Brazilian Journal of Anesthesiology 2023;73(4):491–499

Sociedade Brasileira de Anestesiologia

Brazilian Journal of ANESTHESIOLOGY

SYSTEMATIC REVIEW

Predictive performance of thyromental height for difficult laryngoscopies in adults: a systematic review and meta-analysis

Clístenes Crístian de Carvalho[a,*], **Jayme Marques Santos Neto**[b], **Flávia Augusta de Orange**[a]

[a] *Instituto de Medicina Integral Prof. Fernando Figueira (IMIP), Department of Post-graduation, Recife, PE, Brazil*
[b] *Universidade Federal de Pernambuco (UFPE), Hospital das Clínicas, Department of Anesthesia, Recife, PE, Brazil*

Received 27 February 2021; accepted 20 June 2021
Available online 9 July 2021

KEYWORDS
Airway management;
Laryngoscopy;
Intubation, intratracheal;
Systematic review;
Meta-analysis;
Sensitivity and specificity

Abstract
Background: Thyromental height (TMH) was first reported as a great single test for prediction of difficult laryngoscopies, although further studies have shown variable estimates of its accuracy. We thus performed this meta-analysis to summarize the predictive values of TMH mainly for prediction of difficult laryngoscopies.
Methods: A search in PubMed, EMBASE, LILACS, and Scielo was conducted in June 2020. We included prospective cohorts fully reported with patients $\geq$ 16 years old, providing data on predictive values of TMH for prediction of either difficult laryngoscopies or difficult intubations. Diagnostic properties and association between TMH and Cormack and Lehanes's classification by direct laryngoscopy were evaluated. A random-effects meta-analysis using hierarchical models was performed.
Results: Eight studies evaluating 2844 patients were included. All included studies had high risk of bias and low concern regarding applicability. There was significant heterogeneity among the studies. The pooled diagnostic odds ratio (DOR) and positive (LR+) and negative (LR-) likelihood ratios were as follows: DOR, 57.94 (95% CI: 18.19–184.55); LR+, 11.32 (95% CI: 4.28–29.92); and LR-, 0.23 (95% CI: 0.15–0.35). Summary sensitivity and specificity for studies with common threshold were 82.6 (95% CI: 74–88.8%) and 93.5 (95% CI: 79–98.2%), respectively. The estimated AUC was 81.1%.
Conclusion: TMH arises as a good predictor of difficult laryngoscopies in adult patients from diverse populations presenting better predictive values than most previously reported bedside tests. However, the high risk of bias throughout the studies may have skewed the results of the individual research as well as the summary points of the present meta-analysis.

* Corresponding author.
E-mail: clistenescristian@hotmail.com (C.C. Carvalho).

https://doi.org/10.1016/j.bjane.2021.06.015

Introduction

Managing an unanticipated difficult airway usually comes with suboptimal care, which may take patients to a life-threatening scenario.[1–3] Physicians have then searched for a way of predicting difficult airways in order to build an appropriate approaching strategy. However, the predictive performance of most tests available so far is unreliable, their role in airway prediction is unclear, and anticipating difficult airways still remains challenging.[4–6]

Although it is unclear the extent of the value of airway assessment for airway prediction, anesthesia organizations around the globe recommend performing a physical examination in search for factors that might suggest possible difficult airways.[1,2,7–9] The upper airway assessment is supposed to take multiple features into account since a comprehensive examination is assumed to improve airway prediction and enhance sensitivity as compared to the use of a single test.[7,10,11] Nonetheless, despite the use of multivariable approaches, airway prediction still lacks accuracy to segregate easy and difficult airways.[6,12]

On the other hand, a novel and promising bedside test was recently described by Etezadi et al. It is the thyromental height (TMH) – an easily performed measure, drawn by a ruler with patient in supine position and defined as the height from the anterior border of thyroid cartilage to the notch of mentum.[13] Its predictive threshold for difficult laryngoscopies is regarded by initial research as TMH $\leq$ 5 cm. This measure works as an estimate of the antero-posterior position of larynx, which the authors have reported to present surprising predictive values – higher than most known single tests and multivariable scores.[4,5] However, the relatively small sample size of the study as well as the airway manipulation by training anesthesiologists, and the lack of prespecified threshold may bring some concern regarding the precision of the predictive values as well as the validity and generalizability of their results.[13]

Additionally, other studies have been conducted since the original one with variable estimates of its diagnostic test accuracy, what makes unclear the usefulness of this measure.[14–20] As no systematic review has been found by authors over the role of TMH to predict either difficult laryngoscopies or difficult intubations, the current meta-analysis was designed to better understand the summary performance of the TMH to anticipate both difficult laryngoscopies and difficult orotracheal intubations by direct laryngoscopies in apparently normal patients undergoing general anesthesia.

Methods

Protocol and registration

The current review was designed and prepared according to recommended standards and reported as per the Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA) guidelines.[21–28] A review protocol was prepared and registered before commencing the screening steps (PROSPERO registration number: CRD42020184439) and followed during the study. Search strategy was designed according to the PRISMA guidelines.

