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

INTRODUCTION

As a bedside screening tool, thyromental height test (TMHT) has been shown to be more accurate than the modified Mallampati score, thyromental distance (TMD), and sternomental distance (SMD) in regards to the sensitivity and positive predictive value (PPV). It is independent of the mobility of cervical spine, dentition, and patient's cooperation.

K Venkata Nageswara Rao1, D Dhatchinamoorthi2, Amar Nandhakumar2, N Selvarajan2, Hani Rajesh Akula3, Venkatesan Thiruvenkatarajan4,5
1Department of Anaesthesia, Lokamanya Hospitals Pune, Maharashtra, 2Department of Anaesthesia, Koval Medical Centre and Hospital, Coimbatore, Tamil Nadu, 3Department of Community Medicine, Alluri Sitarama Raju academy of Medical Sciences, Eluru, Andhra Pradesh, India, 4Department of Anaesthesia, The Queen Elizabeth Hospital, Woodville, 5Discipline of Acute Care Medicine, The University of Adelaide, Adelaide, Australia

ABSTRACT

Background and Aims: Thyromental height test (TMHT) is a recently described anatomical bedside screening tool in predicting difficult laryngoscopy. It has been shown to be more accurate than the modified Mallampati score, thyromental distance (TMD), and sternomental distance with regard to sensitivity and positive predictive value (PPV). Airway assessment studies based on the anatomic parameters of the upper airway are limited in the subcontinent population. We attempted this study to evaluate and validate the predictive value of TMHT at 50 mm in an Indian population in predicting difficult laryngoscopy. Methods: This prospective observational study was conducted in a tertiary teaching hospital on 340 patients. TMHT along with other bedside predictors of difficult intubation, including modified Mallampati score, interincisor gap (IIG), TMD, neck circumference (NC), and neck extension were assessed. We compared the sensitivity, specificity, PPV, negative predictive value (NPV), and diagnostic accuracy of TMHT with other bedside tests such as the modified Mallampati score, IIG, TMD, NC, and neck extension individually in predicting difficult laryngoscopy. Any Cormack and Lehane’s intubation grade II b and above was considered to be difficult laryngoscopy. Results: TMHT had the highest sensitivity (84.62%) and specificity (98.97%), and had the most PPV (88%) and NPV (98.63%) when compared with the modified Mallampati score, IIG, TMD, NC, and neck extension individually in predicting difficult laryngoscopy. TMHT was followed by the modified Mallampati score and IIG. Conclusion: TMHT appears promising as a single anatomical measure to predict the risk of difficult laryngoscopy, however, validation will require further studies in more diverse patient populations.

Key words: Airway management, intubation, laryngoscopy, Mallampati, thyromental distance
and hence, not limited by head extension. A cutoff value of approximately 50 mm has been originally described in an Iranian population, and values below this were noted to be highly correlative of difficult laryngoscopy. Using manual measurements, a study in an Indian population\(^2\) was able to reproduce a similar sensitivity, specificity, PPV, and negative predictive value (NPV) profile, as shown by the original study\(^1\) at a cutoff value of 50 mm. Also, the study showed that TMHT was more accurate than modified Mallampati score and TMD. Using digital measurements, another source found a lower sensitivity and PPV for TMHT at 50 mm cutoff in a Turkish population.\(^3\) Although the study produced the best compromise for sensitivity and specificity when the cutoff was lowered to 43.5 mm, the authors could not replicate the originally described efficacy of TMHT. The above-mentioned studies compared the efficacy of TMHT with modified Mallampati score, TMD, SMD, and upper lip bite test.

TMHT is probably an objective assessment less likely to be affected by inter-observer variability.\(^1\) No comparative data is available between the TMHT and other commonly employed screening tests such as interincisor gap (IIG), neck circumference (NC), and neck extension. We attempted this study to evaluate and validate the predictive value TMHT at 50 mm in an Indian population in predicting difficult laryngoscopy. In addition, we compared the sensitivity, specificity, PPV, NPV, and diagnostic accuracy of TMHT with other bedside tests such as the modified Mallampati score, IIG, TMD, NC, and neck extension individually in predicting difficult laryngoscopy.

