JHWCR JOURNAL OF HEALTH, WELLNESS AND COMMUNITY RESEARCH 3007-0570

Original Article Volume 4, Issue 3

# Comparative Analysis of The Efficacy of Upper Lip Bite Test and Thyromental Distance for Prediction of Difficult Intubation

Muhammad Abdul Sattar[1], Abdul Wadood[1], Fatima Noreen[2], Awais Akhtar[2], Inam Ullah[1]

[1] Department of Health Professional Technologies, Faculty of Allied Health Sciences, The University of Lahore, Lahore, Pakistan

[2] Lecturer, Department of Health Professional Technologies, The University of Lahore, Lahore, Pakistan

* Correspondence: Fatima Noreen: fatima.noreen@dhpt.uol.edu.pk

## ABSTRACT

***Background:*** *Unanticipated difficult intubation remains a major contributor to perioperative morbidity and mortality, underscoring the need for reliable bedside airway assessment tools. The Upper Lip Bite Test (ULBT) and Thyromental Distance (TMD) are commonly used preoperative predictors, yet their comparative diagnostic performance remains clinically debated.* ***Objective:*** *To compare the diagnostic accuracy of ULBT and TMD in predicting difficult intubation among adult patients undergoing elective surgery under general anesthesia.* ***Methods:*** *In this comparative cross-sectional observational study, 119 adult patients scheduled for elective surgery requiring endotracheal intubation were consecutively enrolled. Preoperative airway assessment included ULBT and TMD classification using standardized criteria. Difficult intubation was defined as Cormack–Lehane grade III/IV or intubation requiring more than three attempts or prolonged duration. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), likelihood ratios, and accuracy were calculated with 95% confidence intervals.* ***Results:*** *Difficult intubation occurred in 21 patients (17.7%). ULBT demonstrated higher sensitivity than TMD (66.7% vs 47.6%) with comparable specificity (87.8% vs 89.8%). ULBT showed superior NPV (92.5% vs 88.9%) and lower negative likelihood ratio (0.38 vs 0.58), indicating improved rule-out capability, while PPV was moderate for both tests (53.8% vs 50.0%).* ***Conclusion:*** *ULBT provides greater sensitivity and stronger exclusion performance compared with TMD, supporting its preferential use as a primary bedside screening tool within a multimodal airway assessment strategy.*

***Keywords:*** *Difficult intubation; Airway management; Upper Lip Bite Test; Thyromental Distance; Diagnostic accuracy; Preoperative assessment*

## INTRODUCTION

Securing a patent airway remains the cornerstone of safe anesthetic practice, as failure to establish effective ventilation and oxygenation can rapidly result in hypoxic brain injury or death. Despite advances in airway devices, standardized algorithms, and training, unanticipated difficult intubation continues to contribute substantially to perioperative morbidity and mortality (2,13). Major airway complications account for a significant proportion of anesthesia-related adverse events, and even short periods of desaturation during failed or prolonged laryngoscopy attempts may lead to catastrophic outcomes (2). Consequently, accurate preoperative identification of patients at risk for difficult laryngoscopy or intubation remains a critical clinical priority.

Difficult intubation has been variably defined in the literature, most commonly by the American Society of Anesthesiologists as intubation requiring multiple attempts or prolonged time despite optimal conditions (4). In clinical research, it is frequently operationalized using laryngoscope grading systems, particularly the Cormack–Lehane classification, where grades III or IV indicate limited or absent glottic visualization and are

**Received**: 09 December 2025
**Revised**: 20 January 2026
**Accepted**: 09 February 2026
**Published**: 15 February 2026

Citation: Click to Cite

associated with increased intubation difficulty (4,5). The reported incidence of difficult intubation ranges widely from 1% to 18%, largely reflecting heterogeneity in definitions, patient populations, operator expertise, and assessment criteria (5,14). This variability underscores the methodological necessity of clearly defining the reference standard when evaluating predictive tests and highlights the ongoing need for reliable, reproducible bedside screening tools.

Among the numerous preoperative airway assessment methods, the Upper Lip Bite Test (ULBT) and Thyromental Distance (TMD) are widely used because they are simple, rapid, noninvasive, and require no specialized equipment. The ULBT evaluates mandibular mobility and dental architecture by assessing the patient's ability to bite the upper lip with the lower incisors, thereby integrating functional mandibular protrusion and anatomical constraints that directly influence laryngoscopic alignment (7). In contrast, the TMD measures the distance from the mentum to the thyroid notch with the head extended and mouth closed, reflecting the size of the submandibular space and its capacity to accommodate displacement of the tongue during laryngoscopy (17). While TMD has historically been regarded as a standard component of airway assessment, its predictive performance when used as a single parameter has been questioned, particularly regarding sensitivity and interobserver variability (5,9).

