

A Prospective Observational Single-Blind Study Comparing the Upper Lip Bite Test and the Modified Mallampati Test for Predicting Difficult Laryngoscopy in Adult Patients Undergoing General Anaesthesia with Endotracheal Intubation

Uddalak Chattopadhyay¹, Subrata Bisai², Syed Sadaqat Hussain³

¹Associate Professor, Department of Anaesthesiology, HBCH and MPMCC, Varanasi, Uttar Pradesh, India

²Assistant Professor, Department of Anaesthesiology, Raiganj Government Medical College and Hospital, West Bengal, India

³Associate Professor, Department of Anaesthesiology, HBCH and MPMCC, Varanasi, Uttar Pradesh, India

Received: 18-09-2025 / Revised: 17-10-2025 / Accepted: 18-11-2025

Corresponding Author: Dr. Uddalak Chattopadhyay

Conflict of interest: Nil

Abstract:

Background: Unanticipated difficult laryngoscopic tracheal intubation remains a major cause of anaesthesia-related morbidity and mortality. Preoperative airway assessment tools, such as the Modified Mallampati Test (MMT) and the Upper Lip Bite Test (ULBT), are used to predict difficult laryngoscopy, but their diagnostic accuracy varies.

Methods: This prospective, single-blinded observational study included 300 adult patients (ASA I–II) undergoing elective surgery under general anaesthesia with endotracheal intubation at Burdwan Medical College and Hospital. Preoperative airway assessment was performed using MMT and ULBT. Direct laryngoscopy was conducted under standard anaesthetic conditions and graded according to the Cormack-Lehane classification. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy were calculated for both tests.

Results: Of the 300 patients, 52 (17.3%) exhibited difficult laryngoscopic views (grades III–IV). ULBT demonstrated superior diagnostic performance, with sensitivity of 88.46%, specificity of 92.74%, PPV of 71.87%, NPV of 97.45%, and accuracy of 92.00%. In contrast, MMT showed a sensitivity of 19.23%, specificity of 91.93%, PPV of 33.33%, NPV of 84.44%, and accuracy of 79.33%.

Conclusion: The Upper Lip Bite Test showed significantly higher sensitivity and overall diagnostic accuracy than the Modified Mallampati Test for predicting difficult laryngoscopy. ULBT is a more reliable and practical bedside tool for preoperative airway assessment and may enhance patient safety when integrated into routine anaesthetic evaluation.

Keywords: Airway assessment; Difficult laryngoscopy; Upper Lip Bite Test; Modified Mallampati Test; Intubation prediction; Preoperative evaluation; Anaesthesia safety; Cormack-Lehane classification.

This is an Open Access article that uses a funding model which does not charge readers or their institutions for access and distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>) and the Budapest Open Access Initiative (<http://www.budapestopenaccessinitiative.org/read>), which permit unrestricted use, distribution, and reproduction in any medium, provided original work is properly credited.

Introduction

Unanticipated difficult laryngoscopic tracheal intubation remains a major concern for anaesthesiologists, being a leading cause of perioperative complications and adverse outcomes.[1] Difficult or failed intubation contributes significantly to anaesthesia-related mortality and permanent neurological injury.[2] It is classically defined as restricted glottic visualization during laryngoscopy, according to the Cormack and Lehane classification.[3]

The reported incidence of difficult laryngoscopy ranges from 1.5% to 13%, with failed intubation occurring in 0.05%–0.35% of cases and “cannot

ventilate, cannot intubate” situations in 0.01%–0.02%.[4,5] Owing to the serious consequences of failed intubation, pre-anaesthetic airway assessment is considered an essential standard of care.[4] Airway loss after induction remains a major cause of anaesthesia-related morbidity and mortality.[6,7] About 85% of airway management errors result in permanent cerebral injury, and up to 30% of anaesthetic deaths are linked to difficult airway management.[8] Moreover, 17% of malpractice claims involving difficult intubation lacked documented preoperative airway evaluation.[9]

Accurate prediction of difficult intubation allows anaesthesiologists to plan appropriate management strategies.[10] Airway difficulty depends largely on oropharyngeal anatomy and head-neck mobility,[11] commonly assessed through atlanto-occipital extension, thyromental distance, and the Modified Mallampati Classification.[11] Although numerous predictors exist, their reliability remains inconsistent.[5,12] Because difficult laryngoscopy often parallels difficult intubation,[13] a simple, accurate, and reproducible bedside test is essential.

