 4-5mg/kg sodium thiopental and 0.5mg/kg atracurium. The patient's head was placed in a sniffing position and laryngoscopy was performed 3 minutes after atracurium injection, which provides maximum muscle relaxation, using a Macintosh blade laryngoscope (no=4). In all cases, a fourth-year resident of anesthesia with more than three years of continuous experience of intubation, blinded to preoperative clinic evaluations, performed the intubations. Glottis visualization was assessed using a modified Cormack-Lehane (C-L) classification without external laryngeal manipulation. We omitted backward, upward, rightward pressure (BURP) maneuver because it may artificially increase thyromental height in patients while the size of increment may not be the same for all patients. It is crystal clear that the flexibility of tissues of neck and larynx in an individual as well as the force imposed by anesthesia personnel may differ among cases and so it might be a potential confounder in our study. If more than one intubation attempt occurs, the best view that led to intubation was assigned for the C-L grading. This classification involves four grades of glottis visualization: Grade 1; complete of the vocal cords; Grade II; the inferior portion of the glottis; Grade III; only the epiglottis; and Grade IV; a non-visualized epiglottis [11]. In this study, C/L grades III and IV were set as gold standard for classifying a case into difficult laryngoscopy group.

Statistical Analysis

All data were analyzed by SPSS software 17 edition. Chi-square test (Fisher exact test) was used for statistical comparisons. Raw data are presented as number/ percent or mean +/- standard deviation. The results are given based on descriptive analysis. Based on previous studies, the average incidence of difficult laryngoscopy estimated about 10 per cent, so a sample size of 270 was calculated to have at least 90% power to detect the agreement between the C/L grade and the predictors. Validity indexes including sensitivity, specificity, positive predictive value, negative predictive value, and accuracy in order to compare four predictors are calculated. Sex, Age, Weight, and Height of all patients are also recorded and presented. A P value less than 0.05 considered as significant.

Results

Two hundred and seventy ASA physical status I and II patients, consiststing of 137 men (50.7%) and 133 women (49.3%), participated in this study. Descriptive information of the study population is reported (Table1). According to the gold standard test (C-L) of the study, difficult laryngoscopy is reported in forty seven patients (17.4%) with C-L grade III or IV of laryngeal view. Fifty three patients needed two attempts and four patients needed three attempts to have a successful intubation. There was no report of failed intubation.

The rate of difficult intubation were reported 47 (17.4%), 2 (0.7%), 10 (3.7%) and 61 (22.6%) by TMHT, Wilson, Arne and Naguib respectively. In addition, Mallampati test which is an accepted and popular predictor, showed the presence of difficult intubation in 23 patients (8.5%). Statistical comparison of the four predictors (Table 2), revealed that both TMHT and Naguib tests are significantly reliable to predict difficult laryngoscopy ($p<0.001$). The statistical parameters obtained via the study including sensitivity, specificity, accuracy, positive predictive value, negative predictive value, and accuracy are shown (Table 3).

Table 1- Quantitative and qualitative recorded values of the study

	N	Mean	Std. Deviation	Minimum	Maximum
Age (year)	270	45.55	16.34	17.00	85.00
Height(cm)	270	168.99	9.58	145.00	190.00
Weight(kg)	270	77.39	16.77	40.00	175.00
Arne	270	2.74	3.76	0.00	16.00
TSD(cm)	270	6.04	0.88	3.00	8.00
TMD(cm)	270	6.67	0.89	4.00	9.50

Table 1- Quantitative and qualitative recorded values of the study (Continued)

	N	Mean	Std. Deviation	Minimum	Maximum
Naguib	270	2.24	2.75	-8.89	9.78
Wilson	270	0.50	0.83	0.00	4.00

		Frequency	Percentage	Valid Percentage	Cumulative Percentage
Gender	Men	137	50.7	50.7	50.7
	Women	133	49.3	49.3	100.0
Mallampati score	3 or 4	23	8.5	8.5	8.5
	1 or 2	247	91.5	91.5	100.0
Cormack-Lehane Grade	1	168	62.2	62.2	62.2
	2	55	20.4	20.4	82.6
	3	46	17.0	17.0	99.6
	4	1	0.4	0.4	100.0
Attempts	1	213	78.9	78.9	78.9
	2	53	19.6	19.6	98.5
	3	4	1.5	1.5	100.0

TSD, thyrosternal distance; TMD, thyromental distance

Table 2- Chi-square table showing comparison between C-L grades vs. four preoperative Predictors.