**METHODS**

The study was conducted over a 12-month period between 2015 and 2016 as a prospective, observational, single blind evaluation at a tertiary teaching hospital, and was approved by the Hospital Ethics Committee. After obtaining informed consent, 340 patients, American Society of Anesthesiologists grade I–II, aged 18–65 years presenting for various elective surgeries under general anaesthesia were recruited. Patients with a body mass index >35 kg/m\(^2\), those with neuromuscular disorders, craniofacial abnormalities, abnormal dentition, and unable to sit up were excluded.

Participants were premedicated with oral pantoprazole 40 mg and alprazolam 0.5 mg the night before and 2 hours prior to surgery. Preoperative airway assessments were done by a single researcher. TMHT was performed using a depth caliper and measured between the anterior border of the thyroid cartilage (on the thyroid notch just between the two thyroid laminae) and the anterior border of the mentum (on the mental protuberance of the mandible), with the patient lying supine with a closed mouth. A neutral position of head and neck was maintained with a pillow beneath the head [Figures 1 and 2 (reproduced with permission from Wolters Kluwer)]\(^2\). A digital depth gauge was used to measure the height (INSIZE\(^\text{®}\) Electronic depth gauge, Insize Co Ltd; Suzhou, New District China). A height less than 50 mm was considered to be a predictor of difficult laryngoscopy, which was subsequently assessed clinically by different blinded investigators.

The modified Mallampati score (Samsoon and Young’s modification) was assessed in a sitting posture, with the mouth fully opened, the tongue protruded, and without phonation. The IIG was assessed in a sitting position with the back supported and asking the patient to widely open the mouth. A cutoff less than 4.5 cm was considered as a predictor of difficult laryngoscopy.\(^4\) TMD was measured as a straight line using a ruler between the upper border of the thyroid cartilage and the bony point of mentum, a measurement of <6.5 cm was considered to be a predictor of difficult laryngoscopy.\(^4\) A neck circumference above 37.5 cm at the level of thyroid cartilage was deemed as a predictor of difficult laryngoscopy.\(^5\) Neck extension was measured in the sitting position and facing forward with the shoulder and spine supported. Participants were asked to extend the neck without moving shoulders, with the mouth closed. The angle...
traversed between the external auditory canal to the tip of the nose was measured using a goniometer, and movements less than 90 degrees were considered to be significant.

After applying standard monitoring, anaesthesia was induced with fentanyl 1–2 µg/kg IV, propofol 1.5–2 mg/kg IV, and endotracheal intubation was facilitated with muscle relaxation with atracurium 0.5 mg/kg IV. Neuromuscular blockade was assessed after 3 min of mask ventilation using a single twitch response from a peripheral nerve stimulator. Laryngoscopy was accomplished in a sniffing position using a #3 or #4 Macintosh blade by a group of anaesthetists who had at least 3 years of experience. They were blinded to the TMHT assessment. Laryngoscopic view without backward-upward-rightward pressure maneuver was graded as per the modified Cormack–Lehane (CL) scale from I–IV (Grade I: full view of the vocal cords, Grade IIa: partial view of the vocal cords, Grade IIb; only arytenoids and epiglottis seen, Grade III: only epiglottis visible, Grade IV: neither the epiglottis nor glottis visible). Grade I and IIa were categorized as easy visualization and grade IIb and above were categorized as difficult visualization. All the difficult laryngoscopies were visualized by external laryngeal pressure and intubations done using bougie or McCoy blade. The gradings were noted by the same investigator doing the intubation.

Data of the preoperative bedside screening tests and laryngoscopic visualization were used together to assess and validate the TMHT in predicting difficult laryngoscopy.

