Anaesth Intensive Care 1998; **26:** 382-386

Screening Tests for Predicting Difficult Intubation. A Clinical Assessment in Turkish Patients

H. BILGIN*, G. ÖZYURT†

Department of Anaesthesiology, Uludag University Medical School, Bursa, Turkey

SUMMARY

Three methods of predicting difficult intubation were compared prospectively. Mallampati test, Wilson risk-sum and thyromental distance were determined preoperatively and laryngeal views were graded in 500 patients. The sensitivities, specificities, positive and negative predictive values of each test were calculated. The incidence of difficult intubation was found to be 8%. The sensitivities of the Mallampati test, the Wilson risk-sum and the thyromental distance were 43%, 58% and 35% respectively, and the specificities were 93%, 91% and 95% respectively. Significant differences were seen between the sensitivities of the Mallampati test and the Wilson risk-sum ($P<0.001$), the Wilson risk-sum and the thyromental distance ($P<0.001$), the Mallampati test and the thyromental distance ($P<0.05$). Among the different specificities, the only significant difference was observed between the Wilson risk-sum and the thyromental distance ($P<0.05$). There were no significant differences between the positive and negative predictive values of the three screening tests ($P>0.05$). In conclusion, the Wilson risk-sum was the most sensitive test and the thyromental distance has the highest positive predictive value for difficult intubation.

Key Words: INTUBATION, LARYNGOSCOPY: difficult, Mallampati test, thyromental distance, Wilson risk-sum

Unexpected difficulty in visualizing the glottis during tracheal intubation may cause hypoxia leading to severe brain damage and death. In different studies the incidence of difficult or failed intubation ranges from 1% to 13%[1-4]. It has been estimated that in the developed world alone, 600 deaths occur annually from complications at the time of tracheal intubation[5].

Early identification of patients in whom orotracheal intubation may be difficult or impossible may reduce mortality and morbidity. Although many tests have been described, the results of each test vary from one study to another. Therefore it has not been possible to form a judgement as to which test is the best predictor of difficult intubation.

The purpose of this study was to compare the screening tests for difficult intubation described by various authors in terms of their sensitivities, specificities and positive and negative predictive values. We have applied these tests to a group of 500 Turkish patients.

*M.D., Staff Anaesthesiologist.
†M.D., Professor of Anaesthesiology.

Address for Reprints: Professor Gürayten Özyurt, Uludağ Üniversitesi Tip Fakultesi, Anesteziyoloji ve Reanimayson A.B.D., 16059 Bursa, Turkey.

Accepted for publication on 19 December, 1997.

MATERIALS AND METHODS

Written informed consent was obtained from 500 surgical patients (247 male, 253 female ASA-physical status I or II), scheduled for general anaesthesia requiring tracheal intubation. The age and weight of each patient was recorded. None of the patients had known abnormalities of the upper airway or head and neck trauma. Each patient was evaluated preoperatively by one of four individuals (1 staff anaesthesiologist, 3 anaesthesiology residents) having at least two years clinical experience in anaesthesia.

The Mallampati test (MT) was performed by asking each patient to protrude the tongue maximally from a fully open mouth while sitting upright. Patients were divided into three classes:

Class I: Faucial pillars, soft palate and uvula could be visualized.

Class II: Faucial pillars and soft palate could be visualized, but uvula was masked by the base of the tongue.

Class III: Only soft palate could be visualized[6].

Each patient was requested to perform the manoeuvre at least twice to decrease the chance of error.

Secondly, the Wilson risk-sum (WRS) was calculated. Five risk factors of this score were weight, head, neck and jaw movements, mandibular recession and the presence or absence of protruding teeth. Each

factor scored from 0 to 2, giving a total ranging from 0 to 10 (Table 1)[7].