Among existing methods, the Mallampati and Modified Mallampati Tests (MMT) are the most widely used.[14-16] However, large studies show only moderate accuracy.[17-20] In 2002, Khan et al. introduced the Upper Lip Bite Test (ULBT) as a simple alternative assessing mandibular mobility and dental architecture.[1] Unlike the MMT, which estimates tongue size and mouth opening,[21] ULBT evaluates mandibular range of movement.[1]

This study therefore aims to compare the predictive performance of the ULBT and MMT in anticipating difficult laryngoscopy in adults undergoing elective surgery under general anaesthesia, by assessing sensitivity, specificity, positive and negative predictive values, and overall accuracy using the Cormack and Lehane grading system as the reference standard.

Materials and Methods

This prospective, single-blinded observational study was conducted in the Department of Anaesthesiology at Burdwan Medical College and Hospital, India, over a 12-month period from January to December 2013, following approval from the Institutional Ethics Committee. Written informed consent was obtained from all participants prior to enrolment. A total of 300 adult patients, aged over 16 years, classified as American Society of Anesthesiologists (ASA) physical status I or II, and scheduled for elective surgery under general anaesthesia requiring endotracheal intubation were recruited. Exclusion criteria included patients with a history of difficult intubation, craniofacial abnormalities, airway or cervical spine pathology, edentulism, limited mouth opening, or those requiring rapid sequence induction or awake intubation. Sample size was determined based on an expected 8% difference in diagnostic accuracy between the Upper Lip Bite Test (ULBT) and the Modified Mallampati Test (MMT), with a power of 80% and significance level of 5%, yielding a target of 280 patients; 300 were included to account for potential exclusions.

Airway assessment was performed preoperatively using both ULBT and MMT by a trained anaesthesiology resident not involved in intraoperative management. MMT was conducted with the patient seated, mouth fully open and tongue

protruded without phonation, and graded I–IV, with Grades I and II considered predictive of easy laryngoscopy, and Grades III and IV indicative of difficulty. ULBT involved asking patients to bite their upper lip with the lower incisors while seated, and was classified as Class I (biting above vermilion line), Class II (biting below vermilion line), and Class III (inability to bite); Class I and II were considered predictive of easy laryngoscopy, and Class III as difficult. All laryngoscopies were performed in the sniffing position using a Macintosh blade (size 3 or 4) by an experienced anaesthesiologist blinded to the preoperative test results, following standard induction with fentanyl (2 µg/kg), lidocaine (1 mg/kg), propofol (2 mg/kg), and atracurium (0.5 mg/kg). Laryngoscopic views were graded using the Cormack and Lehane (C & L) classification, with Grades I and II considered easy, and Grades III and IV considered difficult. No external laryngeal manipulation was applied during grading.

The primary outcome was the diagnostic performance of ULBT and MMT in predicting difficult laryngoscopy, measured using sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and overall accuracy. Standard formulas were used to compute each metric: sensitivity = $TP / (TP + FN)$; specificity = $TN / (TN + FP)$; PPV = $TP / (TP + FP)$; NPV = $TN / (TN + FN)$; and accuracy = $(TP + TN) / \text{Total}$. Data were analyzed using Microsoft Excel and SPSS software, and results were expressed as mean $\pm$ standard deviation for continuous variables, and percentages for categorical variables. A p-value <0.05 was considered statistically significant.

Results

A total of 300 adult patients undergoing elective surgery under general anaesthesia were included in the final analysis (Table 1). The study cohort comprised 120 males (40%) and 180 females (60%), with a mean age of 28.00 ± 3.28 years, mean height of 154.60 ± 5.47 cm, and mean weight of 61.50 ± 3.93 kg. Most participants (240; 80%) were classified as ASA physical status I, while 60 (20%) were ASA II. A graphical summary of the demographic distribution is presented in Figure 1.

During direct laryngoscopy performed under standardized anaesthetic conditions, the Cormack–Lehane grading system identified 248 patients (82.7%) with easy laryngoscopic views (Grade I: 130 [43.3%], Grade II: 118 [39.3%]) and 52 patients (17.3%) with difficult views (Grade III: 44 [14.7%], Grade IV: 8 [2.7%]) (Table 2).

Preoperative airway evaluation using the Modified Mallampati Test (MMT) classified 270 patients (90%) as having an easy airway (Grade I: 75 [25%]; Grade II: 195 [65%]) and 30 patients (10%) as potentially difficult (Grade III: 25 [8.3%]; Grade IV:

5 [1.7%]) (Table 2). In contrast, the Upper Lip Bite Test (ULBT) identified 236 patients (78.7%) as easy (Class I: 116 [38.7%]; Class II: 120 [40.0%]) and 64 patients (21.3%) as difficult (Class III) (Table 2, Figure 2).