The necessary sample size was estimated to be 311 with a 95% confidence interval and 80% power and alpha level of 0.05, assuming the incidence of difficult intubation to be 9% based on previous studies.[1] Taking into consideration an attrition rate of 10%, the final sample size was rounded off to 340. Statistical analysis was performed using SPSS version (Version 19.0, SPSS Inc; Chicago, IL, USA). The sensitivity, specificity, PPV, NPV for each airway assessment tool were calculated, and the data were analysed using the Student’s t-test, Fisher’s exact test, and Yates Chi-square test as appropriate. Statistical significance was defined as $P < 0.05$. The diagnostic accuracy of the bedside tests was compared by measuring the area under the receiver operating characteristic curve (AUC) with 95% confidence interval. Predefined cut-off values were used to formulate the ROC curve. Data are presented as number (percentage) or mean (standard deviation).

RESULTS

A total of 340 patients were included in the study group. Twenty-four patients had to be excluded as they violated the study protocol as they were managed with a supraglottic airway and hence, data from 316 patients (149 male 167 female) were analysed. Three investigators took part in assessing the laryngoscopic grading. The demographic profile and characteristics of preoperative airway assessment tests are depicted in Tables 1 and 2.

The incidence of difficult laryngoscopy was 8.2% (26/316), of which 23 had CL grade IIb and 3 had CL grade III. None of our patients had a grade IV view, and there were no failed intubations. With BURP manoeuvre, 8 of the 26 patients had an improvement in their views. A bougie/stylet was used to facilitate intubation in 23 of these difficult laryngoscopies. A single attempt intubation was accomplished in 303 patients and the remaining 13 required two attempts.

Among all the tests, TMHT had the highest sensitivity (84.62%), specificity (98.97%), and had the most PPV (88%) and NPV (98.63%). This was

| Table 1: Demographic data of the patients |
|-------------------------------------------|
| **Patient characteristics** | **Value: Mean±SD/n (%)** |
| Gender: Male/female | 149±47.1/167 (52.8) |
| Age (years) | 43.4±13.3 |
| Height (cm) | 162.6±5.9 |
| Weight (kg) | 62.0±7.3 |
| BMI (kg/m²) | 23.4 (2.0) |

SD – Standard deviation; BMI – Body mass index
followed by modified Mallampati class and IIG. TMD had the least sensitivity and PPV [Table 3]. The ROC curve [Figure 3] for the airway assessment tools revealed that AUC value for TMHT was higher than all other parameters (0.92), implying near perfect discrimination. Only TMHT, modified Mallampati class, and IIG had AUC value >0.7 on the ROC, implying that these tests are fair. The AUC values of TMD, NC, and neck extension were below 0.5 indicating that these variables performed no better than chance. A statistical significance ($P < 0.05$) was evident for the diagnostic accuracy of TMHT, modified Mallampati class, and IIG.

**DISCUSSION**

Our results suggest that, as a single test, TMHT at a 50 mm cutoff had the highest sensitivity, specificity, PPV, NPV, and was the most accurate when compared with modified Mallampati score, IIG, TMD, NC, and neck extension. TMHT may have the best ability to predict difficult laryngoscopy as it had the highest PPV (88%). Modified Mallampati score was close to TMHT in terms of sensitivity and specificity which was followed by IIG. Although modified Mallampati score and IIG were found to have a statistically significant diagnostic accuracy, with AUC >0.7, they were just adequate but not as reliable as the TMHT. TMD, NC, and neck extension were found to be poorly reliable as their AUC values were below 0.5.