TABLE 1
Wilson risk-sum scoring system

Risk Factor	Level
Weight	0 <90 kg
	1 90-110 kg
	2 >100 kg
Head and neck movement	0 Above 90°
	1 About 90° (i.e. ±10°)
	2 Below 90°
Jaw movement	0 IG* >5 cm** or SLux*** >0
	1 IG <5 cm and SLux=0
	2 IG <5 cm and SLux <0
Receding mandible	0 Normal
	1 Moderate
	2 Severe
Buck teeth	0 Normal
	1 Moderate
	2 Severe

*IG: inter-incisor gap.
**5 cm is approximately three fingers breadth.
***SLux: subluxation (maximal forward protrusion of the lower incisors beyond the upper incisors).

The thyromental distance (TMD) was measured as the straight line distance between the thyroid notch and the lower border of the mentum with the head in full extension[8]. A ruler was used for the measurements which were then rounded to the nearest 0.5 cm. In an average-sized adult, when the thyromental distance is shorter than six cm with the head fully extended, laryngoscopy is considered "impossible"[8].

After the induction of anaesthesia and administration of a neuromuscular blocking agent, laryngoscopy was performed using a Macintosh blade size 3, with the head in the "sniffing position" on a pillow. Laryngoscopy was graded on a scale of I-IV as described by Cormack and Lehane[9]. Grades III and IV were taken to signify a difficult laryngoscopy. External laryngeal pressure was applied, if necessary, to improve the view on the glottis of those patients who had laryngoscopy grades III and IV.

Intubation of the trachea was considered as difficult:

—if the patient's trachea could not be intubated without the use of a blade other than the Macintosh or Miller blade and special instruments (e.g. stylet, guide, fibreoptic laryngoscope) or the need for external laryngeal pressure;
—when intubation of the trachea could only be performed with an endotracheal tube narrower in diameter than the initially estimated tube size;
—when a patient's trachea had to be intubated by an experienced staff anaesthesiologist after at least three unsuccessful attempts by a senior anaesthesiology resident.

All patients were evaluated on the basis of these tests, and the sensitivity, specificity, positive and negative predictive values of each test were calculated (*sensitivity:* the proportion of difficult intubations correctly predicted to be difficult; *specificity:* the proportion of easy intubations correctly predicted to be easy, *positive predictive value:* the proportion of predicted difficult intubations that actually proved to be difficult; *negative predictive value:* the proportion of predicted easy intubation that actually proved to be easy[10]).

The results were analysed with "Z test" and "McNemar test"[11].

RESULTS

Patients' ages ranged from 16-80 years (mean 45.6 years). Their weights ranged from 45-110 kg (mean 63.1 kg).

Difficult intubation was observed in 40 patients. Of the 500 patients studied 391 were allocated to Mallampati class I, 58 to Mallampati class II and 51 to class III. The distribution of the patients on the basis of laryngoscopic grades is shown in Table 2. The distribution of the 40 patients with difficult intubation in relation to their Mallampati test results and laryngoscopic grading is shown in Table 3. Seventeen of the 51 patients who were expected to have difficult intubation according to the Mallampati test were found to actually have difficult intubations. This result represents a positive predictive value 33%. It identified 51 patients who would have a difficult intubation, but in fact 34 of them had an easy intubation. The test correctly predicted 43% of all difficult intubations (its sensitivity). On the other hand, the Mallampati classification did not predict 57% of difficult intubations. The Mallampati test correctly predicted 93% of the easy intubations (its specificity). Negative predictive value of this test was found to be 95% (Table 4).

There were 437 patients with a total score equal to

TABLE 2
Relationship between the results of three screening tests and laryngoscopic grading

Screening test		Cormack and Lehane laryngoscopic grade			
		Grade I	Grade II	Grade III	Grade IV
MT	Class I	362	29	–	–
	Class II	34	17	7	–
	Class III	2	10	38	1
WRS	Score ≤ 2	394	36	7	–
	Score >2	4	20	38	1
TMD	TMD ≥ 6 cm	391	43	30	–
	TMD <6 cm	7	13	15	1

MT: Mallampati test
WRS: Wilson risk-sum
TMD: Thyromental distance

TABLE 3

Relationship between the results of three screening tests and laryngoscopic grading of 40 patients with difficult intubation

