*British Journal of Anaesthesia* **107** (5): 659–67 (2011)
Advance Access publication 26 September 2011 · doi:10.1093/bja/aer292

BJA

REVIEW ARTICLES

# Poor prognostic value of the modified Mallampati score: a meta-analysis involving 177 088 patients

L. H. Lundstrøm[1,2]*, M. Vester-Andersen[1], A. M. Møller[1], S. Charuluxananan[3], J. L'Hermite[4] and J. Wetterslev[2], the Danish Anaesthesia Database

[1] Department of Anaesthesia and Intensive Care, Herlev Hospital, Copenhagen University Hospital, Herlev Ringvej 75, 2730 Herlev, Denmark
[2] Copenhagen Trial Unit, Centre of Clinical Intervention Research, Copenhagen University Hospital, Rigshospitalet, Denmark
[3] Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand
[4] Division Anesthésie Réanimation Douleur Urgences Groupe Hospitalo-Universitaire Caremeau, Nîmes, France

* Corresponding author. E-mail: lars_hyldborg@hotmail.com

## Editor's key points

- The Mallampati score is used by some clinicians to predict difficult intubation.
- Meta-analysis of 55 studies involving a total of 177 088 patients (two studies account for 149 096 patients).
- Only 35% of patients with a difficult intubation were identified as Mallampati III or IV.
- The modified Mallampati is inadequate as a stand-alone test of a difficult laryngoscopy or intubation.

**Summary.** The modified Mallampati score is used to predict difficult tracheal intubation. We have conducted a meta-analysis of published studies to evaluate the Mallampati score as a prognostic test. A total of 55 studies involving 177 088 patients were included after comprehensive electronic and manual searches. The pooled estimates from the meta-analyses were calculated based on a random-effects model and a summary receiver operating curve. Meta-regression analyses were performed to explore sources of possible heterogeneity between the studies. The summary receiver operating curve demonstrated an area under the curve of 0.75. The pooled odds ratio for a difficult intubation with a modified Mallampati score of III or IV was 5.89 [95% confidence interval (CI), 4.74–7.32]. The pooled estimates of the specificity and sensitivity were 0.91 (CI, 0.91–0.91) and 0.35 (CI, 0.34–0.36), respectively. The pooled positive and negative likelihood ratios were 4.13 (CI, 3.60–4.66) and 0.70 (CI, 0.65–0.75), respectively. The meta-analyses had statistical and clinical heterogeneity ranging from 87.2% to 99.4%. Meta-regression analyses did not identify any significant explanation of the heterogeneity. We conclude that the prognostic value of the modified Mallampati score was worse than that estimated by previous meta-analyses. Our assessment shows that the modified Mallampati score is inadequate as a stand-alone test of a difficult laryngoscopy or tracheal intubation, but it may well be a part of a multivariate model for the prediction of a difficult tracheal intubation.

**Keywords:** anaesthetic techniques, laryngoscopy; complications, intubation tracheal; intubation

Difficult tracheal intubation may be a major cause of severe perioperative morbidity and mortality related to anaesthesia.[1–4] Several studies have focused on one or more patient-related factors which may identify those at risk for difficult tracheal intubation. Among these, the Mallampati score[5] and the modified Mallampati score[6] have been evaluated as risk factors for difficult tracheal intubation. Previous meta-analyses have evaluated the accuracy of the original[7 8] and the modified Mallampati tests[8] to predict a difficult intubation or difficult laryngoscopy. Since then, the modified Mallampati score has been evaluated in several new studies, including two large cohort studies[9 10] with more than 149 000 patients. These large cohort studies represent everyday practice, in a clinical environment with diverse settings, where the preoperative airway evaluation and airway management were performed by multiple anaesthesiologists. Hereby, these studies differ from most of the other studies which had very few evaluators adhering strictly to protocol procedures of both evaluation and settings for the intubation. The aim of this review is to assess the accuracy of the modified Mallampati score as a prognostic test for a difficult tracheal intubation using a meta-analysis and including studies published since the previous meta-analysis.[8]

## Search strategy

We conducted an electronic search covering May 1987, the time since introduction of the modified Mallampati score,[6] through to December 2009. The Cochrane Library, MEDLINE, Science Citation Index, and EMBASE were used as sources for the identification of studies. The search was conducted with the following search string:[11] (sensitivity OR specificity OR screening OR false positive OR false negative OR predictive value of tests OR reference values OR roc analyses OR roc area OR roc characteristics OR roc curve), (intubation OR

endotracheal intubation OR intratracheal intubation OR laryngoscopy OR difficult laryngoscopy OR difficult intubation OR Cormack Lehane), and Mallampati. To determine the studies to be assessed further, two authors (L.H.L. and J.W.) independently scanned the abstract, title, or both sections of every record retrieved. All potentially relevant articles were investigated as full text. In addition, we checked the references from included studies. Any relevant missing information on the study was sought from the original author(s) of the article, if required. The inclusion criteria were: (i) The modified Mallampati score was used. (ii) Studies included prospectively collected data. (iii) The study included adult patients. (iv) Direct laryngoscopy with a standard laryngoscope was performed. (v) The absolute number of true positive, false negative, true negative, and false negative could be extracted from the article, from previous meta-analyses,[7 8] or by contacting the author(s). (vi) The study was reported in English.

## Data extraction

For studies that fulfilled the inclusion criteria, two authors (L.H.L. and M.V.-A.) independently abstracted relevant information and characteristics using standard data extraction templates. When differences in opinion existed, they were resolved by a third party (J.W.). The following information was extracted.

The outcome measure, difficult tracheal intubation, or difficult laryngoscopy defined by Cormack and Lehane score of III and IV[12] or a modified Cormack and Lehane score of IIb, III, and IV.[13] As there is no international consensus of defining an intubation score, the definitions of a difficult tracheal intubation presented in the individual articles were accepted. However, if the authors defined a difficult laryngoscopy using a Cormack and Lehane score as a difficult intubation, we included and reported the Cormack and Lehane score as an outcome measure in our assessment. A difficult laryngoscopy is a surrogate outcome measure for a difficult tracheal intubation. Therefore, if a study both reported an intubation score and the Cormack and Lehane score based on the same population in the same assessment, only the intubation score was extracted for our assessment.

The modified Mallampati score was defined by Samsoon and Young.[6] The view was graded as follows: class I, soft palate, fauces, uvula, and pillars visible; class II, soft palate, fauces, and uvula visible; class III, soft palate and base of the uvula visible; class IV, soft palate not visible at all. The patients were placed in a sitting position with the head in a neutral potion and the assessment was performed without phonation. A modified Mallampati score of III or IV was considered a risk factor for difficult laryngoscopy/intubation. When a single study has evaluated various versions of the modified Mallampati score, the results of the score performed similar to the one reported by Samsoon and Young[6] were extracted for the meta-analyses.

