British Journal of Anaesthesia 1991; **66**: 305–309

COMPARISON OF TWO METHODS FOR PREDICTING DIFFICULT INTUBATION†

J. D. L. OATES, A. D. MACLEOD, P. D. OATES, F. J. PEARSALL, J. C. HOWIE AND G. D. MURRAY

SUMMARY

Two methods of predicting difficult laryngoscopy were compared prospectively. Mallampati class and Wilson risk-sum were determined before operation and laryngeal view graded in 675 patients. Both tests identified five of 12 difficult laryngoscopies; twice as many patients were predicted to be difficult by Mallampati classification than by Wilson risk-sum. Inter-observer variation was minimal using Wilson risk-sum, but considerable for Mallampati classification. We prefer the Wilson risk-sum for assessment of the airway, while noting that both tests have poor sensitivities.

KEY WORDS

Complications: difficult intubation Intubation, tracheal: preoperative assessment.

The unexpected difficult intubation is a major concern for the anaesthetist. Recently, several different methods of predicting possible difficulty in intubation have been devised [1]. Two of the most interesting are those of Mallampati and co-workers [2] and Wilson and colleagues [3]. The Mallampati test classifies the view of pharyngeal structures with the mouth wide open and the tongue protruded. It is a simple test and may be performed quickly. The Wilson test is more complex and assesses five factors, including weight and head and neck mobility.

The purpose of this study was to compare the Mallampati classification with Wilson risk-sum in the prediction of difficult laryngoscopy, and to assess inter-observer variation in performing these tests.

PATIENTS AND METHODS

In a prospective study, 751 patients (448 female) requiring tracheal intubation were seen before operation by one of four observers—three registrars and one senior registrar (training grades). Sex, age and weight of each patient were recorded. All patients were assessed for Mallampati class and Wilson risk-sum.

Mallampati classification was performed as described originally by asking the patient to protrude the tongue maximally while in the sitting position [2]. In class 1, faucial pillars, soft palate and uvula are visible. In class 2, faucial pillars and soft palate may be seen, but the uvula is masked by the base of the tongue. In class 3, only the soft palate is visible. The manoeuvre was repeated to ensure maximum visibility of the pharyngeal structures.

Wilson risk-sum was then ascertained. This scores five factors (weight, head and neck movement, jaw movement, receding mandible and buck teeth) from 0 to 2, giving a total ranging from 0 to 10 [3].

Anaesthetists of all grades, who were not informed of the preoperative Mallampati class and Wilson risk-sum, assessed difficulty of laryngoscopy at intubation. Laryngoscopy was graded on a scale of 1–5 as described by Wilson and colleagues [3]: grade 1 = almost all of cords visible; grade 2 = only half of cords visible; grade 3 = only arytenoids visible; grade 4 = only

JONATHAN D. L. OATES, B.M., D.A.(U.K.), F.F.A.R.C.S.I.; ALEX D. MACLEOD, M.B., CH.B., M.SC., F.C.ANAES.; PHILIP D. OATES, M.B., CH.B., D.A.(U.K.), F.F.A.R.C.S.I.; FIONA J. PEARSALL, M.B., CH.B., D.A.(U.K.); J. CAMERON HOWIE, B.SC.(HONS), M.B., CH.B., M.R.C.P., F.C.ANAES.; Department of Anaesthetics, Victoria Infirmary, Glasgow G42 9TY. GORDON D. MURRAY, M.A., PH.D., Medical Statistics Unit, Department of Surgery, Western Infirmary, Glasgow G11 6NT. Accepted for Publication: September 25, 1990.

Correspondence to J.D.L.O.

† Presented in part to the Scottish Society of Anaesthetists, Peebles, April 1990.

TABLE I. *Wilson risk-sum values (0–5) and laryngoscopy grades (1–5) in 677 patients*

Wilson risk-sum	Laryngeal view				
	1	2	3	4	5
0	388	70	11	5	1
1	104	35	6	0	1
2	26	13	1	1	0
3	4	3	2	3	0
4	1	0	0	0	1
5	1	0	0	0	0

TABLE II. *Mallampati class (1–3) and laryngoscopy grades (1–5) in 675 patients*

Mallampati class	Laryngeal view				
	1	2	3	4	5
1	398	84	17	5	1
2	49	4	2	1	0
3	76	33	0	3	2

epiglottis visible; grade 5 = not even epiglottis visible. Difficult laryngoscopy was defined as grades 4 and 5—that is, no more than the epiglottis could be seen.