Screening test		Cormack and Lehane laryngoscopic grade			
		Grade I	Grade II	Grade III	Grade IV
MT	Class I	–	8	–	–
	Class II	–	10	5	–
	Class III	–	5	11	1
WRS	Score ≤ 2	1	11	5	–
	Score >2	–	5	17	1
TMD	TMD ≥ 6 cm	4	6	16	–
	TMD <6 cm	1	1	11	1

MT: Mallampati test
WRS: Wilson risk-sum
TMD: Thyromental distance

or less than 2 on the Wilson risk-sum, 63 patients had a score greater than 2. The Wilson risk scores of the patients and their distribution according to laryngoscopic grades are shown in Table 2. Twenty-three of the 40 patients who had difficult intubation had a score greater than 2. The Wilson risk scores of patients who had difficult intubation and their distribution according to laryngoscopic grading are shown in Table 3. The sensitivity, specificity, positive and negative predictive values were 58%, 91%, 37% and 96% respectively (Table 4).

In 36 of the 500 patients, the thyromental distance was less than 6 cm. The distribution of the patients according to the length of the thyromental distance and laryngoscopic grading is shown in Table 2. The findings in forty patients with difficult intubations, laryngoscopic grading and length of the thyromental distance are shown in Table 3. The thyromental distance and its significance as a screening test (sensitivity, specificity, positive and negative predictive values) are shown in Table 4.

By the application of the Cormack and Lehane laryngoscopic classification, 398 patients were determined as grade I, 56 patients as grade II, 45 patients as grade III and one patient as grade IV (Table 2). Comparisons of the sensitivity, specificity, positive and negative predictive values for each screening test are shown in Table 5.

In Turkish patients, we found the sensitivity of the Wilson risk-sum to be significantly greater than the Mallampati test and the thyromental distance in predicting difficult intubations ($P<0.001$, $P<0.001$ respectively). The sensitivity of the Mallampati test was significantly greater than the thyromental distance ($P<0.05$). The specificity of the Wilson risk-sum was significantly higher than the thyromental distance ($P<0.05$). We could not find any significant difference between positive and negative predictive values of the three screening tests ($P>0.05$).

DISCUSSION

Predicting difficult intubation can reduce anaesthesia associated morbidity and mortality. In order to be clinically useful, a test predicting difficult intubation must be easily applicable at the bedside and must give reliable results. No test has 100% sensitivity and there will always be some patients with unpredicted difficult intubation, but high sensitivity and high positive predictive value are desirable. A test to predict difficult intubation should have high sensitivity so that it will identify most patients in whom intubation will truly be difficult. It should also have a high

TABLE 4

The incidence of predicted and rated difficult intubations in relation to the three screening tests

Screening test	Suspected to be difficult	Really difficult cases	False positive	False negative	Sens %	Spec %	PPV %	NPV %
MT	51/500	17/40	34	23	43	93	33	95
WRS	63/500	23/40	40	17	58	91	37	96
TMD	36/500	14/40	22	26	35	95	39	94

MT: Mallampati test
WRS: Wilson-risk sum
TMD: Thyromental distance
Sens: Sensitivity
Spec: Specificity
PPV: Positive predictive value
NPV: Negative predictive value

TABLE 5

Comparison of sensitivity, specificity, positive and negative predictive value of screening tests

Screening tests	Sensitivity		Specificity		Positive predictive value		Negative predictive value	
	Z value	*P* value	Z value	*P* value	Z value	*P* value	Z value	*P* value
MT-WRS	–4.74	<0.001	–0.76	NS	–1.06	NS	–0.93	NS
MT-TMD	2.43	<0.05	–1.73	NS	–1.84	NS	0.33	NS
WRS-TMD	7.13	<0.001	–2.47	<0.05	–0.77	NS	1.27	NS

MT: Mallampati test
WRS: Wilson risk-sum
TMD: Thyromental distance
NS: Nonsignificant

positive predictive value so that only a few patients who can be actually intubated easily are subjected to the protocol for management of a difficult intubation.