If possible, the following other data were extracted:

The settings of the Mallampati score by retrieving the position of the head and body, and if the patients phonated during the evaluation.

The number of anaesthetists performing the preoperative airway assessments and the number of anaesthetists involved in tracheal intubations.

It was noted, if the modified Mallampati score was blinded for the anaesthetists performing the airway management.

The participant sampling and inclusion and exclusion criteria of patient population were retrieved.

How the patients were recruited.

## Quality assessment and risk of bias

We applied QUADAS, a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews,[14] as a measure of the quality of the included studies. The quality assessment was based on the following criteria:

(1) Was the modified Mallampati score blinded for the anaesthetist performing the airway management?
(2) Was the conduct of the modified Mallampati test clearly described?
(3) Selection of study population: were the inclusion and exclusion criteria described?
(4) Recruitment of study population described (e.g. consecutive, randomly, case–control)?

Studies fulfilling all four criteria were classified as studies with low risk of bias, if three criteria were fulfilled, they were categorized as medium risk of bias studies. Otherwise they were classified as studies with high risk of bias.

Our meta-analysis included studies reporting difficult laryngoscopy or difficult tracheal intubation. Further, subgroup meta-analyses of difficult laryngoscopy and difficult tracheal intubation were presented separately. However, if a study both reported an intubation score and the Cormack and Lehane score, only the intubation score was extracted for the subgroup analyses. Finally, a subgroup meta-analysis of studies with low and medium risk of bias was presented.

## Statistical analysis

The modified Mallampati score from the pooled estimates in the meta-analyses was described with: sensitivity, specificity, positive likelihood ratio, and negative likelihood ratio. The pooled estimates were presented with 95% confidence interval (CI). If a study was reported with a 0 value in any outcome, 0.5 was added as a continuity correction to all values in the study before performing the meta-analysis. Before conducting a meta-analysis, the degrees of heterogeneity displayed by the $I^2$ of all estimates were calculated.[15] Because of expected clinical diversity and high $I^2$ values, the 'random-effects model'[16] was used incorporating a moment-based between-study variance when calculating the pooled estimates. The calculation of the Spearman correlation coefficient of logit (true-positive rates) *vs* logit (false-positive rates) demonstrated that the sensitivity and specificity were associated across the studies. Therefore, a summary receiver operator characteristics (sROC) curve[17] was conducted. The area under the sROC curve was used as a measure for the

**Table 1** Characteristics of the studies evaluating the modified Mallampati score. Blinding, blinding of the Mallampati score; DL, difficult laryngoscopy; DTI, difficult tracheal intubation; Recruitment, if the recruitment of study population was described; Selection, if the selection of study population was described; Setting, if the setting of the test was described

| | Outcome measure | No. of patients | Blinding | Setting | Selection | Recruitment | Risk of bias |
|---|---|---|---|---|---|---|---|
| Adnet F[20] | DTI | 1171 | | | Yes | Yes | High |
| Ali Z[21] | DL | 66 | Yes | | Yes | Yes | Medium |
| Allahyar E[22] | DL | 203 | Yes | Yes | Yes | Yes | Low |
| Arne J[23] | DTI | 1200 | | | Yes | Yes | High |
| Ayuso MA[24] | DTI | 181 | Yes | | Yes | Yes | Medium |
| Bouaggad A[25] | DTI | 320 | | | Yes | Yes | High |
| Brodsky JB[26] | DTI | 100 | | | Yes | Yes | High |
| Butler P[27] | DL | 220 | | | Yes | | High |
| Cattano D[28] | DTI | 1956 | | | Yes | Yes | High |
| Charuluxananan S[9] | DL | 57 005 | | Yes | Yes | Yes | Medium |
| Constantikes J[29] | DL | 30 | | Yes | Yes | | High |
| Domi R[30] | DL | 426 | Yes | | Yes | | High |
| Eberhart LHJ[31] | DL | 1107 | Yes | | Yes | Yes | Medium |
| Ezri T[32] | DL | 764 | | Yes | Yes | Yes | Medium |
| Ezri T[33] | DL | 1472 | | | Yes | Yes | High |
| Ezri T[34] | DL | 50 | | | Yes | | High |
| Ezri T[35] | DL | 644 | | | Yes | Yes | High |
| Frerk CM[36] | DL | 244 | | Yes | | | High |
| Gercek A[37] | DTI | 500 | Yes | | Yes | | High |
| Gupta S[38] | DTI | 372 | Yes | | | | High |
| Hester CE[39] | DL | 50 | Yes | | Yes | | High |
| Honarman A[40] | DL | 400 | Yes | | Yes | Yes | Medium |
| Huh J[41] | DL | 213 | Yes | | Yes | Yes | Medium |
| Iohom G[42] | DL | 212 | Yes | | Yes | | High |
| Ita CE[43] | DL | 57 | | | Yes | | High |
| Juvin P[44] | DTI | 245 | | | Yes | Yes | High |
| Kamalipour H[45] | DL | 100 | Yes | | Yes | Yes | Medium |
| Kaul TK[46] | DL | 500 | | Yes | Yes | | High |
| Khan ZH[47] | DL | 300 | Yes | | Yes | Yes | Medium |
| Koh LK[48] | DL | 605 | | | Yes | Yes | High |
| Komatsu R[49] | DL | 64 | | | Yes | | High |
| Krishna HM[50] | DL | 200 | Yes | | Yes | Yes | Medium |
| Krobbuaban B[51] | DL | 550 | Yes | | Yes | Yes | Medium |
| L'Hermite J[52] | DTI | 1023 | | Yes | Yes | Yes | Medium |
| Lundstrom LH[10] | DTI | 92 091 | | Yes | Yes | Yes | Medium |
| Mashour GA[53] | DL | 346 | | Yes | Yes | Yes | Medium |
| Mashour GA[54] | DTI | 1133 | | | Yes | Yes | High |
| Mashour GA[55] | DL | 60 | | Yes | | | High |
| Merah NA[56] | DL | 80 | | Yes | Yes | Yes | Medium |
| Merah NA[57] | DL | 380 | | | Yes | Yes | High |
| Noorizad S[58] | DL | 379 | | | | | High |
| Rocke DA[59] | DTI | 1500 | Yes | Yes | | | High |
| Samra SK[60] | DL | 566 | | | Yes | | High |
| Savva D[61] | DTI | 350 | Yes | Yes | Yes | Yes | Low |
| Schmitt H[62] | DL | 128 | | Yes | Yes | Yes | Medium |
| Siddiqi R[63] | DL | 338 | | | Yes | | High |
| Singh R[64] | DL | 300 | | | Yes | | High |
| Topcu I[65] | DTI | 208 | | | Yes | Yes | High |
| Torres K[66] | DTI | 96 | | | | Yes | High |

