Br. J. Anaesth. (1988), **61**, 211–216

PREDICTING DIFFICULT INTUBATION

M. E. WILSON, D. SPIEGELHALTER, J. A. ROBERTSON AND P. LESSER

SUMMARY

The amount of larynx seen at intubation was assessed in 633 adult patients undergoing routine surgery. Various measurements of the head and neck were made in an attempt to discover which features were associated with difficulty with laryngoscopy (defined as the inability to see even the arytenoids). In addition 38 patients, reported by colleagues because they had been "difficult to intubate", were measured. Five useful risk factors, measured at three levels of severity, were identified. A simple predictive rule was developed and tested on a prospective set of 778 patients, in 1.5% of whom laryngoscopy was found to be difficult. Depending on the threshold chosen, the rule allowed the detection of, for example, 75% of the "difficult" laryngoscopies at a cost of falsely identifying 12% of the "not difficult" patients.

A common cause of mortality and morbidity attributable to anaesthesia is difficult or failed intubation. If those patients in whom intubation proves difficult could be identified in advance, it could be arranged that a senior anaesthetist, properly equipped to deal with the problem, could be present. Many features that are believed to indicate difficulty of intubation have been described [1], but a predictive rule that is useful in routine clinical practice has not been formulated as yet and properly evaluated. Our study was designed to collect data, develop a rule and test it. It was considered that failure to intubate would be a poor criterion for a study, since complete failure is a very rare event. Moreover, the difficulty of intubation is influenced by factors other than anatomical ones, for example, the anaesthetist's skill and persistence. Therefore, the difficulty of laryngoscopy, that is, the amount of larynx seen on laryngoscopy, was studied.

PATIENTS AND METHODS

First, a consecutive series of patients, and some additional patients in whom the trachea was known to be difficult to intubate, were measured to establish descriptive statistics so that a simple rule could be developed. Next, this rule was tested on the data from which it was derived and found to be satisfactory. Finally, the rule was tested prospectively on a further group of patients.

MICHAEL E. WILSON, PH.D., F.F.A.R.C.S; PIERS LESSER†, F.F.A.R.C.S; Department of Anaesthesia, Royal United Hospital, Bath BA1 3NG. DAVID SPIEGELHALTER, PH.D., MRC Biostatistics Unit, 5 Shaftesbury Road, Cambridge CB2 2BW. J. ALEXANDER ROBERTSON*, F.F.A.R.C.S., Royal Naval Hospital, Plymouth. Accepted for Publication: February 27, 1987.

Present addresses:

†Department of Anaesthesia, The General Infirmary at Leeds, Leeds.

*Department of Anaesthesia, Airedale District General Hospital, Steeton, Nr Keighley, Yorks.

Correspondence to M. E. W.

Initial study

Data were collected from a total of 633 patients older than 16 years of age undergoing non-emergency surgery at a District General Hospital and Royal Naval Hospital. During the periods of collection consecutive series of patients, who represented the daily work load of one consultant and three registrars, were recruited to the study. A number of measurements and assessments of features that might predict difficult intubation were made on each patient. These features were:

Inter-incisor gap (IG), measured with the mouth fully open; mandibular length (ML) (tempero–mandibular joint to the tip of the lower incisors); anterior mandibular depth (AM) (base of lower incisor to bottom of chin); posterior mandibular depth (PMD) (depth of mandible measured through the skin immediately posterior to the third molar); neck length (NL) (measured

TABLE I. *Definitions of the different grades of difficult laryngoscopy with the number and percentage of patients found in each grade for the initial series (with and without pressure applied to the larynx) and for the prospective series (with pressure applied)*