We found an 8% incidence of difficult intubation in this study and there were no failures to intubate the trachea. Although some authors blame different anthropometric features existing between populations as the cause of the discrepancies in the incidence of difficult intubation in different studies, such differences may be due to the fact that sometimes the cases in which pressure was applied to the larynx are excluded from the "difficult intubation" group. We included patients into the difficult intubation group if pressure to the cricoid cartilage had to be applied during laryngoscopy. In another study, Langenstein and Cunitz defined an intubation as difficult, if a practising anaesthesiologist needed more than three attempts or more than ten minutes for a successful endotracheal intubation[12]. We changed this criterion to endotracheal intubation performed by an experienced staff anaesthesiologist after at least three unsuccessful attempts by a senior anaesthesiology resident. Mallampati et al gave an incidence of 13% for difficult laryngoscopy. The definition of difficult laryngoscopy, however, included grade III laryngoscopy. The incidence was 4.3% if grade III laryngoscopy was excluded[6]. Therefore the incidence of difficult intubation can be different depending on the criteria which are used to define it.

We found the sensitivity of the Mallampati test to be lower than the Wilson risk-sum, but the specificities, positive and negative predictive values of these tests were similar. The sensitivity of the Mallampati test has been reported to be 67.9%, 42% and 56% in other studies[1,13,14]. This value was found to be 43% in our study. The great difference between the reported sensitivities suggests an incorrect evaluation of the patients. Oates et al reported that a critical factor in achieving a reliable score for the Mallampati test is ensuring that the patient opens his mouth and protrudes his tongue maximally. Failure to apply this rigorously is a major pitfall when performing the assessment. In that study, sensitivity, specificity and positive predictive value for the Mallampati test changed from 42% to 50%, 84% to 92% and 4.4% to 10% respectively, depending on which anaesthesiologist made the assessment[1]. In addition, recent investigations have shown that many patients involuntarily phonate during the test, which may significantly alter the Mallampati grading[15]. Tham and colleagues observed no differences on grading with the patients' posture. And they concluded the test is still useful in emergency cases when the patient is supine or unable to sit[16].

The present study found the Wilson risk-sum to have the highest sensitivity and negative predictive value, and the second highest positive predictive value among the three tests. Wilson et al reported a predictive rate of 80% when they were describing this scoring system[7]. However, subsequent studies have failed to support these results. Oates et al reported that both the Wilson risk-sum and the Mallampati test failed to predict as many as 58% of difficult laryngoscopies[1]. Although a Wilson risk-sum score greater than 2 may correctly predict 75% of difficult intubations, this may also lead to an unacceptably high number of false alarms[7]. Adjusting the "difficult intubation" category to a score greater than 4 reduces the number of false alarms, but may also increase the chance of missing difficult intubations. In addition, Wilson et al also found that the weight of the patient was the least useful of the five predictive factors[7]. In another study, by removing weight from the risk-sum, the sensitivity was unchanged at 42% and positive predictive value improved from 8.9% to 10.9%. These authors suggested that the weight component of the Wilson risk-sum might be unnecessary[1]. In the study of Mallampati et al, four patients weighing 130 kg or more had adequate glottic exposure[6]. We did not observe any patient weighing more than 110 kg. Concern has been expressed that the Mallampati test may not distinguish those patients in whom difficult laryngoscopy is caused by limited head and neck mobility[3,17].

We found that thyromental distance had the lowest sensitivity and negative predictive value, and the highest specificity and positive predictive value. In Frerk's study the sensitivity and specificity of the thyromental distance were 90.9% and 81.5% respectively[18]. In contrast, Tse and colleagues reported that the sensitivity of thyromental distance was 32% and the positive predictive value was 20%[4]. Although both of them defined an intubation as difficult when the thyromental distance was 7 cm or less, Tse et al explained that the discrepancy between their findings and those of Frerk's were partly due to the different definitions used for difficult intubation[4]. Butler and Dhara reported that when they used the thyromental distance of 6 cm as the predictor of difficult laryngoscopy, the values for sensitivity, specificity and positive predictive value were 62%, 25% and 16% respectively. However, they found a large spectrum of thyromental distances on either side of 6 cm and there was no correlation with laryngoscopic grading[14]. It should be remembered that there are other factors responsible for a short thyromental distance such as a high anterior larynx, head extension or a receding mandible[18].