When correlated with the actual laryngoscopic findings, ULBT demonstrated significantly better diagnostic performance compared with MMT (Tables 3–5). Specifically, ULBT achieved a sensitivity of 88.46%, specificity 92.74%, positive predictive value (PPV) 71.87%, negative predictive value (NPV) 97.45%, and overall diagnostic accuracy 92.00%. In contrast, MMT yielded a sensitivity of 19.23%, specificity 91.93%, PPV 33.33%, NPV 84.44%, and overall accuracy 79.33%. Comparative visualizations of diagnostic performance between the two tests are illustrated in Figure 3, while Figure 4 depicts confusion matrix heatmaps highlighting ULBT's superior discrimination of true positives and lower false-negative rates.

Notably, ULBT produced substantially fewer false negatives ($n = 6$) compared to MMT ($n = 42$), indicating superior capability in correctly identifying patients with difficult laryngoscopy. No cases of failed intubation or perioperative airway-related complications were reported in any participant.

Receiver Operating Characteristic (ROC) curve analysis (Figure 5) further substantiated the diagnostic advantage of ULBT, demonstrating an area under the curve (AUC) of 0.96, consistent with excellent discriminatory power, whereas the MMT showed a lower AUC of 0.67, reflecting poor sensitivity and limited predictive accuracy.

Overall, the ULBT significantly outperformed the MMT in all major diagnostic indices-sensitivity, specificity, PPV, NPV, and accuracy-confirming its superior reliability and clinical utility as a preoperative screening tool for predicting difficult laryngoscopy in adult patients undergoing elective surgery under general anaesthesia.

Table 1: Demographic data of the study patient

ASA Physical status- ASA PS I: number (%)	240 (80%)
ASA PS II: number (%)	60 (20%)
Sex- Male: number (%)	120 (40%)
Female: number (%)	180 (60%)
Age in years (Mean $\pm$ Standard deviation)	28.00 $\pm$ 3.28
Height in cm (Mean $\pm$ Standard deviation)	154.60 $\pm$ 5.47
Weight in kg (Mean $\pm$ Standard deviation)	61.50 $\pm$ 3.93

Table 2: Frequencies of airway assessment classifications

Grade	ULBT	MMT	C & L
I	116 (38.7%)	75 (25%)	130 (43.3%)
II	120 (40.0%)	195 (65%)	118 (39.3%)
III	64 (21.3%)	25 (8.3%)	44 (14.7%)
IV	NA	5 (1.7%)	8 (2.7%)