*Continued*

**Table 1** *Continued*

| | Outcome measure | No. of patients | Blinding | Setting | Selection | Recruitment | Risk of bias |
|---|---|---|---|---|---|---|---|
| Tuzuner-Oncul A[67] | DTI | 208 | Yes | Yes | | | High |
| Vani V[68] | DL | 50 | | Yes | Yes | | High |
| Wong SH[69] | DL | 411 | Yes | | Yes | Yes | Medium |
| Yamamoto K[70] | DL | 3680 | Yes | | Yes | Yes | Medium |
| Yeo SW[71] | DL | 560 | | | Yes | Yes | High |
| Yildiz TS[72] | DTI | 1674 | Yes | | Yes | | High |

description of diagnostic accuracy of the Mallampati test. To ensure precise pooled estimates, the pooled sensitivity was derived from the sROC curve using corresponding pooled specificity.[17 18] Subsequently, the pooled sensitivity and specificity were used to calculate the positive and negative likelihood ratios. To explore sources of heterogeneity in the studies, meta-regression analyses using the Moses–Shapiro–Littenberg method[17] was performed. All covariates associated with a *P*-value of $<0.10$ in the univariate analyses were included in a subsequent multivariate meta-regression analysis. In the multivariate analysis, a *P*-value of $<0.05$ was considered significant. Possible bias was assessed by the method described by Egger and colleagues.[19] SPSS v. 17.0 (SPSS Inc., Chicago, IL, USA), Comprehensive Meta-Analysis version 2.2.048 (Borenstein M, Hedges L, Higgins J, Rothstein H. Biostat, Englewood NJ, USA), and MetaDiSc version 1.4 (Zamora J, Muriels A, Abraira V, Madrid, Spain) were used for the analyses.

## Results

A total of 55 studies including 177 088 patients[9 10 20–72] met the inclusion criteria for the meta-analysis. In three of the included studies, the necessary data were extracted by contacting the authors[9 52 66] (Table 1). A detailed description of the studies is presented in Supplementary Appendix S1.

The prognostic performance of the modified Mallampati score in the individual studies (Table 2) showed a statistically significant association between logit (true-positive rate) and logit (false-positive rate) (Spearman's correlation coefficient: 0.36, $P=0.007$). We therefore constructed a symmetric sROC curve (Fig. 1). The area under the curve (AUC) was 0.75, which categorized the diagnostic test as good.[73] The pooled estimates of the total cohort and subgroup populations are presented in Table 3. A pooled diagnostic odds ratio of 5.89 (4.74–7.32, 95% CI) demonstrates an almost six-fold increased risk of a difficult tracheal intubation for patients with a modified Mallampati class III or IV compared with those with class I or II. Only 35% (34–36%, 95% CI) of the patients who had difficult tracheal intubation were correctly identified with a modified Mallampati class III or IV. Among the patients who underwent a tracheal intubation without difficulties, 91% (91–91%, 95% CI) were correctly predicted to be easy. For the patients with a modified Mallampati class III or IV, the ratio between the number of patients with a difficult and an easy tracheal intubations was 4.13 (3.60–4.66, 95% CI), indicating how the odds of a difficult tracheal intubation are increased if the modified Mallampati test indicates difficulties. For the patients with a modified Mallampati class I or II, the ratio between the number of patients with a difficult and an easy tracheal intubation was 0.70 (0.65–0.75, 95% CI), indicating how the odds of a difficult tracheal intubation are decreased if the modified Mallampati score does not predict difficulties.

The prevalence of a difficult intubation or a difficult laryngoscopy varied between 0.7% and 31.3%. The sensitivity and specificity ranged from 0.00 to 1.00 and 0.44 to 1.00, respectively. The positive and negative likelihood ratio varied from 0.46 to 48.38 and from 0.13 to 1.18, respectively. The diagnostic odds ratio ranged from 0.38 to 161.00. There was a great heterogeneity among the studies as $I^2$ in the meta-analyses of all the pooled estimates ranged from 87.2% to 99.4%. To explore possible causes of heterogeneity across the studies, we performed univariate meta-regression analyses based on various characteristics of the individual studies (Supplementary Appendix S1). The meta-regression analyses did not identify any significant explanation of the heterogeneity. Our evaluation of possible bias evaluated by the Egger and colleagues'[19] regression intercept demonstrated no evidence of small study bias ($t=0.71$, $P=0.48$).

## Discussion

In this meta-analysis, the pooled frequency of difficult tracheal intubation was 6.8%, which exceeds the 5.8% found in a previous analysis.[7] The pooled estimates demonstrated that only 35% of the patients, who underwent tracheal intubation with difficulties, were correctly identified with a modified Mallampati test. The pooled positive likelihood ratio was 4.1. A clinical test is considered to be diagnostically accurate if it has a positive likelihood ratio of $>10$.[18] The results of the sROC curve show that the accuracy of the test was only just categorized as 'good',[73] as the AUC was only marginally $>0.75$. Thus, as a stand-alone test, the meta-analysis demonstrated that the modified Mallampati score was an inadequate predictor of a difficult laryngoscopy or tracheal intubation. Hereby, we concur with previous analyses,[7 8] although our assessments are not directly comparable with one of these as it did not distinguish between the original and the modified Mallampati score.[7] However, our meta-analysis of the modified Mallampati test differs substantially from the results reported by Lee and colleagues,[8] who