Difficulty		Initial series		Prospective series
Grade	Amount of larynx seen	No laryngeal pressure	With laryngeal pressure	With laryngeal pressure
1	Almost all of cords	248 (57.5%)	153 (75.7%)	636 (81.7%)
2	Only half of cords	97 (22.5%)	24 (11.9%)	101 (13.0%)
3	Only arytenoids	46 (10.7%)	13 (6.4%)	29 (3.7%)
4	Only epiglottis	39 (9.0%)	11 (5.4%)	10 (1.3%)
5	Not even epiglottis	1 (0.2%)	1 (0.5%)	2 (0.3%)
Totals		431	202	778

with the head in the intubating position, from tip of mastoid process to the medial end of the clavicle); neck circumference (NC); subluxation (SLux) (maximal forward protrusion of the lower incisors beyond the upper incisors); portion of mandible anterior to a line drawn from the tip of the upper incisors to the cricoid with the head in the intubating position (A); portion of mandible posterior to a line drawn from the tip of the upper incisors to the cricoid with the head in the intubating position (B); distance between sternal notch and cricoid with head in the intubating position (SC); with the head in the intubating position, the mid-point between the suprasternal notch and the tip of the chin is marked. The position of the cricoid is measured in relation to this mark (L); maximum range of head and neck flexion (HFNF) measured from the neutral erect position; maximum range of head and neck extension (HENE) measured from the neutral erect position; maximum range of head and neck movement (HFNF + HENE).

Initially, head and neck movement was measured with a special protractor. However, when it became clear that a simpler procedure was adequate, this was adopted. For this, the patient was asked to extend fully the head and neck while a pencil was stood vertically on the forehead. The orientation of the pencil was adjusted so that it was parallel to a distant window frame. Then, while the pencil was held firmly in position, the head and neck were fully flexed and the pencil was sighted against the horizontal of the window frame to judge if it had moved through 90°.

The severity of receding mandible or long upper incisors (buck teeth) was estimated on a subjective three-point scale (0 = normal; 1 = moderate; 2 = severe). The reliability of these scales was confirmed with a series of photographs. Agreement was good, and the chance that a second colleague would agree with the assessment of a first colleague was at least 70%.

The difficulty of laryngoscopy was assessed at intubation. With the head in the optimum position, a Macintosh laryngoscope bearing a standard blade was inserted into the throat and the proportion of the laryngeal aperture seen was recorded on the 1–5 scale defined in table I.

In the initial study the difficulty of laryngoscopy was estimated in 202 patients with backward pressure applied to the larynx to improve the view and in 431 patients without the pressure. In addition, 38 patients were measured after they had been reported "difficult to laryngoscope" by other anaesthetists.

Statistical analysis

Two groups of patients were formed—a "normal" group containing all 581 cases of grade 1–3 (with or without pressure) and a "difficult" group containing all 50 patients who were grade 4 or 5 despite attempts to improve the view by pressure on the larynx (the 12 patients in the initial series plus the 38 patients reported by colleagues). Features that might predict difficulty with intubation were examined individually for their ability to discriminate between normal and difficult groups. Clinical judgement and pairwise plotting of features were used to determine if particular combinations suggested high risk. A discriminant analysis of original and derived features (e.g. neck circumference/neck length) was carried out to find a predictive index based on a minimal number of measurements. Finally, an

TABLE II. *Continuous measurements compared between "normal" and "difficult" patients. P = Significance of two-sided two-sample t test*

Feature	"Normal" (n = 581) n	Mean	SD	"Difficult" (n = 50) n	Mean	SD	P
Age (yr)	365	44.8	19.8	19	46.3	14.0	0.74
Height (cm)	557	168.7	10.4	50	167.4	10.4	0.41
Weight (kg)	561	68.1	12.8	50	72.1	17.2	< 0.05
IG (cm)	580	4.6	1.0	50	3.8	0.7	< 0.05
ML (cm)	370	10.5	1.0	20	10.2	0.8	< 0.001
AM (cm)	120	3.4	0.7	20	3.7	0.4	0.15
PMD (cm)	370	3.1	0.5	20	2.9	0.4	0.06
NL (cm)	553	14.2	1.8	50	13.7	1.7	0.08
B (cm)	301	5.3	0.9	20	5.8	1.0	< 0.02
A+B (cm)	178	9.0	1.0	13	8.3	1.2	< 0.02
SLux (mm)	461	4.8	3.7	50	−0.5	4.7	< 0.001
NC (cm)	263	36.4	3.5	47	37.9	3.9	< 0.01
HFNF°	136	65.2	14.1	20	51.6	18.5	< 0.001
HENE°	136	61.9	17.0	20	42.1	19.9	< 0.001
HFNF°+HENE°	136	127.1	23.2	20	93.6	35.3	< 0.001
SC	124	58.1	15.0	16	54.5	13.2	0.36
L	89	−7.3	10.0	10	−11.2	13.4	0.27
NC/NL	263	2.5	0.4	47	2.8	0.4	< 0.001
Wt/NL	540	4.8	1.1	50	5.3	1.4	< 0.01
NC/Ht	262	0.55	0.05	47	0.57	0.06	< 0.02