TABLE III. *Sensitivity, specificity and positive predictive value (PPV) for Wilson risk-sum and Mallampati classification in 675 patients assessed by one of four observers and in 543 patients for whom data from observer four are excluded. A Wilson risk-sum of 2 or more, or Mallampati class 3 are predictors, and laryngoscopy grades 4 or 5 are diagnosis, of difficult laryngoscopy*

	Wilson	Mallampati
All observers (675 patients)		
Sensitivity	0.42	0.42
Specificity	0.92	0.84
PPV (%)	8.9	4.4
Observers 1–3 (543 patients)		
Sensitivity	0.40	0.50
Specificity	0.95	0.92
PPV (%)	13.3	10.0

Sensitivity, specificity and positive predictive value were calculated for each test. Sensitivity is the proportion of difficult laryngoscopies which were predicted to be difficult. Specificity is the proportion of straightforward laryngoscopies which were predicted to be easy. Positive predictive value is the proportion of patients predicted to be difficult who proved to be difficult [4].

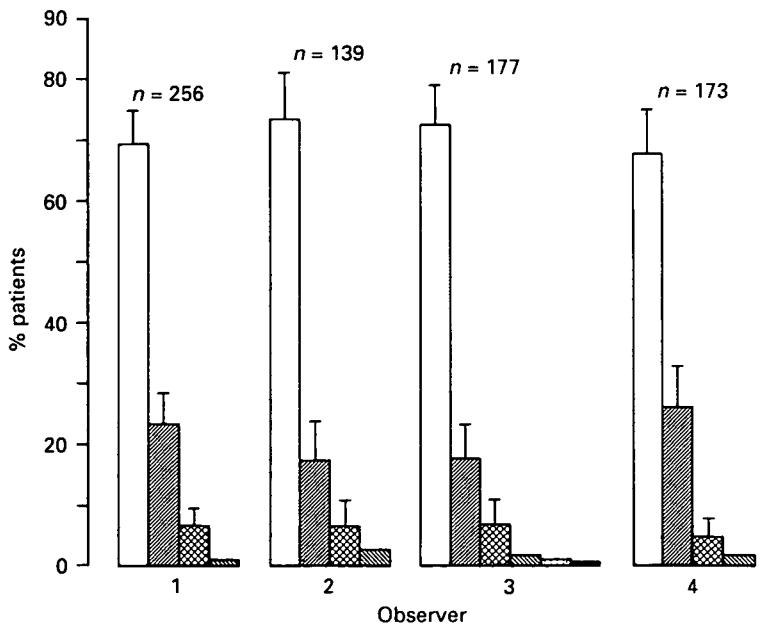

FIG. 1. Distribution of scores for Wilson risk-sum by observers 1–4 (95% confidence intervals). *n* = Number of patients observed. Risk-sums: □ = 0; ▨ = 1; ⊠ = 2; ▧ = 3; ⊡ = 4; ■ = 5.

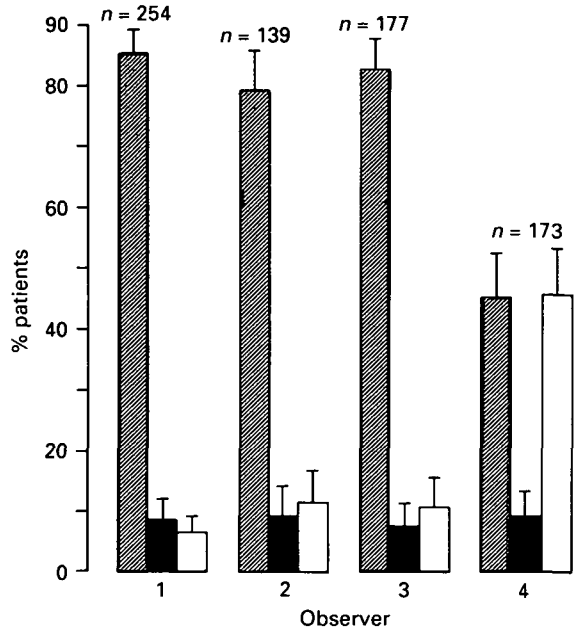

FIG. 2. Distribution of scores for Mallampati class by observers 1–4 (95% confidence intervals). n = Number of patients observed. Mallampati class: ▨ = 1; ■ = 2; □ = 3.

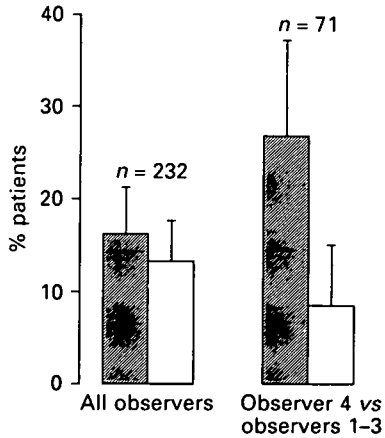

FIG. 3. Percentage (95% CI) of patients, assessed by two observers, for whom the observers disagreed over scoring difficult or easy laryngoscopy for Mallampati classification (▨) and Wilson risk-sum (□). Left: all observers. Right: Observer 4 *vs* observers 1–3. n = Number of patients observed.