**Table 2** Individual studies' evaluating the diagnostic performance of the modified Mallampati score. DL, difficult laryngoscopy; DOR, diagnostic odds ratio; DTI, difficult tracheal intubation; *n*, number of patients; NegLR, negative likelihood ratio; PosLR, positive likelihood ratio; Sens., sensitivity; Spec., specificity; Weight, weight of each study in a pooled diagnostic odds ratio using a random-effect model

| | True positive (*n*) | False positive (*n*) | False negative (*n*) | True negative (*n*) | Prevalence DTI/DL (%) | DOR | Weight (%) | Sens. | Spec. | PosLR | NegLR |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Adnet F[20] | 24 | 54 | 66 | 1027 | 7.7 | 6.92 | 2.6 | 0.27 | 0.95 | 5.34 | 0.77 |
| Ali Z[21] | 5 | 13 | 6 | 42 | 16.7 | 2.69 | 1.4 | 0.46 | 0.76 | 1.92 | 0.71 |
| Allahyar E[22] | 11 | 42 | 26 | 124 | 18.2 | 1.25 | 2.2 | 0.30 | 0.75 | 1.18 | 0.94 |
| Arne J[23] | 39 | 168 | 11 | 982 | 4.2 | 20.72 | 2.4 | 0.78 | 0.85 | 5.34 | 0.26 |
| Ayuso MA[24] | 29 | 26 | 25 | 101 | 29.8 | 4.51 | 2.4 | 0.54 | 0.80 | 2.62 | 0.58 |
| Bouaggad A[25] | 7 | 8 | 10 | 295 | 5.3 | 25.81 | 1.6 | 0.41 | 0.97 | 15.60 | 0.60 |
| Brodsky JB[26] | 7 | 26 | 5 | 62 | 12.0 | 3.34 | 1.6 | 0.58 | 0.71 | 1.97 | 0.59 |
| Butler P[27] | 10 | 38 | 8 | 164 | 8.2 | 5.40 | 1.9 | 0.56 | 0.81 | 2.95 | 0.55 |
| Cattano D[28] | 28 | 169 | 28 | 1731 | 2.9 | 10.24 | 2.6 | 0.50 | 0.91 | 5.62 | 0.55 |
| Charuluxananan S[9] | 703 | 3122 | 1398 | 51782 | 3.7 | 8.34 | 3.1 | 0.34 | 0.94 | 5.88 | 0.71 |
| Constantikes J[29] | 2 | 15 | 0 | 13 | 6.7 | 4.36 | 0.4 | 1.00 | 0.46 | 1.56 | 0.36 |
| Domi R[30] | 30 | 10 | 38 | 348 | 16.0 | 27.47 | 2.2 | 0.44 | 0.97 | 15.79 | 0.58 |
| Eberhart LHJ[31] | 92 | 381 | 39 | 595 | 11.8 | 3.68 | 2.9 | 0.70 | 0.61 | 1.80 | 0.49 |
| Ezri T[32] | 68 | 196 | 13 | 487 | 10.6 | 13.00 | 2.5 | 0.84 | 0.71 | 2.93 | 0.23 |
| Ezri T[33] | 116 | 354 | 36 | 966 | 10.3 | 8.79 | 2.9 | 0.76 | 0.73 | 2.85 | 0.32 |
| Ezri T[34] | 3 | 6 | 6 | 35 | 18.0 | 2.92 | 1.1 | 0.33 | 0.85 | 2.28 | 0.78 |
| Ezri T[35] | 11 | 208 | 32 | 393 | 6.7 | 0.65 | 2.4 | 0.26 | 0.65 | 0.74 | 1.14 |
| Frerk CM[36] | 9 | 43 | 2 | 190 | 4.5 | 19.88 | 1.2 | 0.82 | 0.82 | 4.43 | 0.22 |
| Gercek A[37] | 2 | 15 | 31 | 452 | 6.6 | 1.94 | 1.2 | 0.06 | 0.97 | 1.89 | 0.97 |
| Gupta S[38] | 15 | 8 | 9 | 340 | 6.5 | 70.83 | 1.8 | 0.63 | 0.98 | 27.19 | 0.38 |
| Hester CE[39] | 1 | 10 | 8 | 31 | 18.0 | 0.39 | 0.7 | 0.11 | 0.76 | 0.46 | 1.18 |
| Honarman A[40] | 22 | 12 | 13 | 353 | 8.8 | 49.78 | 2.1 | 0.63 | 0.97 | 19.12 | 0.38 |
| Huh J[41] | 3 | 12 | 23 | 175 | 12.2 | 1.90 | 1.4 | 0.12 | 0.94 | 1.80 | 0.95 |
| Iohom G[42] | 8 | 22 | 12 | 170 | 9.4 | 5.15 | 1.9 | 0.40 | 0.89 | 3.49 | 0.68 |
| Ita CE[43] | 0 | 4 | 1 | 52 | 1.8 | 3.89 | 0.4 | 0.00 | 0.93 | 3.17 | 0.81 |
| Juvin P[44] | 2 | 75 | 3 | 165 | 2.0 | 1.47 | 1.0 | 0.40 | 0.69 | 1.28 | 0.87 |
| Kamalipour H[45] | 4 | 0 | 11 | 85 | 15.0 | 66.91 | 0.5 | 0.27 | 1.00 | 48.38 | 0.72 |
| Kaul TK[46] | 31 | 24 | 11 | 434 | 8.4 | 50.96 | 2.2 | 0.74 | 0.95 | 14.09 | 0.28 |
| Khan ZH[47] | 14 | 94 | 3 | 189 | 5.7 | 9.38 | 1.5 | 0.82 | 0.67 | 2.48 | 0.26 |
| Koh LK[48] | 14 | 45 | 17 | 529 | 5.1 | 9.68 | 2.3 | 0.45 | 0.92 | 5.76 | 0.60 |
| Komatsu R[49] | 6 | 16 | 14 | 28 | 31.3 | 0.75 | 1.7 | 0.30 | 0.64 | 0.83 | 1.10 |
| Krishna HM[50] | 13 | 42 | 4 | 141 | 8.5 | 10.91 | 1.6 | 0.77 | 0.77 | 3.33 | 0.31 |
| Krobbuaban B[51] | 48 | 193 | 21 | 288 | 12.5 | 3.41 | 2.6 | 0.70 | 0.60 | 1.73 | 0.51 |
| L'Hermite J[52] | 24 | 104 | 36 | 859 | 5.9 | 5.51 | 2.6 | 0.40 | 0.89 | 3.70 | 0.67 |
| Lundstrom LH[10] | 1042 | 5898 | 3620 | 81531 | 5.1 | 3.98 | 3.2 | 0.22 | 0.93 | 3.31 | 0.83 |
| Mashour GA[53] | 7 | 79 | 10 | 250 | 4.9 | 2.22 | 1.9 | 0.41 | 0.76 | 1.72 | 0.77 |
| Mashour GA[54] | 55 | 153 | 84 | 841 | 12.3 | 3.60 | 2.9 | 0.40 | 0.85 | 2.57 | 0.71 |
| Mashour GA[55] | 5 | 16 | 1 | 38 | 10.0 | 11.88 | 0.7 | 0.83 | 0.70 | 2.81 | 0.24 |
| Merah NA[56] | 7 | 3 | 1 | 69 | 10.0 | 161.00 | 0.7 | 0.88 | 0.96 | 21.00 | 0.13 |
| Merah NA[57] | 8 | 7 | 5 | 360 | 3.4 | 82.29 | 1.4 | 0.62 | 0.98 | 32.26 | 0.39 |
| Noorizad S[58] | 11 | 81 | 18 | 269 | 7.7 | 2.03 | 2.2 | 0.38 | 0.77 | 1.64 | 0.81 |
| Rocke DA[59] | 19 | 378 | 13 | 1090 | 2.1 | 4.21 | 2.4 | 0.59 | 0.74 | 2.31 | 0.55 |
| Samra SK[60] | 24 | 69 | 24 | 449 | 8.5 | 6.51 | 2.5 | 0.50 | 0.87 | 3.75 | 0.58 |
| Savva D[61] | 11 | 113 | 6 | 220 | 4.9 | 3.57 | 1.9 | 0.65 | 0.66 | 1.91 | 0.53 |
| Schmitt H[62] | 25 | 53 | 8 | 42 | 25.8 | 2.48 | 2.1 | 0.76 | 0.44 | 1.36 | 0.55 |
| Siddiqi R[63] | 3 | 53 | 4 | 278 | 2.1 | 3.93 | 1.2 | 0.43 | 0.84 | 2.68 | 0.68 |
| Singh R[64] | 0 | 46 | 2 | 252 | 0.7 | 1.09 | 0.4 | 0.00 | 0.85 | 1.07 | 0.99 |
| Topcu I[65] | 5 | 11 | 7 | 185 | 5.8 | 12.01 | 1.5 | 0.42 | 0.94 | 7.42 | 0.62 |