Eligibility criteria

Inclusion criteria were as follows: 1) prospective cohorts or clinical trials in full reports; 2) patients aging 16 years or older; 3) data available on thyromental height for prediction of difficult laryngoscopies or endotracheal intubation that was measured in the same manner for all patients in each individual study; 4) endotracheal intubation by direct laryngoscopies under general anesthesia. Exclusion criteria were as follows: 1) studies published in language other than English, Spanish, or Portuguese; 2) inability to abstract relevant data.

Literature search

We conducted a computerized search (11/05/2020) through PubMed, LILACS, EMBASE, and Scielo, updated on August 8, 2020. The search strategy used was (''thyromental height'' [all fields]) OR (''altura tir?omentoniana'' [all fields]) without any limitation. We also searched the reference lists of included studies. The title and abstract of each citation were reviewed in duplicate by two of the reviewers (CC and JSN), with full-text retrieval of any citation that either reviewer considered potentially relevant for assessing the predictive performance of thyromental height for prediction of both difficult laryngoscopy and difficult intubation.

Study selection

Two reviewers (CC and JSN) independently assessed the full text of each retrieved citation. The selection was based on the eligibility criteria. Disagreements were resolved by consensus among all authors.

Data-collection process and data items

We collected or calculated data in duplicate by independent reviewers through a standardized form on author, year of publication, type of study, age, sex, height, weight, body mass index (BMI), true positive, false negative, false positive, true negative, sample size of difficult airway groups, sample size of easy airway groups, and threshold used. Where data was missing, the corresponding author was e-mailed. Where relevant data was not presented or data was conflicting and the corresponding author did not reply our contact after one month, the study was excluded. Duplicated datasets were compared, and disagreements were resolved by consensus among all authors.

Diagnostic accuracy measures

Primary outcome was sensitivity and specificity for difficult laryngoscopies classified according to Cormack and Lehane's grading system. The best view achieved during the manipulations was used for definition of difficult laryngoscopies. Secondary outcome was the sensitivity and specificity for difficult intubation classified according to definition used by authors from each study other than Cormack and Lehane's classification. We alternatively evaluated diagnostic odds ratio, and positive and negative likelihood ratios.

Risk of bias in individual studies

Two authors (CC and JSN) assessed risk of bias and generalizability using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool for diagnostic tests.[29] Studies were evaluated for risk of bias with signalling questions for (1) patient selection; (2) index test; (3) reference standard; and (4) patient flow and timing between the index test and reference standard (coded 'yes', 'no' or 'unclear'), with the first three domains also considered in terms of applicability (representativeness), which was coded 'high', 'low' or 'unclear'. Disagreements were resolved by discussion and consensus of all authors.

Summary measures and synthesis of results

Analyses were conducted in Review Manager[30] (RevMan, London, UK, v5.3.5), and R software[31] tools (R Foundation for Statistical Computing, Vienna, Austria), as appropriate. For R analyses, we used the ''mada'', the ''meta'', the ''metafor'', the ''metaviz'', and the ''dmetar'' packages.[32] The R code for such analyses is available in https://rpubs.com/clistenescarvalho/753921. Diagnostic properties of thyromental height for difficult laryngoscopy were extracted or calculated. Forest plots were constructed for OR, LR+, and LR- (Fig. 1), as well as for sensitivity and specificity (Fig. 2) of thyromental height for diagnostic of difficult laryngoscopies. Summary ROC (SROC) plot with confidence regions for primary study estimates was built (Fig. 3). Heterogeneity was evaluated qualitatively and quantitatively by Cochran's Q-test along with correlation test by Spearman method to assess presence of ''threshold effect''. A Baujat plot (Supplement Fig. 1) was also built to identify the contribution of each study to both overall heterogeneity and results. Pooled estimates based on DerSimonian and Laird random-effects models were calculated where heterogeneity was present. SROC curves were also generated using a bivariate random-effects approaching through a linear mixed model (Supplement Fig. 2).[23,32] Summary sensitivity and specificity were estimated for studies with the same threshold (5 cm).

We performed sensitivity analyses to explore the impact of two features on our results: performance of external laryngeal manipulation, and risk of bias.

Risk of publication bias across studies

Publication bias was investigated using funnel plots and the Durval and Tweedie trim-and-fill approach, a method that first identifies potentially unpublished estimates based on funnel plot asymmetry and then includes these unpublished estimates in a revised pooled value. We employed the Egger's weighted regression method with precision (1/standard error) and log odds ratio plotted (Supplement Fig. 3). The intercept value in Egger's regression method provides an estimate of asymmetry of the funnel plot, with positive values indicating a trend towards higher levels of test accuracy in studies with smaller sample size. The threshold of significance was set at $p < 0.100$ for this method as this test has low power.

Results

Study selection

Our initial electronic search identified 26 articles with 17 remaining after deduplication. These 17 papers had their full text assessed and after applying eligibility criteria, nine articles were thoroughly read. One of these nine articles presented conflicting data and then a total of 8 studies were included in the analysis (Fig. 4). Studies were not included or excluded mainly for the following reasons (Fig. 4): pediatric population (2); diverse language (1); letter (1); full report not available (4); duplicate data (9); and conflicting data[33] (1).