Table 3: Comparison of Gradings between ULBT and Laryngoscopic view

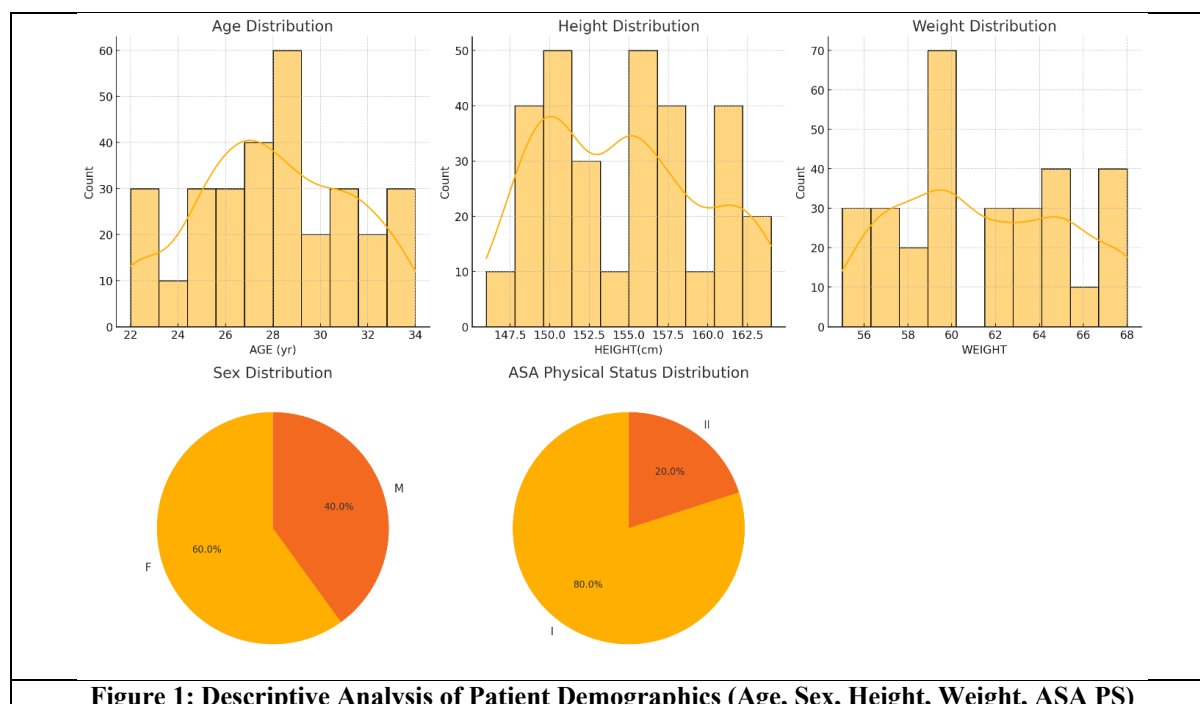
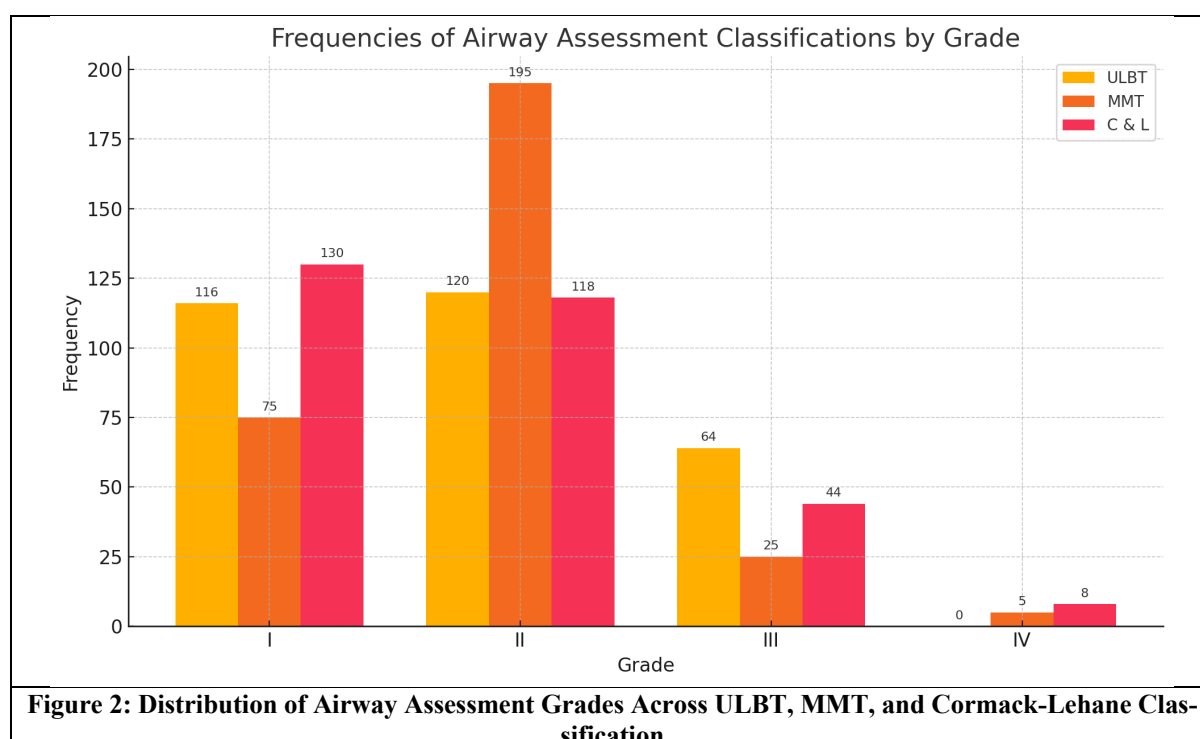
ULBT	Laryngoscopic Grade: Easy (I, II)	Laryngoscopic Grade: Difficult (III, IV)	Total
ULBT Easy (I, II)	230 (a) TN	06 (b) FN	236 (a+b), [78.7%]
ULBT Difficult (III, IV)	18 (c) FP	46 (d) TP	64 (c+d), [21.3 %]
Total	248(a+c), [82.7%]	52(b+d), [17.3%]	300 (a+b+c+d)