*Continued*

**Table 2** *Continued*

| | True positive (*n*) | False positive (*n*) | False negative (*n*) | True negative (*n*) | Prevalence DTI/DL (%) | DOR | Weight (%) | Sens. | Spec. | PosLR | NegLR |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Torres K[66] | 8 | 15 | 12 | 61 | 20.8 | 2.71 | 1.8 | 0.40 | 0.80 | 2.03 | 0.75 |
| Tuzuner-Oncul A[67] | 19 | 30 | 13 | 146 | 15.4 | 7.11 | 2.2 | 0.59 | 0.83 | 3.48 | 0.49 |
| Vani V[68] | 1 | 4 | 7 | 38 | 16.0 | 1.36 | 0.7 | 0.13 | 0.91 | 1.31 | 0.97 |
| Wong SH[69] | 6 | 150 | 1 | 254 | 1.7 | 10.16 | 0.8 | 0.86 | 0.63 | 2.31 | 0.23 |
| Yamamoto K[70] | 38 | 1723 | 18 | 1901 | 1.5 | 2.33 | 2.6 | 0.68 | 0.53 | 1.43 | 0.61 |
| Yeo SW[71] | 3 | 25 | 8 | 524 | 2.0 | 7.86 | 1.4 | 0.27 | 0.95 | 5.99 | 0.76 |
| Yildiz TS[72] | 28 | 172 | 52 | 1422 | 4.8 | 4.45 | 2.7 | 0.35 | 0.89 | 3.24 | 0.73 |

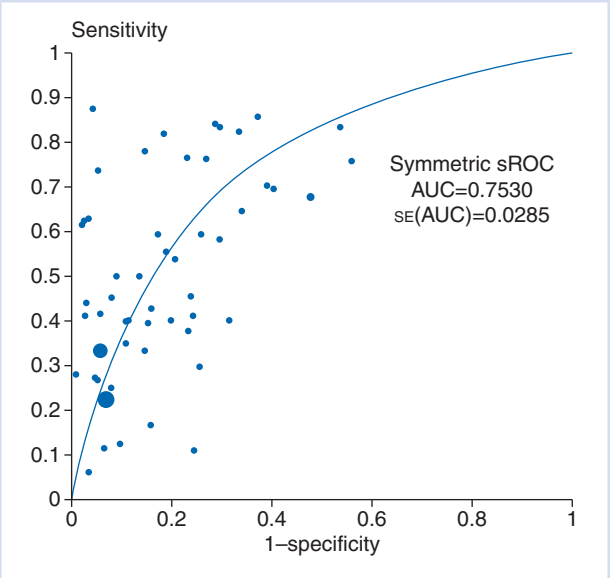


**Fig 1** The sROC curve of 55 studies evaluating the modified Mallampati score as a predictor for difficult tracheal intubation or difficult laryngoscopy. AUC, area under the curve; SE, standard error.

described pooled estimates of sensitivity significantly greater than our findings. Between 55% and 76% of the patients, who underwent tracheal intubation with difficulties, were correctly identified in advance using a modified Mallampati test for their assessment, while the equivalent estimate decreases to 35% in our assessment. The estimated pooled diagnostic odds ratio in our assessment was 5.9. The corresponding diagnostic odds ratios previously reported[8] for increased risk of a difficult tracheal intubation for patients with a modified Mallampati class III or IV compared with those with class I or II ranged was higher at 6.5–10.4. Likewise, the AUC reported[8] ranged from 0.78 to 0.84, while in our assessment, it was only 0.75.

There are several possible explanations for these differences. In some of the studies, the outcome measure was defined as a difficult tracheal intubation, although the real outcome measure was the visibility of anatomical characteristics of the original Cormack and Lehane score.[12] In contrast to Lee and colleagues' study,[8] we used the Cormack and Lehane score as outcome measure in these studies. Further, our meta-analysis combined both difficult laryngoscopy and difficult tracheal intubation as one outcome measure while theirs did not. If a study both reported an intubation score and the Cormack and Lehane score based on the same population in the same assessment, only the intubation score was extracted for our assessment. We did that to avoid a wrong sampling error.[74] One aim of Lee and colleagues' study[8] was to distinguish between a difficult laryngoscopy and a difficult tracheal intubation, and they evaluated the outcome measures separately. We felt that it was correct to retrieve data on both outcome measures from the same patient population. Therefore, this difference may impact the comparison of our assessment with the previous one. However, the major reason for the differences may be the increased number of studies and patients included in our updated meta-analyses. Thus, the number of studies increased from 28 to 55 and the number of patients increased nearly 10-fold from 17 902 to 177 088.

The meta-analyses all had a high statistical heterogeneity. Additionally, the pooled estimates from our meta-analyses may be influenced by a high degree of clinical diversity. Thus, who and how the test was performed, and the type of patient population evaluated, varied considerably between the individual studies. For example, unblinded studies and case–control studies may tend to overestimate the diagnostic accuracy.[75] However, our exploration of the statistical heterogeneity with meta-regression analyses of numerous factors failed to identify the reasons for the statistical heterogeneity.