Study characteristics

Of these 8 studies, all presented data on difficult laryngoscopies according to CL grading system (n = 2844). Only one paper reported data on difficult intubation by definition other than CL classification (n = 120)[20] (Table 1), preventing us of undertaking the meta-analysis over this outcome. The studies were prospective cohorts conducted in Iran, Turkey, India, Bangladesh, Australia, Japan, and Egypt. Studied patients were recruited from those schedules for surgical procedures under general anesthesia in tertiary centers. All studies included only apparently normal patients, judged not to be at high risk of difficult airways by the attending anesthesiologists. One study included only patients undergoing coronary artery bypass surgery,[15] and another one included only elderly patients.[20] Two studies did not use digital, but regular, rulers for TMH.[16,20] The type of blade used was reported in 7 studies.[13−15,17−20] Macintosh's blade was used in all these studies. Four of them used either sizes 3 or 4[15,17−19]; one used only size 4[13]; one used either sizes 4 or 5[14]; and one did not mention the size used.[20] Miller's blade was used in one study when no laryngeal view was achieved with Macintosh's.[13] Seven studies reported direct laryngoscopies were performed by experienced anesthesiologists,[14,18−20] whilst one, by residents.[13] One study did not report if the airway manager was aware of the index test.[15] Four studies classified glottic view supported by external manipulation of larynx,[13,15,18,20] two did not applied it,[14,17] and two did not mention it.[16,19] Seven studies used CL grades 3 or 4 as difficult laryngoscopies and one[17] used CL 2b or higher to define difficulty. Two studies[17,18] reported measures of neuromuscular blockade depth, whilst six[13−16,19,20] did not report it. Sniffing position was employed in six studies,[13,15,17−20] whilst two[14,16] did not report data on it.

Risk of bias within studies

According to QUADAS-2 tool, all eight included studies had low applicability concerns. All studies were also at risk of

Table 1 Baseline characteristics of individual studies.

Author	Year	Design	Mean age; y	Sex	Height	Weight	BMI	DL; n (%)	Total sample size	Threshold; cm
Mostafa	2020	Cohort	68	Male: 57% Female: 43%	NA	NA	27.08	15 (12%)	120	5.7
Panjiar	2019	Cohort	37.19	Male: 43.6% Female: 56.4%	158.4	61.07	24.52	55 (10%)	550	5.1
Yabuki	2019	Cohort	50.6	Male: 18% Female: 82%	159.6	58.6	22.9	6 (1%)	609	5
Rao	2018	Cohort	43.4	Male: 47% Female:53%	162.6	62	23.4	26 (8.2%)	316	5
Nurullah	2018	Cohort	45.4	Male: 50.4% Female: 49.6%	NA	NA	NA	43 (31%)	139	5
Jain	2017	Cohort	56.7	Male: NA Female: NA	162.6	65.3	24.72	32 (9.3%)	345	5
Selvi	2017	Cohort	48.49	Male: 51% Female: 49%	NA	77.65	NA	37 (8.2%)	451	5
Etezadi	2013	Cohort	44.5	Male: 47.5% Female: 52.5%	166.1	72	25.8	23 (7.3%)	314	5

y, years of age; BMI, body mass index; DL, difficult laryngoscopy; NA, not available.