TABLE III. *Final risk factors compared between "normal" and "difficult" patients. P = probability that difference between normal and difficult groups could arise by chance. For weight, the risk levels 0 and 1 were pooled and the difference tested by Fisher's Exact test; all other differences were tested with χ^2*

Risk factor	Risk level	Normal	Difficult	P
Weight	0	533 (95%)	45 (90%)	
	1	27 (5%)	3 (6%)	0.05
	2	1 (0.2%)	2 (4%)	
Head and neck movement	0	297 (91%)	27 (54%)	
	1	21 (6%)	11 (22%)	0.001
	2	8 (3%)	12 (24%)	
Jaw movement	0	457 (92%)	19 (38%)	
	1	36 (7%)	17 (34%)	0.001
	2	2 (0.4%)	14 (28%)	
Receding mandible	0	506 (97%)	29 (58%)	
	1	16 (3%)	16 (32%)	0.001
	2	1 (0.2%)	5 (10%)	
Buck teeth	0	504 (96%)	32 (64%)	
	1	18 (3%)	12 (24%)	0.001
	2	2 (0.4%)	6 (12%)	

extremely simple scoring system was investigated, and its predictive properties simulated.

Prospective trial

Eight anaesthetists (five consultants, one senior registrar and two registrars) agreed to use the scoring system and to assess the difficulty of laryngoscopy on a consecutive series of 778 patients for 2 months. Only one anaesthetist (M.W.) knew the rule that had previously been developed from the scoring system.

RESULTS

Initial study

Table I displays the number of patients who were classified in each grade of difficulty. The incidence of "difficult" laryngoscopy, that is, failure to see even the arytenoids, was 9.3%. However, if the larynx was pushed backwards to improve the view, the incidence of "difficulty" was 5.9%.

Summary statistics for features that might predict difficult laryngoscopy are shown in tables II and III. The data are presented, with the statistical significance of observed differences, for the two previously defined "normal" and "difficult" groups of patients. Many features were measured to see if they might be useful. In general, collection of a feature ceased when it was apparent that it had no value. The exact measurement of head and neck angles with a special apparatus was also discarded in favour of a simpler overall assessment.

TABLE IV. *The three levels of the final five risk factors. (* 5 cm is approximately three fingers' breadth)*

Risk factor	Level	
Weight	0	< 90 kg
	1	90–110 kg
	2	> 110 kg
Head and neck movement	0	Above 90°
	1	About 90° (ie. ± 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

TABLE V. *Effect of varying the criterion for identifying "difficult" patients. True positive (%) = proportion of truly "difficult" patients correctly identified; false positive (%) = proportion of "not-difficult" patients wrongly identified as being at risk*

Risk sum criterion	True positive (%)		False positive (%)	
	Initial	Prospective	Initial	Prospective
⩾6		8	0	0
⩾5		17	0	0.3
⩾4	36	42	0	0.8
⩾3	52	50	1	4.6
⩾2	72	75	6	12.1
⩾1	86	92	24	26.2

Creating a predictive index

Linear discriminant analysis revealed that five risk factors, measured at three levels (table IV) provided significant discrimination; no others improved the prediction.

Two approaches to allocating a risk score to a new patient were investigated. The standard method is to add up the linear discriminant coefficients for the risk factors. The alternative, simpler method is to add up the observed risk levels (table IV), thus creating a "Risk-sum". Although one might expect the more complex discriminant analysis to perform better than the risk-sum method, it was found that the performance of the two methods was virtually indistinguishable when applied to the 344 patients in whom there was a complete set of measurements. The percentage of true positive and false positive predictions for each Risk-sum is shown in table V.

Prospective trial

Of the 778 patients studied, 12 (1.5%) were "difficult" (table I). Each patient was assessed using the simple scoring system and the percentage of true positive and false positive predictions for each Risk-sum are shown in table V.