The number of patients evaluated in each study varied considerably. Two studies[9 10] evaluated a total of 149 096 patients which accounts for 84%. The accumulated weight of the two studies in the random-effects model evaluating the diagnostic odds ratio was only 6.3%. This discrepancy suggests that it is reasonable to emphasize the between-study variance when pooling the estimates as it is done in the random-effects model. The random-effect model used for pooling diagnostic studies may have important shortcomings when large cohort studies comprising more than 80% of the included patients are inappropriately down-weighted.[76 77] The ultimate goal of a prognostic test is to guide clinicians in everyday practice, in

**Table 3** Pooled estimates of the total cohort and subgroup populations. Prev., prevalence

| | No. of studies | No. of patients | AUC (SE) | Pooled estimates | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | Prev. DTI/DL (%) | Diagnostic odds ratio | Specificity | Sensitivity | Positive likelihood ratio | Negative likelihood ratio |
| Total population | 55 | 177 088 | 0.75 (0.03) | 7.6 (6.6–8.8) | 5.89 (4.74–7.32) | 0.91 (0.91–0.91) | 0.35 (0.34–0.36) | 4.13 (3.60–4.66) | 0.70 (0.65–0.75) |
| Subgroup populations | | | | | | | | | |
| DTI | 20 | 104 784 | 0.77 (0.04) | 6.8 (5.3–8.7) | 6.33 (4.71–8,49) | 0.93 (0.92–0.93) | 0.32 (0.31–0.33) | 4.61 (3.91–4.30) | 0.73 (0.66–0.80) |
| DL | 35 | 72 304 | 0.75 (0.04) | 8.0 (6.1–10.3) | 5.58 (3.92–7.93) | 0.90 (0.89–0.90) | 0.38 (0.36–0.40) | 3.83 (3.06–4.60) | 0.69 (0.61–0.77) |
| Low/medium bias | 20 | 159 198 | 0.75 (0.03) | 8.5 (6.8–10.5) | 5.12 (3.74–7.00) | 0.92 (0.92–0.92) | 0.31 (0.30–0.32) | 3.85 (3.13–4.56) | 0.75 (68.2–0.82) |

a clinical environment with diverse settings. Studies conducted with few evaluators adhering strictly to a protocol may exaggerate the prognostic value. Therefore, large database studies may convey a more realistic picture of the prognostic value achieved by the Mallampati test. In contrast, the smaller studies adhering strictly to protocols may describe what is ultimately possible if education and training are optimized. Furthermore, it seems appropriate to emphasize that due to the apparent high precision of some of the study estimates of diagnostic accuracy,[9 10] and these estimates discrepancy with those of other studies, the statistical heterogeneity measured by $I^2$ may be exaggerated.[78]

Our assessment only deals with studies predicting difficult tracheal intubation with a standard laryngoscope. However, in a clinical context, the impact of these studies and our meta-analysis may have changed, because of the current introduction of videolaryngoscopes into anaesthetic practice. The clinical use of videolaryngoscopes may change the accuracy of predictors of difficult tracheal intubation and require a different definition of difficult tracheal intubation. Thus, predictors of a difficult tracheal intubation such as the modified Mallampati score may require re-evaluation in the future.

Our assessment demonstrated that the prognostic value of the modified Mallampati score was poorer than that estimated by previous meta-analyses. The modified Mallampati score is inadequate as a stand-alone test of a difficult laryngoscopy or tracheal intubation, but it may well have a role as a part of a multivariate model for the prediction of a difficult tracheal intubation using a standard laryngoscope.[23 52] Numerous studies have failed to present specific risk factors that themselves are able to identify a difficult intubation or laryngoscopy. Therefore, it seems rational to focus even more than hitherto on the development, testing and modification of multivariate models from and in large-scale cohort studies, hereby making the prognostication operational in everyday clinical practice.

## Supplementary material

Supplementary material is available at *British Journal of Anaesthesia* online.

## Conflict of interest

A.M.M. is a member of the Editorial Board of the *British Journal of Anaesthesia*.

## Funding

This study has been supported by the Danish Anaesthesia Database, Gentofte, Denmark; Department of Anaesthesia and Intensive Care, Herlev Hospital, Denmark; Danish Society of Anaesthesia and Intensive Medicine, Rigshospitalet, Copenhagen, Denmark; The Lundbeck Foundation, Denmark (j.nr. 381/06]; the Danish National Research Council of Health, Denmark (271-06-0586); and the Copenhagen Trial Unit, Rigshospitalet, Copenhagen, Denmark.

## References


1 Cooper GM, McClure JH. Anaesthesia chapter from *Saving mothers' lives; reviewing maternal deaths to make pregnancy safer*. *Br J Anaesth* 2008; **100**: 17–22

2 Hove LD, Steinmetz J, Christoffersen JK, Moller A, Nielsen J, Schmidt H. Analysis of deaths related to anesthesia in the period 1996–2004 from closed claims registered by the Danish Patient Insurance Association. *Anesthesiology* 2007; **106**: 675–80

3 Peterson GN, Domino KB, Caplan RA, Posner KL, Lee LA, Cheney FW. Management of the difficult airway: a closed claims analysis. *Anesthesiology* 2005; **103**: 33–9

4 Rosenstock C, Moller J, Hauberg A. Complaints related to respiratory events in anaesthesia and intensive care medicine from 1994 to 1998 in Denmark. *Acta Anaesthesiol Scand* 2001; **45**: 53–8

5 Mallampati SR, Gatt SP, Gugino LD, *et al.* A clinical sign to predict difficult tracheal intubation: a prospective study. *Can Anaesth Soc J* 1985; **32**: 429–34

6 Samsoon GL, Young JR. Difficult tracheal intubation: a retrospective study. *Anaesthesia* 1987; **42**: 487–90

7 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**: 429–37

8 Lee A, Fan LT, Gin T, Karmakar MK, Ngan Kee WD. A systematic review (meta-analysis) of the accuracy of the Mallampati tests to predict the difficult airway. *Anesth Analg* 2006; **102**: 1867–78

9 Charuluxananan S, Suraseranivongse S, Punjasawadwong Y, *et al.* The Thai Anesthesia Incidents Study (THAI Study) of anesthetic outcomes: I description of methods and populations. *J Med Assoc Thai* 2005; **7**(Suppl.): 1–13