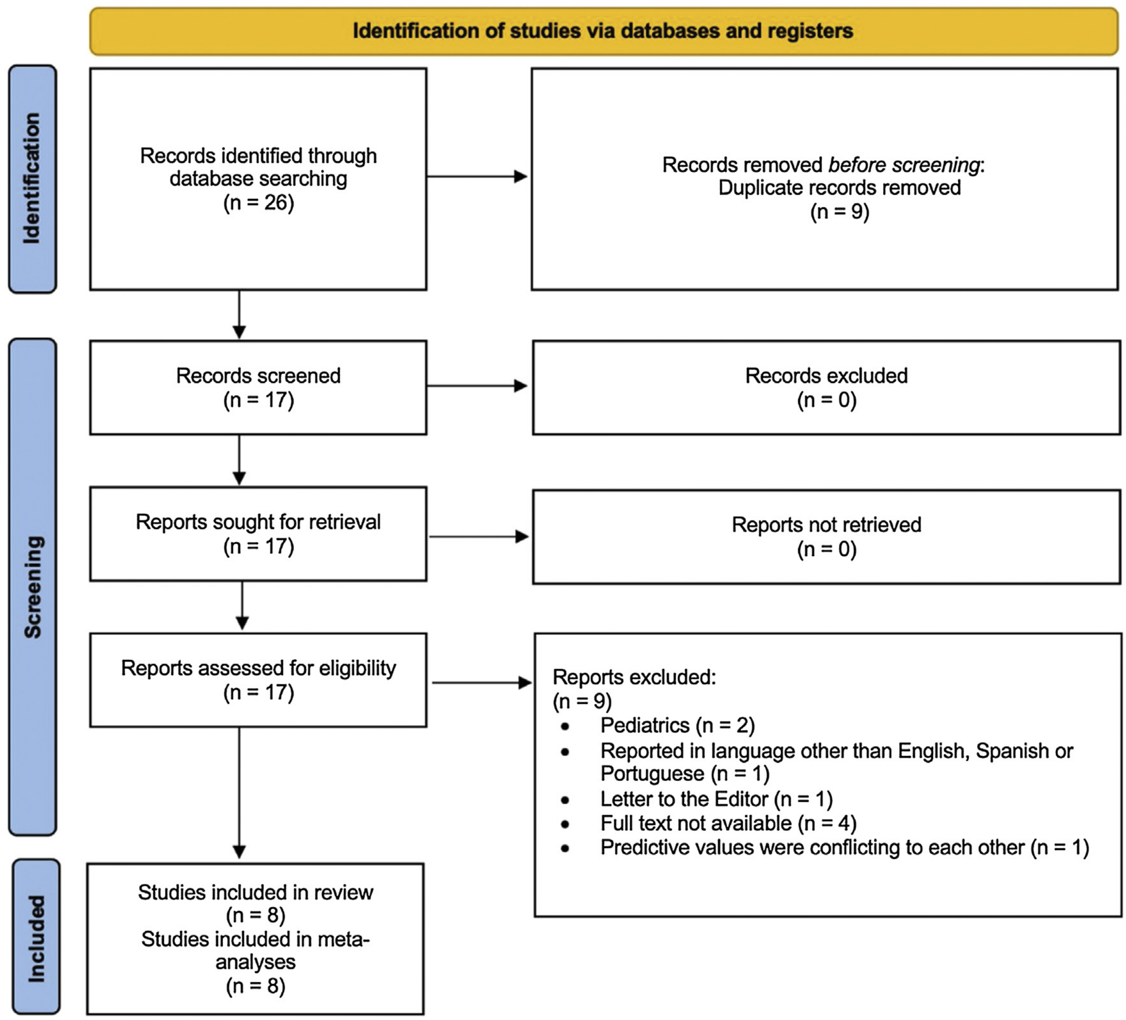

Figure 1 Systematic review flow diagram (PRISMA flow chart).

bias mainly due to concerns over patient's selection and the reference standard (Fig. 5). Six studies[14–19] were assumed to present inappropriate patient selection with exclusion of patients difficult to define THM such as obese and those with anatomic alterations as well as known to be at high risk of difficult airways, thus skewing the test accuracy. Four studies were at high risk of bias regarding index test, two[13,19] because of non-prespecified threshold (overestimating the test accuracy), and two[16,20] due to diverse and non-objective ways of measures. Only one single study was considered at low risk of bias regarding reference standard.[18] The main concerns over the reference standard were performer experience, standardization and measure of neuromuscular blockade depth, application of external manipulation over the larynx, performance of sniffing position, threshold of Cormack and Lehane's classification for difficulty, and awareness of index test. These features may have made studies heterogeneous for this domain.

Results of individual studies

Summaries of individual studies characteristics are presented in Table 1 and Figs. 1–3.

Synthesis of results

Sensitivity and specificity ranged from 50%[18] to 95%,[16] and 52%[14] to 99%,[13,17] respectively (Fig. 2).

Summary sensitivity and specificity for studies with the same threshold[13–19] were 82.6 (95% CI: 74–88.8%) and 93.5 (95% CI: 79–98.2%), respectively. Summary sensitivity and specificity for studies with prespecified threshold (5 cm)[14–18] were 84.1% (95% CI: 68.3–92.9%) and 90.4% (65–98%), respectively.

Only one paper[20] presented data on difficult intubation by classification other than Cormack and Lehane's grade system, preventing us of performing the meta-analysis over this outcome.

For difficult laryngoscopies, the frequency ranged from 1%[18] to 31%.[16] A summary of studies' patients baseline characteristics as well as further studies information are presented in Table 1.

Only one study did not present significant association between TMH and difficult laryngoscopies, when supported by external laryngeal manipulation (Fig. 1).[18] When laryngoscopy view was graded without laryngeal manipulation, the association was present.[18]

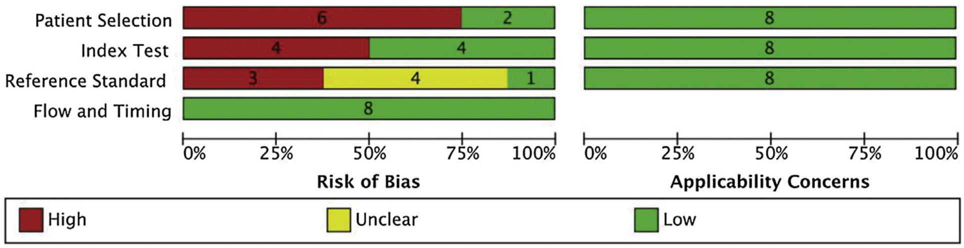

Figure 2 Risk of bias and applicability concerns graph for difficult laryngoscopy: summary of review authors' judgment about each domain presented as percentages across included studies.