Scores for Mallampati class and Wilson risk-sum were examined to assess the distribution of scores assigned by each observer. In addition, 232 patients were assessed before operation by two of the observers independently, so that a direct between-observer comparison of scoring could be made.

Ninety-five percent confidence intervals (95% CI) were calculated where appropriate.

RESULTS

Complete data were available for 675 of 751 patients. Patient age ranged from 16 to 87 yr (mean 46.6 yr). Weight ranged from 30 to 133 kg (mean 65.1 kg); 38 patients weighed 90 kg or more and one patient weighed more than 110 kg.

Tables I and II show Wilson risk-sum and Mallampati class, respectively, for each grade of laryngoscopy. Twelve patients were graded 4 or 5 at laryngoscopy, giving an incidence of difficult laryngoscopy of 1.8% (95% CI 0.77–2.75). Of these 12 patients, five had a Wilson risk-sum of 2 or more and five were Mallampati class 3. Sensitivities, specificities and positive predictive values for the two tests are shown in Table III.

Figures 1 and 2 show scoring of Wilson risk-sum and Mallampati class, respectively, by each observer. For all four observers, approximately 70% of patients had a Wilson risk-sum of zero while 30% had a risk-sum of one or more. Observers one, two and three classified over 75% of patients Mallampati 1, whereas observer four classified 45% of patients Mallampati 1 and 46% Mallampati 3.

For the 232 patients in whom pairs of the observers predicted, independently, straightforward or difficult laryngoscopy, we examined the number of occasions when the pair of observers agreed or disagreed in their predictions (fig. 3). There was no significant difference in the number of agreements and disagreements between the two observers when the data for all observers are used. However, when observer four was compared with the other observers, there were significantly more disagreements as to prediction of difficulty for the Mallampati classification compared with Wilson risk-sum.

Excluding the data for observer four, there remained 543 patients, of whom nine were found to be laryngoscopy grade 4 or 5. Sensitivities, specificities and positive predictive values for the two tests in these patients are shown in table III.

Seventy-three patients had restricted head and neck movement and seven of these scored two points on Wilson criteria (head and neck movement less than 90°). Three patients with restricted

head and neck movement presented difficulty at laryngoscopy; one was Mallampati class 1 and two were class 3.

DISCUSSION

The incidence of difficult laryngoscopy in this study was 1.8%. This agrees with the 1–3% quoted in previous work [5]. Wilson and colleagues [3] found an incidence of 1.5% using the same criteria for difficult laryngoscopy as this study. In contrast, Mallampati and co-workers [2] gave an incidence of 13%. Their definition of difficult laryngoscopy, however, included grade 3 laryngoscopy. The incidence was 4.3% if grade 3 laryngoscopy is excluded.

Examining the data for all four observers in our study, a Wilson risk-sum of 2 or more correctly predicted only five of the 12 difficult laryngoscopies. For a similar incidence of difficult laryngoscopy, Wilson and colleagues predicted nine of 12 difficult laryngoscopies. The positive predictive values were the same (8.9%).

Comparison with the work of Mallampati and co-workers requires clarification because of the different criteria used. In their study, class 3 predicted 14 of the 28 patients having grade 3 or 4 laryngoscopy ("inadequate exposure of the glottis"). Re-examining their data with class 3 as a predictor and taking grade 4 ("glottis including corniculate cartilages could not be exposed") to define difficult laryngoscopy, five of nine difficult laryngoscopies were predicted correctly. Positive predictive value was 33%. We were able to predict correctly only five of 12, with a positive predictive value of 4.4%.

In our study, using the data for all four observers, sensitivities for Wilson risk-sum and Mallampati classification were the same (0.42). Positive predictive value for Wilson risk-sum was double that for Mallampati classification (8.9% and 4.4%, respectively). This implies that to detect the same proportion of difficult laryngoscopies, twice as many patients were predicted to be difficult using the Mallampati classification as the Wilson risk-sum. However, our estimates of positive predictive values were based on small numbers, with consequent wide confidence intervals.

Observer four classified a significantly greater proportion of patients Mallampati class 3 than the other observers (fig. 2). This was confirmed by examining the number of times pairs of observers agreed or disagreed in predicting difficult laryngoscopy (fig. 3). There was a significantly greater proportion of disagreement between observers when observer four was one of the pair. Our findings on inter-observer variation in using the Mallampati classification support those of Wilson and John [6].

After discussion with observer four, we feel that the likely explanation of this is that full patient co-operation was not obtained. A critical factor in achieving a reliable score for Mallampati classification is ensuring that the patient opens the mouth and protrudes the tongue maximally. Failure to apply this rigorously is a major pitfall when performing the assessment. In addition, Mallampati class may be affected if the patient inadvertently phonates during the assessment [7]. We found that Wilson risk-sum and Mallampati classification were similar in their predictive value when the data for observer four were excluded.