The TMHT observations from our study are comparable to the original data from an Iranian population (314 patients), and subsequent data from an Indian population (345 patients) utilising similar cutoffs at 50 mm.$^{[1,2]}$ The sensitivity (84.6%), specificity (98.9%), PPV (88%), and NPV (98.6%) of TMHT from our study is comparable to the original Iranian study that produced the corresponding values as 82.6%, 99.35, 90.45, and 98.6% respectively.$^{[1]}$ The high PPV (88%) along with an accuracy close to 100% (98.63%) from our data might indicate that a high proportion of difficult laryngoscopy could be predicted when the

**Table 2: Data of the preoperative airway assessment parameters and the modified Cormack-Lehane grades**

| Bed side tests                | Value: Mean±SD/n (%) |
|------------------------------|-----------------------|
| Thyromental height (cm)      | 5.3±0.3               |
| Intercisor gap (cm)          | 5.1±0.2               |
| Thyromental distance (cm)    | 6.3±0.3               |
| Neck circumference (cm)      | 37.1±0.9              |
| Neck extension (°)           | 86.8±4.9              |
| Mallampati Class I           | 190 (60.1)            |
| Mallampati Class II          | 111 (35)              |
| Mallampati Class III         | 15 (4.7)              |
| Mallampati Class IV          | 0                     |
| Modified CL grades           |                       |
| CL grade I                   | 215 (68.03)           |
| CL grade II                  | 75 (23.7)             |
| CL grade IIb                 | 23 (7.2)              |
| CL grade III                 | 3 (0.9)               |
| CL grade IV                  | 0                     |
| TMD                          |                       |

**Table 3: Diagnostic validity profiles and measures of diagnostic accuracy of airway assessment tests for predicting DL (n=316)**

| Airway assessment test       | Sensitivity (%) | Specificity (%) | PPV (%) | NPV (%) | Accuracy (%) | AUC         | P       |
|------------------------------|-----------------|-----------------|---------|---------|--------------|-------------|---------|
| TMHT                         | 84.62 (65.13-95.64) | 98.97 (97.01-99.79) | 88.00 (68.78-97.45) | 98.63 (96.52-99.62) | 97.7 | 0.92 (0.83-1.00) | 0.000* |
| Modified Mallampati Score    | 73.08%           | 81.03 (76.04-85.38) | 25.68 (16.22-37.16) | 97.11 (94.13-98.33) | 80.3 | 0.77 (0.67-0.87) | 0.000* |
| IIG                          | 69.23 (48.21-85.67) | 71.03 (65.44-76.19) | 17.65 (10.81-26.45) | 96.26 (92.77-98.37) | 70.8 | 0.70 (0.59-0.80) | 0.000* |
| Neck extension               | 11.54 (2.45-30.15) | 74.48 (69.06-79.40) | 3.90 (0.81-10.97) | 90.38 (85.91-93.80) | 69.3 | 0.43 (0.33-0.53) | 0.238  |
| Neck circumference           | 65.38 (44.33-82.79) | 32.41 (27.06-38.13) | 7.98 (4.72-12.47) | 91.26 (84.06-95.93) | 35.1 | 0.49 (0.37-0.60) | 0.852  |

Data presented as values (95% CI for sensitivity; specificity; PPV; NPV; AUC – Area under the curve indicating the diagnostic accuracy; *P<0.05. PPV – Positive predictive value; NPV – Negative predictive value; CI – Confidence interval; TMHT – Thyromental height test; IIG – Intercisor gap; TMD – Thyromental distance; DL – Difficult laryngoscopy

Figure 3: Receiver operating characteristic curve comparison of the airway assessment tests
TMH is below 50 mm. Unsurprisingly, the reported mean TMH was similar (close to 60 mm) across the aforementioned studies: 63 mm from us, 57 mm from the other Indian study (raw data from the authors), and 59 mm from the original Iranian study. Likewise, the populations had similar height profile: 162 cm reported by us and the other Indian study, and 166 cm in the Iranian population. Interestingly, a recent study re-evaluated TMH in a Turkish population and derived a best compromise at 43 mm describing a sensitivity, specificity, and NPV of 64.8%, 78%, and 96.1%, respectively.[3] The study also produced a low PPV of 20.87% and failed to show the efficacy of TMHT, as revealed by us and others.[1,2] It has to be noted, however, that their mean TMH was only 45 mm, roughly 15 mm less noted by other studies. The measurement was done in a similar head and neck position reported by other studies. Nonetheless, the measurements were done digitally by a trained technician unlike the other reports where it was done manually by anaesthetists.