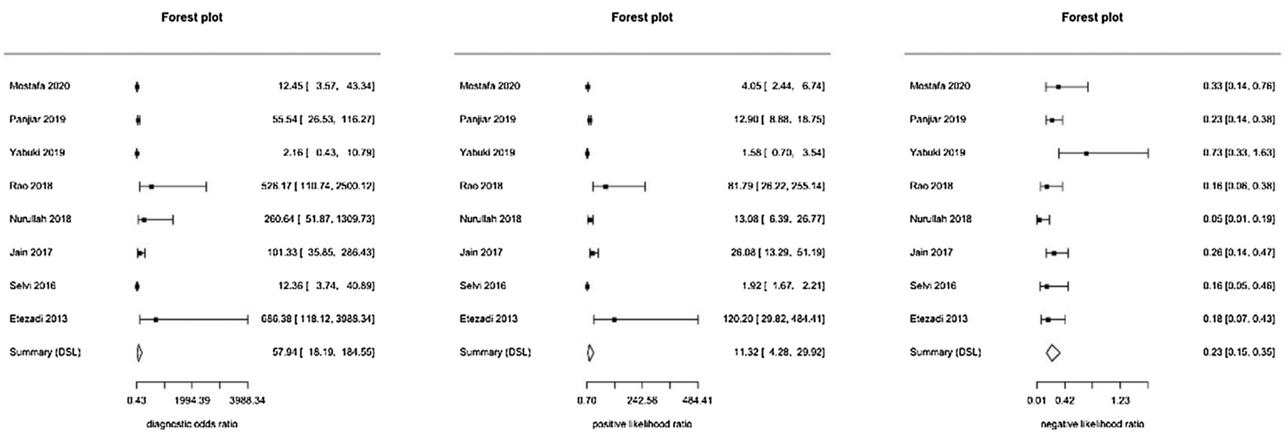

Figure 3 Forest plots for univariate analyses with summary measures of diagnostic odds ratio, positive and negative likelihood ratios.

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Nurullah 2018	41	7	2	89	0.95 [0.84, 0.99]	0.93 [0.86, 0.97]
Selvi 2017	34	198	3	216	0.92 [0.78, 0.98]	0.52 [0.47, 0.57]
Rao 2018	22	3	4	287	0.85 [0.65, 0.96]	0.99 [0.97, 1.00]
Etezadi 2013	19	2	4	289	0.83 [0.61, 0.95]	0.99 [0.98, 1.00]
Panjiar 2019	43	30	12	465	0.78 [0.65, 0.88]	0.94 [0.91, 0.96]
Jain 2017	24	9	8	304	0.75 [0.57, 0.89]	0.97 [0.95, 0.99]
Mostafa 2020	11	19	4	86	0.73 [0.45, 0.92]	0.82 [0.73, 0.89]
Yabuki 2019	3	191	3	412	0.50 [0.12, 0.88]	0.68 [0.64, 0.72]

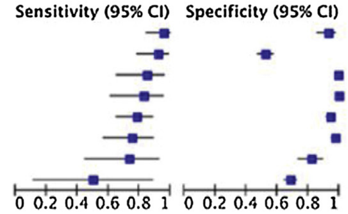

Figure 4 Forest Plot of thyromental height for difficult laryngoscopy, sorted by descending sensitivity. TP, true positive; FP, false positive; FN, false negative; TN, true negative.

The positive and negative likelihood ratios ranged from 1.58[18] to 120.20[13] and 0.05[16] to 0.73,[18] respectively. The summary points for univariate analyses were as follows: DOR, 57.94 (95% CI: 18.19–184.55); LR+, 11.32 (95% CI: 4.28–29.92); and LR-, 0.23 (95% CI: 0.15–0.35). Further information about univariate analyses are presented in Figure 1.

The estimated area under ROC curve for all 8 studies was 81.1% (Supplement Fig. 2).

Heterogeneity and risk of publication bias across studies

There was significant heterogeneity in both sensitivity ($p = 0.038$) and specificity ($p < 0.001$) of thyromental height for difficult laryngoscopies. It may have limited the robustness of pooled estimates, which therefore must be regarded with caution. The hypothesis of threshold effect as a possible reason for the caught heterogeneity was tested and rejected (Spearman correlation estimate 0.119; $p = 0.793$). A single study[18] was most implicated in the heterogeneity (Supplement Fig. 1).

In a visual inspection of the funnel plot (Supplement Fig. 3) used to check for publication bias in this meta-analysis, studies were distributed asymmetrically around the pooled estimate, suggesting a possibility for publication bias. It was confirmed by a regression test with p-value < 0.0148. The trim-and-fill method also suggested that there were possible unpublished studies (Supplement Fig. 3).

Sensitivity analysis

We conducted four separate sensitivity analyses to investigate if any of the following features significantly altered the pooled results: external laryngeal manipulation and risk of bias regarding patient selection, index test, and reference standard. The only variable that affected the results was the high risk of bias regarding reference standard (ML false positive rate, fixed-effect coefficient -3.15; 95% CI: -5.59 to -0.71; $p = 0.011$; 3 studies; 1180 participants).