DISCUSSION

In the initial study, when pressure was applied to the larynx, the incidence of difficult laryngoscopy was 5.9% (95% confidence limits 3.1–10.3%). In the prospective study involving a different group of anaesthetists the incidence was 1.5% (95% confidence limits 0.7–2.3%). This difference probably reflects either the eagerness of the investigators in the initial study to find examples of difficult intubation or a higher proportion of experienced anaesthetists in the second study. Nonetheless, even 1.5% is considerably greater than the 0.05% suggested by Cormack and Lehane [2] from their colleagues' "guestimates". The incidence of difficult *intubation* should be less than this because a tracheal tube may be passed through the cords even if they cannot be seen. On the other hand, sometimes the cords can be seen but there is difficulty passing the tube (e.g. because of an awkward peg tooth). Previous reports of the incidence of difficult intubation are quoted as 1%, 2.3% and 3.6% [1]. In obstetric anaesthesia 2.7% of intubations are reported to be difficult (G. Lyons, personal communication) and the estimates for failed intubation are 0.05–0.3% [3, 4].

The linear discriminant analysis revealed that at least five factors contributed to difficult laryngoscopy, confirming clinical experience that no one factor can predict difficulty. Thus the adverse influence of one factor may be offset by other favourable features. For example, a patient may have severely limited movement of the head and neck and yet intubation may be easy because of great mobility of the lower jaw.

At intubation the neck needs to be flexed, with the head extended [5]; this is a complex manoeuvre and accurate measurement is difficult. X-ray evidence has indicated that limitation of the atlanto-occipital joint may be responsible for difficult intubation [6–9]. Although we have been unable to measure this clinically, we agree that some patients prove unexpectedly difficult to intubate because the head cannot be sufficiently extended on the neck.

Both SLux and IG are individually fairly powerful predictors of difficulty. However, the plot of IG against SLux showed a concentration of "difficult" patients in the bottom left hand corner which indicates an especial risk if both IG and SLux are low. Conversely, a low IG may not be dangerous if SLux is zero or greater; a finding in accord with clinical experience. Thus a predictive factor, "jaw movement", was created which depended on both IG and SLux (table IV).

Various derived ratios such as the neck circumference divided by neck length were designed to try and quantify "short fat neck" which is widely believed to be associated with difficult intubation. Although these derived ratios do predict difficult laryngoscopy (table II), none is simple to use; fortunately, weight itself was found to be satisfactory.

At laryngoscopy the line of sight extends from the laryngeal aperture to the tips of the upper incisors, which limit the view posteriorly. The view will be obstructed if, despite elevation with the laryngoscope blade, the tongue protrudes backwards across the line of sight. The expression "high anterior larynx" is sometimes offered as an explanation for a difficult intubation. Although it is not clear how the position is so judged, the implication is that the lower end of the line of sight is adversely shifted so that the tongue obstructs the view. We have tried many ways to estimate the position of the larynx relative to the tongue and other structures, but without much success. The measurement B, which estimates the length of the mandible behind the line of sight, is helpful but is associated with considerable inter-observer variation. Unusually long upper incisors (buck teeth) may adversely affect the position of the upper end of the line of sight and a receding mandible may indicate that the tongue is positioned more posteriorly than usual, blocking the view. Our analysis confirms the importance of both these factors and not surprisingly receding mandible, buck teeth, inter-incisor gap and subluxation were found to be correlated with one another (especially IG and SLux). Zuck [10] suggested that excessive anterior convexity of the cervical spine may be responsible for an "anteriorly situated larynx". Unfortunately, we have not found a suitable clinical measure.

Neither we (using external estimates of the posterior mandibular depth) nor others (using x-rays) [3, 9] have been able to confirm the predictive value of the posterior mandibular depth [6, 11].

A proposed index should be evaluated on a prospective data set, since the choice of features used in the creation of the index may be influenced by chance factors in the original data. Consequently, the predictive performance of an index might decrease when applied to new data. However, our index performed well on the prospective series of patients (table V). The true positive proportions were remarkably similar and the lower false positive proportions in the initial series may reflect the greater willingness of the authors, who carried out that study, to classify a patient as "difficult".