We now know that certain anatomical features are associated with difficult intubation. Short muscular neck, receding mandible, protruding maxilla or upper incisors, long high arched palate with a narrow mouth, cervical rigidity, and an increased posterior depth of the mandible are some of these. Although fully aware of these factors, we could not accurately predict difficult intubations preoperatively. In a recent study, the combination of screening tests for predicting difficult intubation was reported as unreliable[4]. So, which test is the most reliable in clinical practice? We propose the Wilson risk-sum whose sensitivity is the highest and positive predictive value is the second highest among these three tests. The thyromental distance has the highest positive predictive value.

REFERENCES

1. Oates JDL, Macleod AD, Oates PD, Pearsall FJ, Howie JC, Murray GD. Comparison of two methods for predicting difficult intubation. Br J Anaesth 1991; 66:305-309.
2. Rocke DA, Murray WB, Rout CC, Gouws E. Relative risk analysis of factors associated with difficult intubation in obstetric anesthesia. Clinical investigation. Anesthesiology 1992; 77:67-73.
3. Samsoon GLT, Young JRB. Difficult tracheal intubation: a retrospective study. Anaesthesia 1987; 42:487-490.
4. Tse JC, Rimm EB, Hussain A. Predicting difficult endotracheal intubation in surgical patients scheduled for general anaesthesia. A prospective blind study. Anesth Analg 1995; 81:254-258.
5. Brain AIJ. Further developments of the laryngeal mask. Anaesthesia 1989; 44:530.
6. Mallampati SR, Gatt SP, Gugino LD, et al. A clinical sign to predict tracheal intubation: a prospective study. Can Anaesth Soc J 1985; 32:429-434.
7. Wilson ME, Spiegelhalter D, Robertson JA, Lesser P. Predicting difficult intubation. Br J Anaesth 1988; 61:211-216.
8. Mathew M, Hanna LS, Aldrete JA. Preoperative indices to anticipate difficult tracheal intubation. Anesth Analg 1989; 68:S187.
9. Cormack RS, Lehane J. Difficult tracheal intubation in obstetrics. Anaesthesia 1984; 39:1105-1111.
10. Morton RF, Hebel JR. A study guide to epidemiology and biostatistics including 100 multiple choice questions. 2nd ed. University Park Press, Baltimore 1984; 59-66.
11. Everitt BS. The analysis of contingency tables. Halsted Press, Division of John Willy and Sons, New York 1980; 20-22.
12. Langenstein H, Cunit ZG. Difficult intubation in adults. Anaesthetist 1996; 45:372-383.
13. Yamamoto K, Tsubokawa T, Shibata K, Ohmura S, Nitta S, Kobayashi T. Predicting difficult intubation with indirect laryngoscopy. Anesthesiology 1997; 86:316-321.
14. Butler PJ, Dhara SS. Predicting of difficult laryngoscopy: An assessment of the thyromental distance and Mallampati predictive test. Anaesth Intensive Care 1992; 20:139-142.
15. Oates JDL, Oates PD, Pearsall FJ, Macleod AD, Howie JC. Phonation affects Mallampati classification. Anaesthesia 1990; 45:984-1004.
16. Tham EJ, Gildersleve CD, Sanders LD, Mapleson WW, Vaughan RS. Effects of posture, phonation and observer on Mallampati classification. Br J Anaesth 1992; 68:32-38.
17. Bellhouse CP, Dore C. Criteria for estimating the likelihood of difficulty of endotracheal intubation with the Macintosh laryngoscope. Anaesth Intensive Care 1988; 16:329-337.
18. Frerk CM, Till CBW, Bradley AJ. Difficult intubation: Thyromental distance and the atlanto-occipital gap. Anaesthesia 1996; 51:738-740.