Discussion

The present systematic review showed quite good diagnostic accuracy of TMH for difficult laryngoscopies, according to the pooled sensitivity and specificity, along with other predictive values (sensitivity, 82.6%; specificity, 93.5%; LR+, 11.32; LR-, 0.23; bal acc, 88%). However, relevant imprecision was in place − with large 95% confidence intervals, significant heterogeneity, and scarcity of well-designed studies to validate its performance.

The predictive performance of TMH found in our meta-analysis was better than those from all predictors recently evaluated in other systematic review.[4] Roth et al. assessed the diagnostic accuracy of seven tests for difficult laryngoscopies: Mallampati test, modified Mallampati test, Wilson risk score, thyromental distance, sternomental distance, mouth opening, and upper lip bite test. The test with highest sensitivity was upper lip bite test (sensitivity, 67%; specificity, 92%; LR+, 8.37; LR-, 0.36; balanced accuracy [bal acc], 79.5%), whilst the test with highest specificity and LR+ was Wilson risk score (sensitivity, 51%; specificity, 95%; LR+, 10.2; LR-, 0.52; bal acc, 73%). The single test with highest specificity was mouth opening (sensitivity, 22%; specificity, 94%; LR+, 3.66; LR-, 0.83; bal acc, 58%). Thus, as compared to our results for TMH, most predictors presented close specificity, but all showed considerably lower sensitivity and balanced accuracy. Our result for LR+ of TMH was higher than that of most predictors, and similar to that of Wilson risk score. Also, TMH presented lower LR- than all seven predictors. Therefore, TMH performed better in our series of studies than other seven predictors did in a different meta-analysis. As in our analysis, few studies included in the systematic review of Roth et al. were at low risk of bias.[4]

Another meta-analysis evaluated 12 bedside tests for airway prediction: upper lip bite test, Wilson score, hyomental distance, retrognathia, impaired mandibular protrusion, ratio of height to thyromental distance, impaired neck mobility, sternomental distance, modified Mallampati, impaired mouth opening, thyromental distance, and palm print. The test with highest sensitivity was the palm print (sensitivity, 77%; specificity, 84%; LR+, 3; LR-, 0.28; bal acc, 80.5%). The test with highest specificity was retrognathia (sensitivity, 19%; specificity, 98%; LR+ 6; LR-, 0.85; bal acc, 58.5%). The test with highest LR+ was upper lip bite test (sensitivity, 60%; specificity, 96%; LR+ 14; LR-, 0.42; bal acc, 78%). Wilson risk score presented sensitivity of 43%, specificity of 95%, LR+ of 9.1, LR- of 0.6, and balanced accuracy of 69%, which was comparable to the results of Roth et al. As before, the predictive values of TMH in our meta-analysis also were better than that of all predictors in the meta-analysis of Detsky et al.[5] This superior TMH predictive performance draw our attention to what seems to be a major role of antero-posterior larynx position for the occurrence of difficult laryngoscopies.

Despite the great performance shown by TMH as a single test, some studies have demonstrated that multivariable

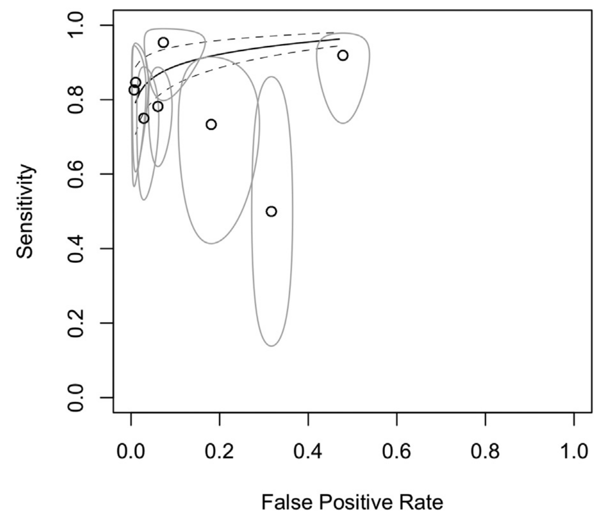

Figure 5 Summary receiver operating curve (SROC) plot along with SROC curve by proportional hazard model approach. Open circle (o) represents false positive rate (x-coordinate) and sensitivity (y-coordinate) of individual studies. Size of bubbles reflects precision of estimate. AUC = 95.2%.

assessments are superior for airway prediction, mainly by enhancing sensitivity values.[11,34,35] It makes sense since there are a multitude of factors determining the emergence of a difficult airway and has led authors to recommend the assessment of multiple features during the physical examination.[7,10,35] It is in accordance with the findings of Selvi et al., who encountered higher sensitivity for Mallampati and thyromental height when combined as compared to each one alone.[14] This way, a multivariable approach including TMH might be a good alternative for airway prediction and further studies investigating predictive scores and models with TMH would be worthwhile.