Model		C-L Grade		Total	Chi2, P_value
		Easy	Difficult		
TMHT	Easy	205	4	209	154.45, p<0.001
	Difficult	18	43	61	
	Total	223	47	270	
Naguib	Easy	194	29	223	17.27, p<0.001
	Difficult	29	18	47	
	Total	223	47	270	
Arne	Easy	220	40	260	0.304, p=0.582
	Difficult	3	7	10	
	Total	223	47	270	
Wilson	Easy	222	46	268	1.48,
	Difficult	1	1	2	P(fisher exact
	Total	223	47	270	test)=0.318

TMHT, Thyromental height test; C-L, Cormack-Lehane

Table 3- Comparison of the final results for the four predictors to predict the occurrence of a grade 3 or 4 of C-L Classification.

Model	TP	TN	FP	FN	Sensitivity	PPV	Specificity	NPV	Accuracy
TMHT (≤50mm)	43	205	18	4	91.49%	70.49%	91.93%	98.09%	91.85%
Naguib (<0)	18	194	29	29	38.3%	38.3%	87%	87%	78.52%
Arne (>11)	7	220	3	40	14.89%	70%	98.65%	84.62%	84.07%
Wilson (≥4)	1	222	1	46	2.12%	50%	99.55%	82.83%	82.59%

TMHT= Thyromental height test; TP= true positive; TN= true negative; FP= false positive; FN= false negative; PPV= positive predictive value NPV= negative predictive value

Discussion

This is the first study through which three multifactorial clinical models were compared with a univariate clinical test named TMHT. The TMHT, as a univariate anatomical index, is a more accurate predictor of difficult laryngoscopy

than other multivariable models. The incidence of difficult laryngoscopy in our study was 17.4% which is comparable with other researches [1,3,6,9-11]. However, there is an important question that why the incidence of difficult laryngoscopy increased in this study in comparison to the study conducted by Etezadi and associates in 2013 at the same hospital (17.4% Vs 7.3%) [10]. The answer may be clear if we consider the difference in designed methodology of the current study in which the application of BURP has been omitted during laryngoscopy attempts. While BURP was applied for all secondary laryngoscopy attempts in the previous study done by the investigators. It has been already revealed that application of BURP maneuver will push the glottis aperture downward and into the visual axis of laryngoscopist and subsequently make laryngoscopy easier and may lower the C-L grade [16-18].

A test to predict difficult intubation should have a high sensitivity, so that it will exactly predict most patients in whom intubation will truly be difficult. It means that it also has a high Positive Predictive Value (PPV) and a low rate of False Negative (FN) predictions so that only few patients with airways actually difficult to intubate are subjected to the protocol for management of an easy airway. It is noteworthy to mention that having a high specificity as well as a high Negative Predictive Value (NPV) finally may lead to a low rate of False Positive (FP) prediction. A FP result may increase our concern of having being faced with a difficult airway that actually is not that difficult to manage. The consequence of a FN prediction may lead to a disastrous situation. Thus, decreasing FN predictions (increased sensitivity) is more important than lowering the rate of FP predictions (increased specificity) for patients [14-19]. In other words, we don't lose anything with too much FP prediction except money, while we may encounter a mortality even with one FN prediction.

According to table 2 we find that both TMHT and Naguib test are reliable to predict laryngoscopy, while if we take a look at table 3, we can suggest that the TMHT is superior to other models in terms of accuracy. TMHT shows the highest sensitivity, although the specificity of both Arne and Wilson are higher than TMHT. Naguib et al. reported the sensitivity of three multivariate clinical models (Naguib, Arne and Wilson) as 81.4%, 54.6 %, and 40.2%, respectively which are significantly higher than what the current study shows [12]. However, what is common between results of this study and the Naguib's study is the trend of sensitivity values among multivariate tests that the highest and the lowest sensitivities belong to the Naguib and Wilson respectively. In addition, the Naguib's scoring system gets the lowest accuracy (78.52%) among multivariate tests in the current study.

TMHT has the highest NPV (98.09%) among four models (87%, 84.62% and 82.83% for Naguib, Arne and Wilson respectively). In addition, the rate of PPV is reported 38.3%, 70%, 50% and 80.49% for Naguib, Arne, Wilson and TMHT respectively, that is the highest for TMHT in comparison to others.

Regarding the importance of time management especially in preoperative clinic, the need for a simple test to predict difficult laryngoscopy is getting more important than ever. It's obvious that application of a single factor test like TMHT is less time-consuming and looks more implementable than using multifactorial models which are very lengthy to perform. In addition, we should pay attention to this point that having more variables in a predictor test may lead to a higher likelihood of measurement bias. In a prior research, it has been already shown that there is a very small inter-observer variability with TMHT [10].