Though the sensitivity and specificity of modified Mallampati score in our study (73% and 81%) are comparable with that of an earlier metaanalysis by Shiga et al.[4] (49% and 86%, respectively), recent data shows that modified Mallampati score is a poor predictor of difficult intubation.[7-9] Mallampati assessment is susceptible for incorrect evaluation and gross interobserver variability.[10] The classification is prone to error with phonation which usually occurs involuntarily resulting in poor differentiation between various grades.[11]

At a 4.5-cm cutoff, our sensitivity and specificity of IIG (69% and 71%) are reasonably comparable to that of the previously published data (42% and 97%, respectively).[12]

TMD (6.5 cm) and neck extension (90°) had the least sensitivity and PPV, implying that both these predictors cannot be used individually for predicting difficult laryngoscopy. TMD has been the most questioned of all the bedside tests.[13] A meta-analysis comprising 35 studies including 50,760 patients[8] concluded that the diagnostic value of TMD was unsatisfactory due to wide range in sensitivity, possibly due to different cutoff points (4–7 cm). Although cervical spine movements greater than 90 degrees has been contemplated warranting easy intubation,[14] gross interobserver variability[15] and access to goniometer limit its clinical utility.

The usefulness of NC in predicting difficult laryngoscopy/intubation is dependent on the BMI of the population assessed with most studies evaluating patients with BMI >35 kg/m². In patients with BMI >40 kg/m², at an NC of 40 cm, there is a 5% probability of difficult intubation; increasing up to 35% at an NC of 60 cm.[16] Gender-specific differences and the variable amount of pretracheal soft tissue limit the usefulness of NC unless ultrasound quantification of soft tissue is done.[17] There has been recent interest in the literature exploring the role of ultrasound in predicting difficult airway.[18,19]

A large sample size and an attempt to re-evaluate TMHT at a known cutoff value are some of the strengths of our study. Nonetheless, TMHT may have a role in physically and mentally disabled patients, as well as in those who cannot cooperate for other tests such as the modified Mallampati score and upper lip bite test. Our study had few limitations. The results from our specific ethnic group should be extrapolated cautiously across population with different morphological features. Airway assessment tools were assessed as single tests. Nonetheless, a recent systematic review highlighted that combination of tests has limited value.[20] An observer bias in the subjective measurements and an interobserver variability of the laryngoscopic grading are likely. The limitations of TMHT are also worth mentioning. The requirements of calipers and tapes to conduct the test is a constraint.

Future studies should try and define ethnicity-specific cutoffs and validate them as there are variations in craniofacial features and body habitus across different races.[21,22] TMHT data on caucasian population is lacking. The ideal way of measuring the airway parameters should be defined and pilot testing for interobserver variability should be strongly encouraged before embarking clinical evaluation. It is worth reiterating that clinicians should move away from the notion of focusing on a single airway assessment tool in predicting difficult intubation. As echoed in the All India Difficult Airway Association guidelines,[23] safer management of patients including recognising those at risk should be a priority rather than relying on a single bedside airway assessment parameter.

**CONCLUSION**

TMHT was the most sensitive and accurate test in predicting difficult laryngoscopy when compared against modified Mallampati score, IIG, TMD, NC, and
neck extension. TMHT appears promising as a single anatomical measure to predict the risk of difficult laryngoscopy, however, validation will require further studies in more diverse patient populations.