Concern has been expressed that the Mallampati classification may not discriminate those patients in whom difficult laryngoscopy is caused by limited head and neck mobility [8, 9]. Mallampati and co-workers identified four patients with moderately restricted mobility of the neck [2]. Two of these had "inadequate glottic exposure". The Mallampati classifications of these patients were not specified. In our study, of the three patients who had limited head and neck movement and who presented difficulty at laryngoscopy, two were Mallampati class 3. In view of the small number, we remain guarded about the validity of the classification in such patients.

The initial study by Wilson and colleagues [3] found that weight was the least useful of the five predictive factors. In the study by Mallampati and co-workers [2], the four patients weighing 130 kg or more had adequate glottic exposure. We found that, in removing weight from the risk-sum, the sensitivity was unchanged at 0.42 and positive predictive value improved to 10.9%. This suggests that the weight component of the Wilson risk-sum might be unnecessary. Of the 38 patients who scored points for weight, however, only one patient scored 2 points (that is, weight more than 110 kg). We are cautious about extending our findings to patients weighing more than 110 kg and further investigation is required in this group.

Failure to intubate the trachea and maintain oxygenation remains a significant cause of perioperative death associated with anaesthesia. This has been shown repeatedly in studies of an-

aesthetic mortality. *The Report of a Confidential Enquiry into Perioperative Deaths* found that of the three deaths attributable wholly to anaesthesia, one was caused by inability to intubate the trachea [10]. Similarly, the *Report on Maternal and Perinatal Deaths in Scotland 1981–1985* identified one of the three deaths directly caused by anaesthesia as being a result of failure to achieve tracheal intubation [11].

We were unable to predict more than 50% of the difficult laryngoscopies. This is of particular importance to the relatively inexperienced anaesthetist presented with a patient at risk of regurgitation. Increasing the sensitivity of the test by lowering its threshold, however, also increases the proportion of false positives. For example, in our study if, instead of a Wilson risk-sum of 2 or more, a risk-sum of 1 or more was taken as a predictor of difficult laryngoscopy, 50% of difficult cases were identified, but the false positive rate increased from 0.08 to 0.29. Of the patients who were predicted to be difficult, only 3% proved to be difficult. This has obvious implications for the provision of senior anaesthetic cover, and deserves wider discussion.

In conclusion, we were unable to reproduce the results of the original studies. We found that the sensitivities of Wilson risk-sum and Mallampati classification were similar, but poor. At best it was possible to identify only 50% of the difficult laryngoscopies, with a high false positive rate. In view of the significant inter-observer variation in performing the Mallampati classification, we prefer the Wilson risk-sum in airway assessment.

REFERENCES

1. Vaughan RS. Airways revisited. *British Journal of Anaesthesia* 1989; **62**: 1–3.
2. Mallampati SR, Gatt SP, Gugino LD, Desai SP, Waraksa B, Freiberger D, Liu PL. A clinical sign to predict difficult tracheal intubation: a prospective study. *Canadian Anaesthetists' Society Journal* 1985; **32**: 429–434.
3. Wilson ME, Spiegelhalter D, Robertson JA, Lesser P. Predicting difficult intubation. *British Journal of Anaesthesia* 1988; **61**: 211–216.
4. The interpretation of diagnostic data. In: Sackett DL, Haynes RB, Tugwell P, eds. *Clinical Epidemiology: A Basic Science for Clinical Medicine*. Boston/Toronto: Little, Brown and Company, 1985; 59–138.
5. Latto IP. Management of difficult intubation. In: Latto IP, Rosen M, eds. *Difficulties in Tracheal Intubation*. London: Baillière Tindall, 1987; 99–141.
6. Wilson ME, John R. Problems with the Mallampati sign. *Anaesthesia* 1990; **45**: 486–487.
7. Oates JDL, Oates PD, Pearsall FJ, MacLeod AD, Howie JC. Phonation affects Mallampati class. *Anaesthesia* 1990; **45**: 984.
8. Samsoon GLT, Young JRB. Difficult tracheal intubation: a retrospective study. *Anaesthesia* 1987; **42**: 487–490.
9. Bellhouse CP, Doré C. Criteria for estimating likelihood of difficulty of endotracheal intubation with the Macintosh laryngoscope. *Anaesthesia and Intensive Care* 1988; **16**: 329–337.
10. Buck N, Devlin HB, Lunn JN. *The Report of a Confidential Enquiry into Perioperative Deaths*. London: Nuffield Provincial Hospitals Trust, 1987; 76–77.
11. *Report on Maternal and Perinatal Deaths in Scotland 1981–1985*. Edinburgh: Her Majesty's Stationery Office, 1989; 20.