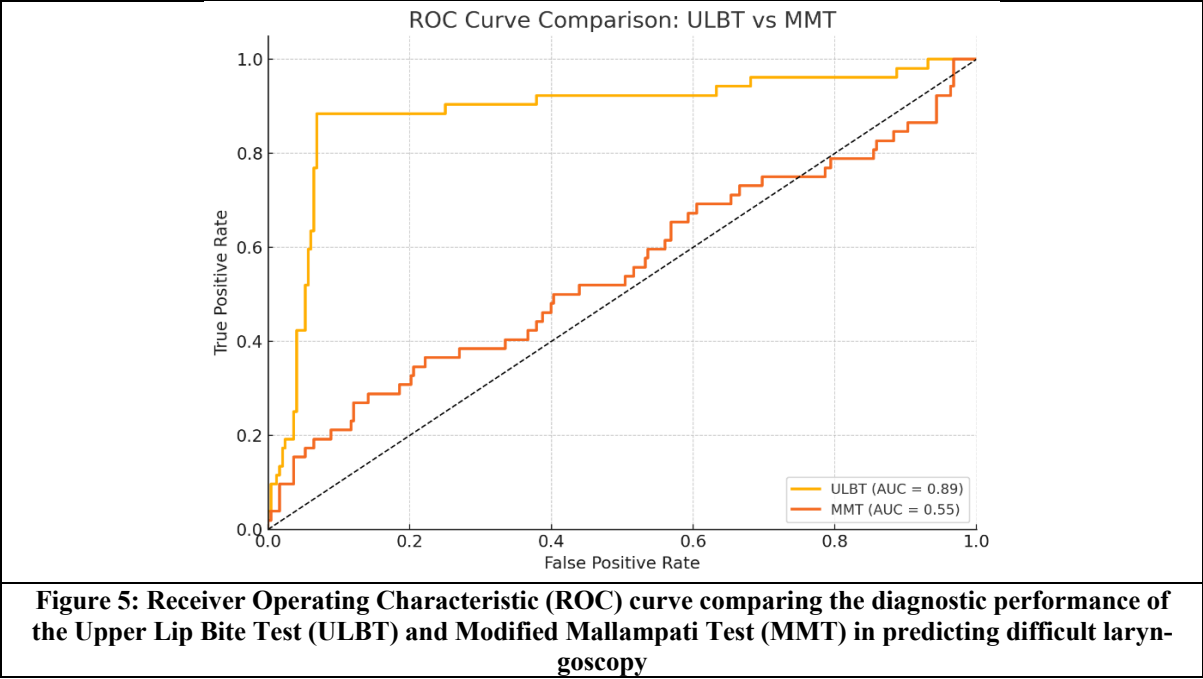
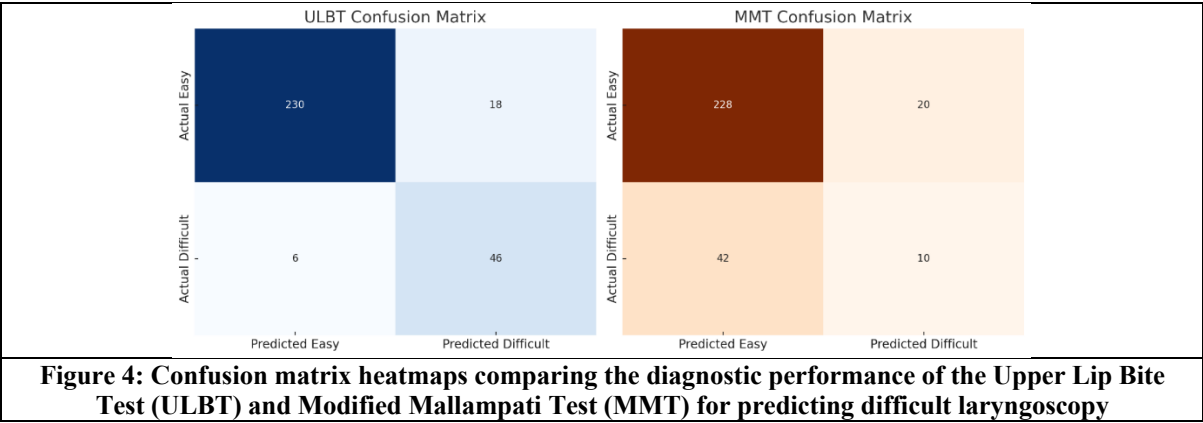
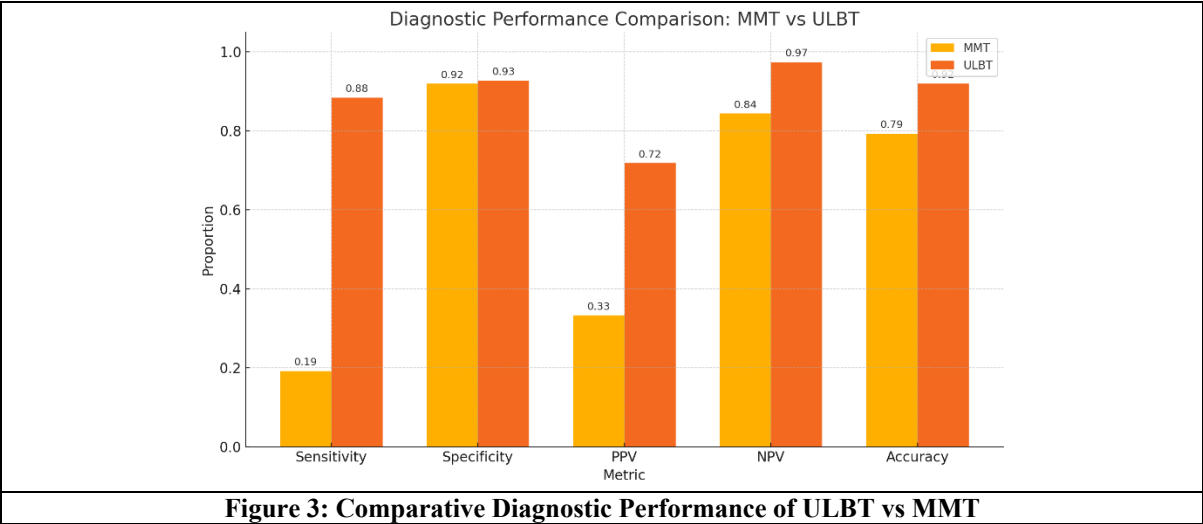
Table 4: Comparison of Gradings between MMT and Laryngoscopic view