It is important to state that in the present meta-analysis, the largest study[18] was also the only one regarded as at low risk of bias for direct laryngoscopies. In this study, the frequency of difficult laryngoscopies was the lowest (1%) and the performance of the TMH was the worst, with no significant association being presented between TMH and DL when applying external laryngeal manipulation. The best predictive performance of TMH, on the other hand, was presented in a study with uncommon high frequency of difficult laryngoscopies (31%)[16] − which by the way was not the one conducted with residents. This behavior is in accordance with the well-known dependence of a test accuracy on the prevalence of disease in the population.[36] These aspects along with the remaining risk of bias and heterogeneity across the studies bring further concern over the validity of our pooled results.

Another intriguing characteristic was how largely the frequency of difficult laryngoscopies ranged (1−31%) among studies. It raises some concern about the reliability of Cormack and Lehane’s classification as a tool to define difficulty of laryngoscopy. Many instances may have been responsible for this variability such as some tool subjectivity and the non-standardized use of external laryngeal manipulation, patients’ positioning, threshold for difficult laryngoscopy, and depth of neuromuscular blockade as well as the use of different types of laryngoscope blades. On the other hand, manipulator experience did not seem to be a major concern in this regard since the only study conducted with non-experienced anesthetists − residents − showed second to the lowest frequency (7.3%).[13]

Notwithstanding, the association between TMH and difficult laryngoscopies was consistent and replicable throughout most studies. Furthermore, the overall predictive performance of the test was great and promising, even when compared to either single or composite scores. From these findings, TMH might be regarded at least as a risk factor for difficult laryngoscopies, while larger and more well-designed studies are yet not available.

Apart from the weaknesses discussed so far for the present meta-analysis, such as significant heterogeneity and high risk of bias across individual studies, other limitations should be taken into account. We excluded four studies not fully reported and one due to language, what reduced the total number of patients evaluated. The results were based on a small set of small studies, what also led to a large 95% CI and further compromised the reliability of our pooled results. Since there was no threshold effect, the estimated AUC should be regarded with caution. We did not assess the role of TMH in particular scenarios and populations such as in obese or elderly patients, in the emergency setting during anesthesia, in the ICU, in the emergency departments, and outside hospitals. Therefore, caution should be exercised in translating our results to these clinical settings. Also, we did not distinguish performance differences between different types of rulers for measuring TMH as well as between different types of laryngoscopy blades.

Conclusion

We found great predictive values of TMH for difficult laryngoscopies in adult patients submitted to general anesthesia for elective procedures. The pooled estimates of both sensitivity and specificity were higher than the estimates of all other predictors performed in different meta-analyses. It gives support to the use of TMH during routine airway assessment. However, our summary results originated from a small set of small and heterogeneous studies with high or unclear risk of bias in many domains. This way, further studies with larger sample sizes and more well-designed methods would be necessary to better understand the actual predictive performance of TMH.

Conflicts of interests

The authors declare no conflicts of interest.

Acknowledgments

The authors are grateful to Jaideep J Pandit, DPhil, Nuffield Department of Anaesthesia, University of Oxford, Oxford, United Kingdom, for his advices during the study performance and for reviewing the paper.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.bjane.2021.06.015.

References

1. Cook TM, Woodall N, Frerk C. Major complications of airway management in the UK: Results of the Fourth National Audit Project of the Royal College of Anaesthetists and the Difficult Airway Society. Part 1: Anaesthesia. Br J Anaesth. 2011;106:617−31.
2. Frerk C, Mitchell VS, McNarry AF, et al. Difficult Airway Society 2015 guidelines for management of unanticipated difficult intubation in adults. Br J Anaesth. 2015;115:827−48.
3. Higgs A, McGrath BA, Goddard C, et al. Guidelines for the management of tracheal intubation in critically ill adults. Br J Anaesth. 2018;120:323−52.
4. Roth D, Pace NL, Lee A, et al. Bedside tests for predicting difficult airways: an abridged Cochrane diagnostic test accuracy systematic review. Anaesthesia. 2019;74:915−28.
5. Detsky ME, Jivraj N, Adhikari NK, et al. Will this patient be difficult to intubate? the rational clinical examination systematic review. JAMA. 2019;321:493−503.
6. Nørskov AK, Rosenstock CV, Wetterslev J, et al. Diagnostic accuracy of anaesthesiologists' prediction of difficult airway management in daily clinical practice: a cohort study of 188 064 patients registered in the Danish Anaesthesia Database. Anaesthesia. 2015;70:272−81.
7. Apfelbaum JL, Hagberg CA, Caplan RA, et al. Practice Guidelines for Management of the Difficult Airway: An Updated Report by the American Society of Anesthesiologists Task Force on Management of the Difficult Airway. Anesthesiology. 2013;118:251−70.
8. Cook TM, Woodall N, Harper J, et al. Major complications of airway management in the UK: results of the Fourth National Audit Project of the Royal College of Anaesthetists and the Difficult Airway Society. Part 2: intensive care and emergency departments. Br J Anaesth. 2011;106:632−42.
9. Law JA, Broemling N, Cooper RM, et al. The difficult airway with recommendations for management - Part 2 - Difficult airway. Can J Anesth. 2013;60:1119−38.
10. Pandit JJ, Heidegger T. Putting the ''point'' back into the ritual: a binary approach to difficult airway prediction. Anaesthesia. 2017;72:283−8.
11. Vannucci A, Cavallone LF. Bedside predictors of difficult intubation : a systematic review. Minerva Anestesiol. 2016;82:69−83.
12. Hung O, Law JA, Morris I, et al. Airway assessment before intervention: What we know and what we do. Anesth Analg. 2016;122:1752−4.
13. Etezadi F, Ahangari A, Shokri H, et al. Thyromental height: A new clinical test for prediction of difficult laryngoscopy. Anesth Analg. 2013;117:1347−51.
14. Selvi O, Kahraman T, Senturk O, et al. Evaluation of the reliability of preoperative descriptive airway assessment tests in prediction of the Cormack-Lehane score: A prospective randomized clinical study. J Clin Anesth. 2017;36:21−6.
15. Jain N, Das S, Kanchi M. Thyromental height test for prediction of difficult laryngoscopy in patients undergoing coronary artery bypass graft surgical procedure. Ann Card Anaesth. 2017;20:207−11.
16. Nurullah M, Alam MS, Hossen M, et al. Prediction of difficult airway by thyromental height test-a comparison with modified mallampati test. Bangladesh J Med Sci. 2018;17:455−61.
17. Rao KVN, Dhatchinamoorthi D, Nandhakumar A, et al. Validity of thyromental height test as a predictor of difficult laryngoscopy:

A prospective evaluation comparing modified Mallampati score, interincisor gap, thyromental distance, neck circumference, and neck extension. Indian J Anaesth. 2018;62:603–8.
18. Yabuki S, Iwaoka S, Murakami M, et al. Reliability of the thyromental height test for prediction of difficult visualisation of the larynx: A prospective external validation. Indian J Anaesth. 2019;63:270–6.
19. Panjiar P, Kochhar A, Bhat K, et al. Comparison of thyromental height test with ratio of height to thyromental distance, thyromental distance, and modified Mallampati test in predicting difficult laryngoscopy: A prospective study. J Anaesthesiol Clin Pharmacol. 2019;35:390–5.
20. Mostafa M, Saeed M, Hasanin A, et al. Accuracy of thyromental height test for predicting difficult intubation in elderly. J Anesth. 2020;34:217–23.
21. Macaskill P, Gatsonis C, Deeks JJ, et al. Analysing and presenting results. In: Deeks JJ, Bossuyt PM, Gatsonis C, editors. Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0. The Cochrane Collaboration; 2010. Available at: http://srdta.cochrane.org/. (Accessed April 1, 2020).
22. Kim KW, Lee J, Choi SH, et al. Systematic review and meta-analysis of studies evaluating diagnostic test accuracy: A practical review for clinical researchers–part I. General guidance and tips. Korean J Radiol. 2015;16:1175–87.
23. Lee J, Kim KW, Choi SH, et al. Systematic review and meta-analysis of studies evaluating diagnostic test accuracy: A practical review for clinical researchers–part II. Statistical methods of meta-analysis. Korean J Radiol. 2015;16:1188–96.
24. Bramer WM, de Jonge GB, Rethlefsen ML, et al. A systematic approach to searching: An efficient and complete method to develop literature searches. J Med Libr Assoc. 2018;106:531–41.
25. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev. 2015;4:1.
26. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. PLoS Med. 2009;6:e1000097.
27. McGrath TA, Alabousi M, Skidmore B, et al. Recommendations for reporting of systematic reviews and meta-analyses of diagnostic test accuracy: A systematic review. Syst Rev. 2017;6:194.
28. Mclnnes MDF, Moher D, Thombs BD, et al. Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies: The PRISMA-DTA Statement. JAMA. 2018;319:388–96.
29. Whiting PF, Rutjes AWS, Westwood ME, et al. Quadas-2: A revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155:529–36.
30. Review Manager (RevMan) [Computer program]. Version 5.3. Copenhagen, Denmark: The Nordic Cochrane Centre,. The Cochrane Collaboration; 2014.
31. R CoreTeam. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2013.
32. Shim SR, Kim S-J, Lee J. Diagnostic test accuracy: application and practice using R software. Epidemiol Health. 2019;41:e2019007.
33. Palczynski P, Bialka S, Misiolek H, et al. Thyromental height test as a new method for prediction of difficult intubation with double lumen tube. PLoS One. 2018;13:1–13.
34. Shiga T, Wajima Z, Inoue T, et al. Predicting difficult intubation in apparently normal patients: a meta-analysis of bedside screening test performance. Anesthesiology. 2005;103:429–37.
35. Nørskov AK. Pre-operative Airway Assessment – Experience Gained from a Multicentre Cluster Randomised Trial and the Danish Anaesthesia Database. [PhD thesis]. Copenhagen, Denmark: University of Copenhagen; 2015.
36. Fischer JE, Bachmann LM, Jaeschke R. A reader's guide to the interpretation of diagnostic test properties: clinical examples of sepsis. Intensive Care Med. 2003;29:1043–51.