**Financial support and sponsorship**
Nil.

**Conflicts of interest**
There are no conflicts of interest.

**REFERENCES**

1. Rezadi F, Ahangari A, Shokri H, Najafi A, Khajavi MR, Daghigh M, et al. Thyromental height: A new clinical test for prediction of difficult laryngoscopy. Anesth Analg 2013;117:1347-51.

2. 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.

3. Selvi O, Kahraman T, Senturk O, Tulgar S, Serifsoy E, Ozer Z. 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.

4. Hosain Khan Z, Mohammadi M, Eghtesadi P. Upper lip bite test with thyromental, sternomental and interincisor distance in predicting difficult intubation: A comparative study. Tehran Univ J Med Sci 2006:64:102-10.

5. Liaskou C, Vouzounerakis E, Moirasgenti M, Trikoupi A, Staikou C. Anatomic features of the neck as predictive markers of difficult direct laryngoscopy in men and women: A prospective study. Indian J Anaesth 2014;58:176-82.

6. Shiga T, Wajima Z, Inoue T, Sakamoto A. Predicting difficult intubation in apparently normal patients: A meta-analysis of bedside screening test performance. Anaesthesiology 2005;103:429-37.

7. Lee A, Fan LT, Gin T, Karmakar MK, Ngor Kee WD. A systematic review (meta-analysis) of the accuracy of the Mallampati test in predicting difficult airway. Anesth Analg 2006;102:1867-78.

8. Krohnbaan B, Divyapake 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.

9. 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.

10. Eberhart LH, Arndt C, Cierpka T, Schwaneckamp J, Wulf H, Putze C, et al. 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.

11. Chohedri AH, Allahari E, Tabari M. The upper lip bite test: Prediction of difficult endotracheal intubation. Prof Med J 2005;12:440-5.

12. Arné J, Descoins P, Fusiardi J, Ingrand P, Ferrier B, Boudigues D, 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.

13. Randell T. Prediction of difficult intubation. Acta Anaesthesiol Scand 1996;40:1016-23.

14. Wilson ME, Spiegelhalter D, Robertson JA, Lesser P. Predicting difficult intubation. Br J Anaesth 1988;61:211-6.

15. Ukami Y, Takenaka I, Nakamura M, Fukuyama H, Aoyama K, Kadoya T. The reliability of the bellhouse test for evaluating extension capacity of the occipitoatlantoaxial complex. Anesth Analg 2002;95:1437-41.

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

17. Ezri T, Gewurtz G, Sessler DI, Medalion B, Szmuk P, Hagberg C, et al. Prediction of difficult laryngoscopy in obese patients by ultrasound quantification of anterior neck soft tissue. Anaesthesia 2003;58:1111-4.

18. Caballero P, Alvarez-Sala R, García-Río F, Prados C, Hernán MA, Villamor J, et al. CT in the evaluation of the upper airway in healthy subjects and in patients with obstructive sleep apnea syndrome. Chest 1998;113:111-6.

19. Abe T, Kawakami Y, Sugita M, Yoshikawa K, Fukunaga T. Use of B-mode ultrasound for visceral fat mass evaluation: Comparisons with magnetic resonance imaging. Appl Human Sci 1995;14:133-9.

20. Vannucci A, Cavallone LF. Bedside predictors of difficult intubation: A systematic review. Minerva Anestesiol 2016;82:69-83.

21. Farkas LG, Katic MJ, Forrest CR, Alt KW, Bagic I, Baltadjiev G, et al. International anthropometric study of facial morphology in various ethnic groups/races. J Craniofac Surg 2005;16:615-46.

22. Balakrishnan KP, Chockalingam PA. Ethnicity and upper airway measurements: A study in South Indian population. Indian J Anaesth 2017;61:622-6.

23. Myatra SN, Shah A, Kundra P, Patwa A, Ramkumar V, Divatia JV, et al. All India Difficult Airway Association 2016 guidelines for the management of unanticipated difficult tracheal intubation in adults. Indian J Anaesth 2016;60:885-98.