MMT	Laryngoscopic Grade: Easy (I, II)	Laryngoscopic Grade: Difficult (III, IV)	Total
MMT grade: Easy (I, II)	228 (a) TN	42 (b) FN	270 (a+b), [90%]
MMT grade: Difficult (III, IV)	20 (c) FP	10 (d) TP	30 (c+d), [10%]
Total	248(a+c), [82.7%]	52(b+d), [17.3%]	300 (a+b+c+d)

Table 5: Comparison of analytical results between ULBT and MMT

Parameters	ULBT (n=300)	MMT (n=300)
True Positive	46	10
False Positive	18	20
True Negative	230	228
False Negative	06	42
Sensitivity (%)	88.46	19.23
Specificity (%)	92.74	91.93
Positive predictive value (%)	71.87	33.33
Negative predictive value (%)	97.45	84.44
Accuracy (%)	92.00	79.33

**Figure 1: Descriptive Analysis of Patient Demographics (Age, Sex, Height, Weight, ASA PS)****Figure 2: Distribution of Airway Assessment Grades Across ULBT, MMT, and Cormack-Lehane Classification**



Discussion

Failure of the anesthesiologist to maintain a patent airway after induction of general anesthesia remains

one of the most significant causes of anesthesia-related morbidity and mortality. [22-24] Owing to the potentially serious complications associated with

failed tracheal intubation-such as hypoxemic cardiopulmonary arrest-considerable emphasis has been placed on identifying patients at risk for difficult laryngoscopy and intubation. A reliable screening test must be both accurate and reproducible, providing high sensitivity and specificity to minimize false predictions and prevent life-threatening outcomes. However, no single assessment tool achieves complete sensitivity or specificity.[12] Numerous methods, including the Mallampati test, thyromental distance, inter-incisor gap, mandibular length, chin protrusion, atlanto-occipital extension, and the Upper Lip Bite Test (ULBT), have been proposed, yet all have inherent limitations. Consequently, unanticipated difficult laryngoscopy continues to pose a clinical challenge, although combining multiple predictors may enhance diagnostic precision.

Difficult or failed tracheal intubation remains a major cause of death and permanent neurological injury during anesthesia.[25] The reported incidence of difficult laryngoscopy and intubation varies from 1.5% to 13% among patients undergoing general anesthesia,[26-29] reflecting differences in study methodologies and reference standards such as Cormack and Lehane grades, number of laryngoscopic attempts,[30] and use of the BURP maneuver.[1] Other factors, including cricoid pressure, head position, laryngoscope blade type, and operator experience, further influence visualization. In the present study, all intubations were performed by a single experienced anesthesiologist under standardized conditions, resulting in an incidence of 17.3%, which is slightly higher than previously reported values. [26,28]

The Mallampati Test (MT), introduced in 1985,[21] remains one of the most commonly used bedside predictors of difficult intubation. Despite its widespread use, its diagnostic accuracy has been debated. [12,20,31,32] A meta-analysis of 42 studies encompassing 34,513 patients by Lee et al. [31] demonstrated poor-to-moderate predictive performance. Although the Modified Mallampati Test (MMT) continues to be regarded as the “gold standard” in airway assessment,[16] its reliability has frequently been questioned. Mallampati et al. [21] originally established a correlation between visibility of oropharyngeal structures and ease of intubation, but inter-rater reliability was not reported. Variability in technique, patient cooperation, and interpretation contribute to inconsistent results.[12] Oates et al. [7] noted significant inter-observer variation, showing that phonation and incomplete mouth opening can alter grading. Other studies have similarly described the MMT as overly subjective, with high false-positive rates and limited specificity. [33,34]