10 Lundstrom LH, Moller AM, Rosenstock C, Astrup G, Gatke MR, Wetterslev J. Avoidance of neuromuscular blocking agents may increase the risk of difficult tracheal intubation: a cohort study of 103 812 consecutive adult patients recorded in the Danish Anaesthesia Database. *Br J Anaesth* 2009; **103**: 283–90

11 Deville WL, Buntinx F, Bouter LM, *et al.* Conducting systematic reviews of diagnostic studies: didactic guidelines. *BMC Med Res Methodol* 2002; **2**: 9

12 Cormack RS, Lehane J. Difficult tracheal intubation in obstetrics. *Anaesthesia* 1984; **39**: 1105–11

13 Yentis SM, Lee DJ. Evaluation of an improved scoring system for the grading of direct laryngoscopy. *Anaesthesia* 1998; **53**: 1041–4

14 Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003; **3**: 25

15 Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *Br Med J* 2003; **327**: 557–60

16 DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986; **7**: 177–88

17 Moses LE, Shapiro D, Littenberg B. Combining independent studies of a diagnostic test into a summary ROC curve: data-analytic approaches and some additional considerations. *Stat Med* 1993; **12**: 1293–316

18 Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. *J Am Med Assoc* 1994; **271**: 703–7

19 Egger M, Davey SG, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *Br Med J* 1997; **315**: 629–34

20 Adnet F, Racine SX, Borron SW, *et al.* A survey of tracheal intubation difficulty in the operating room: a prospective observational study. *Acta Anaesthesiol Scand* 2001; **45**: 327–32

21 Ali Z, Bithal PK, Prabhakar H, Rath GP, Dash HH. An assessment of the predictors of difficult intubation in patients with acromegaly. *J Clin Neurosci* 2009; **16**: 1043–5

22 Allahyary E, Ghaemei SR, Azemati S. Comparison of six methods for predicting difficult intubation in obstetric patients. *Iran Red Crescent Med J* 2008; **10**: 194–201

23 Arne J, Descoins P, Fusciardi J, *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

24 Ayuso MA, Sala X, Luis M, Carbo JM. Predicting difficult orotracheal intubation in pharyngo-laryngeal disease: preliminary results of a composite index. *Can J Anaesth* 2003; **50**: 81–5

25 Bouaggad A, Nejmi SE, Bouderka MA, Abbassi O. Prediction of difficult tracheal intubation in thyroid surgery. *Anesth Analg* 2004; **99**: 603–6

26 Brodsky JB, Lemmens HJ, Brock-Utne JG, Vierra M, Saidman LJ. Morbid obesity and tracheal intubation. *Anesth Analg* 2002; **94**: 732–6

27 Butler PJ, Dhara SS. Prediction of difficult laryngoscopy: an assessment of the thyromental distance and Mallampati predictive tests. *Anaesth Intensive Care* 1992; **20**: 139–42

28 Cattano D, Panicucci E, Paolicchi A, Forfori F, Giunta F, Hagberg C. Risk factors assessment of the difficult airway: an Italian survey of 1956 patients. *Anesth Analg* 2004; **99**: 1774–9

29 Constantikes J. Predicting difficult tracheal intubation using a modified Mallampati sign: a pilot study report. *CRNA* 1993; **4**: 16–20

30 Domi R. A comparison of Wilson sum score and combination Mallampati, tiromental and sternomental distances for predicting difficult intubation. *Maced J Med Sci* 2009; **2**: 141–4

31 Eberhart LH, Arndt C, Cierpka T, Schwanekamp J, Wulf H, Putzke C. The reliability and validity of the upper lip bite test compared with the Mallampati classification to predict difficult laryngoscopy: an external prospective evaluation. *Anesth Analg* 2005; **101**: 284–9

32 Ezri T, Warters RD, Szmuk P, *et al.* The incidence of class 'zero' airway and the impact of Mallampati score, age, sex, and body mass index on prediction of laryngoscopy grade. *Anesth Analg* 2001; **93**: 1073–5

33 Ezri T, Medalion B, Weisenberg M, Szmuk P, Warters RD, Charuzi I. Increased body mass index per se is not a predictor of difficult laryngoscopy. *Can J Anaesth* 2003; **50**: 179–83

34 Ezri T, Gewurtz G, Sessler DI, *et al.* Prediction of difficult laryngoscopy in obese patients by ultrasound quantification of anterior neck soft tissue. *Anaesthesia* 2003; **58**: 1111–4

35 Ezri T, Weisenberg M, Khazin V, *et al.* Difficult laryngoscopy: incidence and predictors in patients undergoing coronary artery bypass surgery versus general surgery patients. *J Cardiothorac Vasc Anesth* 2003; **17**: 321–4

36 Frerk CM. Predicting difficult intubation. *Anaesthesia* 1991; **46**: 1005–8

37 Gercek A, Lim S, Isler FB, Eti Z, Gogus FY. The prediction of difficult intubation with bedside scoring systems. *Marmara Med J* 2003; **1**: 16–9

38 Gupta S, Pareek S, Dulara SC. Comparison of two methods for predicting difficult intubation in obstetric patients. *Middle East J Anesthesiol* 2003; **17**: 275–85

39 Hester CE, Dietrich SA, White SW, Secrest JA, Lindgren KR, Smith T. A comparison of preoperative airway assessment techniques: the modified Mallampati and the upper lip bite test. *AANA J* 2007; **75**: 177–82

40 Honarmand A, Safavi MR. Prediction of difficult laryngoscopy in obstetric patients scheduled for Caesarean delivery. *Eur J Anaesthesiol* 2008; **25**: 714–20

41 Huh J, Shin HY, Kim SH, Yoon TK, Kim DK. Diagnostic predictor of difficult laryngoscopy: the hyomental distance ratio. *Anesth Analg* 2009; **108**: 544–8

42 Iohom G, Ronayne M, Cunningham AJ. Prediction of difficult tracheal intubation. *Eur J Anaesthesiol* 2003; **20**: 31–6

43 Ita CE, Eshiet AI, Akpan SG. Recognition of the difficult airway in normal Nigerian adults (a prospective study). *West Afr J Med* 1994; **13**: 102–4

44 Juvin P, Lavaut E, Dupont H, *et al.* Difficult tracheal intubation is more common in obese than in lean patients. *Anesth Analg* 2003; **97**: 595–600

45 Kamalipour H, Bagheri M, Kamali K, Taleie A, Yarmohammadi H. Lateral neck radiography for prediction of difficult orotracheal intubation. *Eur J Anaesthesiol* 2005; **22**: 689–93

46 Kaul TK, Deepika PG, Singh A, Bansal M. Prediction of difficult intubation—analysis in 500 adult patients. *J Anaesthesiol Clin Pharmacol* 1998; **14**: 323–8