Although a test that picks up 75% of difficult laryngoscopies (using the Risk-sum criterion of $\geqslant 2$) appears valuable, the number of false alarms may be unacceptable. For example, if a Department anaesthetizes 10000 patients a year and the incidence of difficult laryngoscopy is 1.5%, then one might expect 150 difficult laryngoscopies and 9850 non-difficult laryngoscopies a year. With a false positive rate of 12.1%, 1191 patients a year would be falsely classified as at risk. In other words, skilled assistance would be called unnecessarily 99 times a month so that assistance could be at hand for nine of the 12 difficult laryngoscopies expected each month (the other three being missed); thus only 8% (9/9+99) of alarms would turn out to be necessary. Whether this is helpful or not will depend upon the geography and staffing arrangements of individual hospitals. Increasing the threshold to a criterion of $\geqslant 4$ would reduce the number of false alarms to seven per month at a cost of lowering the number of detected difficult laryngoscopies to five out of 12; thus 42% (5/5+7) of alarms would turn out to be positive.

Ideally, one might wish for a predictive test that picked up nine out of 10 difficult laryngoscopies at a cost of perhaps only one false alarm a week. However, because the incidence of difficult laryngoscopy is so low, this would require a test that was 90% specific and 99.5% sensitive—this is clearly unattainable! An important implication of this study is that, despite exploring many features, it was not possible to identify the difficult patients without the false positive rate increasing unacceptably. Inevitably, patients are at risk because the problem is unexpected. We conclude that it is essential that anaesthetists are aware of the inadequacy of attempts to predict difficult intubation and it is recommended that they practice simulated difficult intubations in the manner described by Cormack and Lehane [2]. Finally, they should be able to fall back upon a clear "difficult intubation drill" or, if necessary, a simple cricothyroidotomy procedure.

ACKNOWLEDGEMENTS

We are grateful to our colleagues at the Royal United Hospital, Bath and the Royal Naval Hospital Plymouth for their enthusiastic support and assistance in collecting data.

REFERENCES

1. Latto IP, Rosen M. *Difficulties in Tracheal Intubation.* Eastbourne: Baillière Tindall, 1985.
2. Cormack RS, Lehane J. Difficult tracheal intubation in obstetrics. *Anaesthesia* 1984; **39**: 1105–1111.
3. Lyons G. Failed intubation. Six years' experience in a teaching maternity unit. *Anaesthesia* 1985; **40**: 759–762.
4. Lyons G, MacDonald R. Difficult intubation in obstetrics. *Anaesthesia* 1985; **40**: 1016.
5. Bannister FB, Macbeth RG. Direct laryngoscopy and tracheal intubation. *Lancet* 1944; **2**: 651–654.
6. White A, Kander PL. Anatomical factors in difficult direct laryngoscopy. *British Journal of Anaesthesia* 1975, **47**: 468–474.
7. Nichol HC, Zuck D. Difficult laryngoscopy—the 'anterior' larynx and the atlanto-occipital gap. *British Journal of Anaesthesia* 1983; **55**: 141–144.
8. Skues M. Anatomical factors complicating direct laryngoscopy. *Anaesthesia Points West* 1984; **18**: 31–33.
9. Levack ID, Masson AHB. Difficult tracheal intubation. *Anaesthesia* 1986; **41**: 332.
10. Zuck D. Difficult tracheal intubation. *Anaesthesia* 1985; **40**: 1016–1017.
11. Evans R, Cormack RS. Difficult intubation. *Anaesthesia Points West* 1984; **17**: 79–80.

Note added in proof. Since the manuscript was accepted for publication Samsoon and Young (*Anaesthesia* 1987; **42**: 487–490) have advocated the Malampatti test (*Canadian Anaesthestists' Society Journal* 1985; **32**: 429–434). They examined 13 patients known to be difficult subjects for intubation and, with the exception of one patient with tracheal stenosis, found that in every patient the uvula and soft palate were obscured by the tongue. Unfortunately, they did not report the incidence of this sign in patients who were not difficult to intubate. However, it is possible to evaluate the performance of the test from data presented in table II of the Malampatti paper. If Class 3 (only soft palate could be visualized) is used as the criterion to predict difficult laryngoscopy, then it can be calculated that the false positive rate was 5% for a true positive rate of only 55%.