Wilson et al. [6] identified five clinical predictors of difficult intubation-weight, head and neck mobility,

jaw movement, receding mandible, and prominent incisors. The MMT evaluates relative tongue size in relation to the oral cavity,[21] whereas the ULBT assesses mandibular subluxation and dentition simultaneously, thereby improving objectivity and reproducibility.[1] The ULBT's three clearly defined classes make it less susceptible to inter-observer variability, allowing for a simple and standardized bedside assessment.[1]

In this study, the diagnostic accuracy of ULBT and MMT was compared using the Cormack and Lehane grading as the reference standard. The ULBT demonstrated superior overall accuracy (92.0%), sensitivity (88.46%), positive predictive value (PPV = 71.87%), and negative predictive value (NPV = 97.45%) compared with the MMT, though both tests showed similar specificity. These results are consistent with prior studies reporting higher accuracy for ULBT (88.7% vs 66.8%).[1] The findings of the present study also align with previously published results, though sensitivity and PPV were slightly higher-likely due to reduced inter-observer bias and ethnic craniofacial variations. [35-37]

In contrast, the MMT showed moderate accuracy (79.33%) and specificity (91.93%) but notably low sensitivity (19.23%), corroborating earlier studies that reported variable predictive ability. The reduced sensitivity in this cohort may be attributed to ethnic differences and uniform scoring by a single observer, as MMT is known for poor inter-observer reliability. [12,25]

The strengths of this study include its standardized methodology, with all airway assessments and laryngoscopies performed by the same investigator, minimizing inter-observer variability and enhancing consistency. However, several limitations should be acknowledged. Being a single-centre study, the findings may not be generalizable. Patient cooperation and comprehension-particularly for the MMT-can influence outcomes, though proper demonstration of the procedure may mitigate these issues. Additionally, the ULBT cannot be applied in certain populations, such as edentulous individuals, patients with restricted mouth opening, or those unable to follow instructions.

Conclusion

This study demonstrates that the Upper Lip Bite Test has significantly higher diagnostic accuracy, sensitivity, positive predictive value, and negative predictive value than the Modified Mallampati Test for predicting difficult laryngoscopy, although both tests showed similar specificity. Given its superior sensitivity and ease of administration, the ULBT appears to be a more effective and reliable tool for preoperative airway assessment. Its simplicity, objectivity, and quick bedside applicability make it a valuable option, either alone or in combination

with other predictive tests. However, its use is limited in specific patient groups, and these constraints must be considered in clinical practice. Overall, the ULBT offers a practical, reproducible method for improving the safety of airway management in anaesthetic practice.

Funding

This study received no external funding.

Ethical Approval

This study was approved by the Institutional Ethics Committee Burdwan Medical College (BMC/PG/307 dated 24.01.2013 and conducted in accordance with the Declaration of Helsinki (2013).