Our study is subject to one limitation: The study population was limited to patients scheduled for elective, non-emergent surgeries. Thus, the results are only applicable to this group of patients.

Conclusion

In summary, although both TMHT and Naguib multivariable model showed promising results in prediction of difficult laryngoscopy, the TMHT was found to be a more accurate predictor of difficult laryngoscopy than other predicting models. However, because TMHT is a recently introduced method, it should be reevaluated through further studies within other patients' population like obstetric patients who are more prone to have a difficult laryngoscopy than others.

Acknowledgements

Authors would like to thank Research Development Center of Sina Hospital for their technical assistance.

References

1. Naguib M, Scamman FL, O'Sullivan C, Aker J, Ross AF, Kosmach S, et al. Predictive Performance of Three Multivariate Difficult Tracheal Intubation Models: A Double-Blind, Case-Controlled Study. Anesth Analg. 2006; 102(3):818-24.
2. King TA, Adams AP. Failed tracheal intubation. Br J Anaesth.1990; 65(3):400-14.
3. National Collaborating Centre for Women's and Children's Health (UK). Pregnancy and Complex Social Factors: A Model for Service Provision for Pregnant Women with Complex Social Factors. London: RCOG Press; 2010 Sep. (NICE Clinical Guidelines, No.110.) References Available from: http://www.ncbi.nlm.nih.gov/books/NBK62619/
4. Wilson ME, Spiegelhalter D, Robertson JA, Lesser P. Predicting difficult intubation. Br J Anaesth. 1988; 61(2): 211-6.
5. Samsoon GL, Young JR. Difficult tracheal intubation: a retrospective study. Anaesthesia. 1987; 42(5): 487-90.
6. Benumof JL. Management of the difficult adult airway: with special emphasis on awake tracheal intubation. Anesthesiology. 1991; 75:1087–110
7. Tse JC, Rimm EB, Hussain A. Predicting Difficult Endotracheal Intubation in Surgical Patients Scheduled for General Anesthesia: A Prospective Blind Study. Anesth Analg. 1995;81(2):254-8.
8. 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(2):140-6
9. Naguib M, Malabarey T, AlSatli RA, Al Damegh S, Samarkandi AH. Predictive models for difficult laryngoscopy and intubation: a clinical, radiologic and three-dimensional computer imaging study. Can J Anaesth. 1999; 46(8): 748-59
10. Etezadi F, Ahangari A, Shokri H, Najafi A, Khajavi MR, Daghigh M, et al. Thyromental height: a new clinical test for prediction of difficult laryngoscopy. Anesth Analg. 2013;117(6):1347-51.
11. Cormack RS, Lehane J. Difficult tracheal intubation in obstetrics. Anaesthesia. 1984; 39(11): 1105-11.
12. 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(6):1197-204.
13. Merah NA, Wong DT, Foulkes-Crabbe DJ, Kushimo OT, Bode CO. Modified Mallampati test, thyromental distance and interincisor gap are the best predictors of difficult laryngoscopy in West Africans. Can J Anaesth. 2005; 52(3):291–6
14. Krobbuaban B, Diregpoke S, Kumkeaw S, Tanomsat M. The predictive value of the height ratio and thyromental distance: four

predictive tests for difficult laryngoscopy. Anesth Analg. 2005; 101(5):1542–5
15. Khan ZH, Mohammadi M, Rasouli MR, Farrokhnia F, Khan RH. The diagnostic value of the upper lip bite test combined with sternomental distance, thyromental distance, and interincisor distance for prediction of easy laryngoscopy and intubation: a prospective study. Anesth Analg. 2009; 109(3): 822–4.
16. Knill RL. Difficult laryngoscopy made easy with a "BURP". Can J Anaesth. 1993; 40(3):279-82.
17. Takahata O, Kubota M, Mamiya K, Akama Y, Nozaka T, Matsumoto H, et al. The efficacy of the "BURP" maneuver during a difficult laryngoscopy. Anesth Analg. 1997; 84(2): 419–21
18. Ulrich B, Listyo R, Gerig HJ, Gabi K, Kreienbühl G. The difficult intubation. The value of BURP and 3 predictive tests of difficult intubation. Anaesthesist. 1998; 47(1):45-50.
19. Shiga T, Wajima Z, Inoue T, Sakamoto A. Predicting difficult intubation in apparently normal patients: a meta-analysis of bedside screening test performance. Anesthesiology. 2005; 103(2): 429–37.