47 Khan ZH, Kashfi A, Ebrahimkhani E. A comparison of the upper lip bite test (a simple new technique) with modified Mallampati classification in predicting difficulty in endotracheal intubation: a prospective blinded study. *Anesth Analg* 2003; **96**: 595–9

48 Koh LK, Kong CE, Ip-Yam PC. The modified Cormack–Lehane score for the grading of direct laryngoscopy: evaluation in the Asian population. *Anaesth Intensive Care* 2002; **30**: 48–51

49 Komatsu R, Sengupta P, Wadhwa A, *et al.* Ultrasound quantification of anterior soft tissue thickness fails to predict difficult laryngoscopy in obese patients. *Anaesth Intensive Care* 2007; **35**: 32–7

50 Krishna HM, Agarwal M, Dali JS, Rampal P, Dua CK. Prediction of difficult laryngoscopy in Indian population: role of ratio of patient's height to thyromental distance. *J Anaesthesiol Clin Pharmacol* 2005; **21**: 257–60

51 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**: 1542–5

52 L'hermite J, Nouvellon E, Cuvillon P, Fabbro-Peray P, Langeron O, Ripart J. The simplified predictive intubation difficulty score: a new weighted score for difficult airway assessment. *Eur J Anaesthesiol* 2009; **26**: 1003–9

53 Mashour GA, Kheterpal S, Vanaharam V, *et al.* The extended Mallampati score and a diagnosis of diabetes mellitus are predictors of difficult laryngoscopy in the morbidly obese. *Anesth Analg* 2008; **107**: 1919–23

54 Mashour GA, Stallmer ML, Kheterpal S, Shanks A. Predictors of difficult intubation in patients with cervical spine limitations. *J Neurosurg Anesthesiol* 2008; **20**: 110–5

55 Mashour GA, Sandberg WS. Craniocervical extension improves the specificity and predictive value of the Mallampati airway evaluation. *Anesth Analg* 2006; **103**: 1256–9

56 Merah NA, Foulkes-Crabbe DJ, Kushimo OT, Ajayi PA. Prediction of difficult laryngoscopy in a population of Nigerian obstetric patients. *West Afr J Med* 2004; **23**: 38–41

57 Merah NA, Wong DT, Ffoulkes-Crabbe DJ, Kushimo OT, Bode CO. Modified Mallampati test, thyromental distance and inter-incisor gap are the best predictors of difficult laryngoscopy in West Africans. *Can J Anaesth* 2005; **52**: 291–6

58 Noorizad S, Mahdian M. Mallampati and thyromental tests to predict difficult intubation. *J Med Sci* 2006; **6**: 169–72

59 Rocke DA, Murray WB, Rout CC, Gouws E. Relative risk analysis of factors associated with difficult intubation in obstetric anesthesia. *Anesthesiology* 1992; **77**: 67–73

60 Samra SK, Schork MA, Guinto FC Jr. A study of radiologic imaging techniques and airway grading to predict a difficult endotracheal intubation. *J Clin Anesth* 1995; **7**: 373–9

61 Savva D. Prediction of difficult tracheal intubation. *Br J Anaesth* 1994; **73**: 149–53

62 Schmitt H, Buchfelder M, Radespiel-Troger M, Fahlbusch R. Difficult intubation in acromegalic patients: incidence and predictability. *Anesthesiology* 2000; **93**: 110–4

63 Siddiqi R, Kazi WA. Predicting difficult intubation—a comparison between Mallampati classification and Wilson risk-sum. *J Coll Physicians Surg Pak* 2005; **15**: 253–6

64 Singh R, Jain A, Mishra S, Kohli P. Clinical evaluation for predicting difficult laryngoscopy in obstetric patients. *J Anaesthesiol Clin Pharmacol* 2009; **25**: 38–42

65 Topcu I, Ovali GY, Yentur EA, Kefi A, Tuncyurek O, Pabuscu Y. Clinical and radiologic evaluations for predicting difficult tracheal intubation [Italian, English]. *Acta Anaesthesiol Italica/Anaesth Intensive Care Italy* 2006; **57**: 285–300

66 Torres K, Patel AA, Stylinski K, *et al.* The body constitution of patients and intubation scales as predictors of difficult intubation considered in relation to the experience of the intubator. *Folia Morphol (Warsz)* 2008; **67**: 171–4

67 Tuzuner-Oncul AM, Kucukyavuz Z. Prevalence and prediction of difficult intubation in maxillofacial surgery patients. *J Oral Maxillofac Surg* 2008; **66**: 1652–8

68 Vani V, Kamath SK, Naik LD. The palm print as a sensitive predictor of difficult laryngoscopy in diabetics: a comparison with other airway evaluation indices. *J Postgrad Med* 2000; **46**: 75–9

69 Wong SH, Hung CT. Prevalence and prediction of difficult intubation in Chinese women. *Anaesth Intensive Care* 1999; **27**: 49–52

70 Yamamoto K, Tsubokawa T, Shibata K, Ohmura S, Nitta S, Kobayashi T. Predicting difficult intubation with indirect laryngoscopy. *Anesthesiology* 1997; **86**: 316–21

71 Yeo SW, Chong JL, Thomas E. Difficult intubation: a prospective study. *Singapore Med J* 1992; **33**: 362–4

72 Yildiz TS, Korkmaz F, Solak M, *et al.* Prediction of difficult tracheal intubation in Turkish patients: a multi-center methodological study. *Eur J Anaesthesiol* 2007; **24**: 1034–40

73 Jones CM, Athanasiou T. Summary receiver operating characteristic curve analysis techniques in the evaluation of diagnostic tests. *Ann Thorac Surg* 2005; **79**: 16–20

74 Altman DG, Bland JM. Statistics notes. Units of analysis. *Br Med J* 1997; **314**: 1874

75 Lijmer JG, Bossuyt PM, Heisterkamp SH. Exploring sources of heterogeneity in systematic reviews of diagnostic tests. *Stat Med* 2002; **21**: 1525–37

76 Higgins JP, Spiegelhalter DJ. Being sceptical about meta-analyses: a Bayesian perspective on magnesium trials in myocardial infarction. *Int J Epidemiol* 2002; **31**: 96–104

77 Li J, Zhang Q, Zhang M, Egger M. Intravenous magnesium for acute myocardial infarction. *Cochrane Database Syst Rev* 2007; CD002755

78 Rucker G, Schwarzer G, Carpenter JR, Schumacher M. Undue reliance on I(2) in assessing heterogeneity may mislead. *BMC Med Res Methodol* 2008; **8**: 79