References

1. 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(2):595-9.
2. Ali MA, Qamar-ul-Hoda M, Samad K. Comparison of upper lip bite test with Mallampati test in the prediction of difficult intubation at a tertiary care hospital of Pakistan. *Journal of Pakistan Medical Association* 2012;62(10):1012.
3. Randell T. Prediction of difficult intubation. *Acta Anaesthesiol Scand* 1996;40(8P2):1016-23.
4. Koch BE. The evolution of nurse anesthesia in the United States. In: *The Wondrous Story of Anesthesia*. New York, NY: Springer New York 2014:271-92.
5. Crosby ET, Cooper RM, Douglas MJ, et al. The unanticipated difficult airway with recommendations for management. *Canadian Journal of Anaesthesia* 1998;45(8):757-76.
6. Wilson ME, Spiegelhalter D, Robertson JA, Lesser P. Predicting difficult intubation. *Br J Anaesth* 1988;61(2):211-6.
7. Oates JD, Macleod AD, Oates PD, et al. Comparison of two methods for predicting difficult intubation. *Br J Anaesth* 1991;66(3):305-9.
8. Benumof JL, Scheller MS. The importance of transtracheal jet ventilation in the management of the difficult airway. *Anesthesiology* 1989;71(5):769-78.
9. Larson SL, Jordan L. Preventable adverse patient outcomes: a closed claims analysis of respiratory incidents. *AANA Journal* 2001;69(5).
10. Enterlein G, Byhahn C. Practice guidelines for management of the difficult airway. *Der Anaesthesist* 2013;62(10):832-5.
11. Hester CE, Dietrich SA, White SW, et al. A comparison of preoperative airway assessment techniques: the modified Mallampati and the upper lip bite test. *AANA Journal* 2007;75(3).
12. Karkouti K, Rose DK, Ferris LE, et al. Inter-observer reliability of ten tests used for predicting difficult tracheal intubation. *Can J Anaesth* 1996;43(6):554-9.
13. Benumof JL. Difficult laryngoscopy: obtaining the best view. *Can J Anaesth* 1994;41(5):361-5.
14. Samsoon GL, Young JR. Difficult tracheal intubation: a retrospective study. *Anaesthesia* 1987;42(5):487-90.
15. Khan RM. *Airway Management – Manual for Clinical Practitioners and Examinees* 2011;4:21.
16. Isono S. Common practice and concepts in anesthesia: time for reassessment. *Anesthesiology* 2001;95(4):825-7.
17. El-Ganzouri AR, McCarthy RJ, Tuman KJ, et al. Preoperative airway assessment: predictive value of a multivariate risk index. *Anesth Analg* 1996;82(6):1197-204.
18. Rocke DA, Murray WB, Rout CC, et al. Relative risk analysis of factors associated with difficult intubation in obstetric anesthesia. *Anesthesiology* 1992;77(1):67-73.
19. Yamamoto K, Tsubokawa T, Shibata K, et al. Predicting difficult intubation with indirect laryngoscopy. *Anesthesiology* 1997;86(2):316-21.
20. Cattano D, Panicucci E, Paolicchi A, et al. Risk factors assessment of the difficult airway: an Italian survey of 1956 patients. *Anesth Analg* 2004;99(6):1774-9.
21. Mallampati SR, Gatt SP, Gugino LD, et al. A clinical sign to predict difficult tracheal intubation; a prospective study. *Canadian Anaesthetists' Society Journal* 1985;32(4):429-34.
22. Merah NA, Foulkes-Crabbe DJ, Kushimo OT, et al. Prediction of difficult laryngoscopy in a population of Nigerian obstetric patients. *West African Journal of Medicine* 2004;23(1):38-41.
23. Kuczkowski KM, Reisner LS, Benumof JL. Airway problems and new solutions for the obstetric patient. *J Clin Anesth* 2003;15(7):552-63.
24. Davies JM, Weeks S, Crone LA, et al. Difficult intubation in the parturient. *Can J Anaesth* 1989;36(6):668-74.
25. Hilditch WG, Kopka A, Crawford JM, et al. Inter observer reliability between a nurse and anaesthetist of tests used for predicting difficult tracheal intubation. *Anaesthesia* 2004;59(9):881-4.
26. Siddiqi R, Kazi WA. Predicting difficult intubation--a comparison between Mallampati classification and Wilson risk-sum. *J Coll Physicians Surg Pak* 2005;15(5):253-6.

27. Hung O, Mills J. Predictions and clinical decisions: a fine balance. *Can J Anesth* 2000;47(8):721-4.
28. Wong SH, Hung CT. Prevalence and prediction of difficult intubation in Chinese women. *Anaesthesia and Intensive Care* 1999;27(1):49-52.
29. Ezri T, Evron S, Hadad H, et al. Tracheostomy and endotracheal intubation: a short history. *Harefuah* 2005;144(12):891-3.
30. Brodsky JB, Lemmens HJ, Brock-Utne JG, et al. Morbid obesity and tracheal intubation. *Anesth Analg* 2002;94(3):732-6.
31. Lee A, Fan LT, Gin T, et al. A systematic review (meta-analysis) of the accuracy of the Mallampati tests to predict the difficult airway. *Anesth Analg* 2006;102(6):1867-78.
32. Shiga T, Wajima ZI, Inoue T, et al. Predicting difficult intubation in apparently normal patients: a meta-analysis of bedside screening test performance. *Anesthesiology* 2005;103(2):429-37.
33. Savva D. Prediction of difficult tracheal intubation. *Br J Anaesth* 1994;73(2):149-53.
34. Frerk CM. Predicting difficult intubation. *Anaesthesia* 1991;46(12):1005-8.
35. Hanihara T, Ishida H, Dodo Y. Characterization of biological diversity through analysis of discrete cranial traits. *Am J Phys Anthropol* 2003;121(3):241-51.
36. Badawi-Fayad J, Cabanis EA. Three-dimensional procrustes analysis of modern human craniofacial form. *Anat Rec (Hoboken)* 2007;290(3):268-76.
37. Perez SI, Bernal V, Gonzalez PN. Differences between sliding semi-landmark methods in geometric morphometrics, with an application to human craniofacial and dental variation. *Journal of Anatomy* 2006;208(6):769-84.