Comparative investigations suggest that ULBT may demonstrate higher specificity and negative predictive value than TMD, particularly in ruling out difficult laryngoscopy in apparently normal patients (1,7). Prospective and systematic evaluations have reported that ULBT correlates more consistently with laryngoscopic view than linear anatomical measurements alone (10–12,15). However, both tests exhibit imperfect sensitivity, and no single bedside assessment has achieved sufficient discriminative accuracy to independently predict all difficult intubations (3,5). Furthermore, variations in cutoff values for TMD, differences in endpoint definitions (difficult laryngoscopy versus difficult intubation), and population-specific craniofacial characteristics limit the generalizability of existing findings (8,18,19). Importantly, much of the available evidence originates from heterogeneous settings, and there remains a need for context-specific data generated within local surgical populations to inform clinical decision-making and improve perioperative risk stratification.

From a PICO perspective, the population of interest comprises adult patients undergoing elective surgery under general anesthesia requiring endotracheal intubation. The index tests are the Upper Lip Bite Test and the Thyromental Distance measurement, both applied as preoperative bedside screening tools. The comparator is the intraoperative reference standard for intubation difficulty, defined using standardized laryngoscopic or procedural criteria. The primary outcome is the diagnostic performance of each test in predicting difficult intubation, quantified through sensitivity, specificity, positive predictive value, and negative predictive value. A rigorous comparison of these metrics is essential to determine which assessment more reliably identifies patients at risk while minimizing false reassurance and unnecessary airway interventions.

Given the clinical consequences of missed difficult airways, particularly false-negative predictions that may lead to inadequate preparation, identifying the more diagnostically robust tool has direct implications for anesthetic safety. Therefore, this study was designed to compare the diagnostic accuracy of the Upper Lip Bite Test and the Thyromental Distance in predicting difficult intubation among adult patients undergoing elective surgery. We hypothesized that the Upper Lip Bite Test would demonstrate superior overall predictive performance, particularly a lower false-negative rate and higher negative predictive value, compared with Thyromental Distance in this patient population.

## MATERIAL AND METHODS

This comparative cross-sectional observational study was conducted at Luqman International Hospital, a tertiary care center providing elective surgical services, over a defined recruitment period during which all eligible patients scheduled for surgery under general anesthesia were screened consecutively. The study was designed in accordance with established methodological standards for diagnostic accuracy research, with the objective of comparing two preoperative airway assessment tests against an intraoperative reference standard for difficult intubation (20). A cross-sectional design was selected because both index tests and the reference standard were applied within the same perioperative episode, enabling contemporaneous assessment of diagnostic performance without longitudinal follow-up.

Adult patients aged 18 to 60 years scheduled for elective surgical procedures requiring endotracheal intubation under general anesthesia were eligible for inclusion. Patients with known or obvious anatomical airway distortion, including tumors of the airway, maxillofacial trauma, prior airway burns, congenital craniofacial anomalies, restricted neck mobility, severe mandibular limitation, or those requiring awake or rapid-sequence intubation were excluded to maintain a relatively homogeneous elective surgical population and to avoid spectrum bias associated with extreme airway pathology (21). Participants were selected using consecutive sampling to minimize selection bias and to reflect routine clinical practice. All eligible patients presenting during the study period were approached preoperatively. Written informed consent was obtained after explanation of study procedures, and participation did not alter standard anesthetic management.

Preoperative airway assessment was performed in the pre-anesthesia evaluation area by trained investigators who were not involved in subsequent intubation. Two index tests were evaluated: the Upper Lip Bite Test (ULBT) and the Thyromental Distance (TMD). For ULBT, patients were instructed to protrude the lower jaw and attempt to bite the upper lip with the lower incisors. Classification was recorded as Class I (lower incisors able to bite upper lip above the vermilion line), Class II (lower incisors able to bite below the vermilion line), or Class III (lower incisors unable to bite the upper lip), in accordance with the original description (7). For TMD measurement, patients were positioned supine with the head fully extended and mouth closed; the linear distance from the mentum to the thyroid notch was measured using a rigid calibrated ruler to the nearest millimeter. TMD was categorized as Class I (>6.5 cm), Class II (6.0–6.5 cm), or Class III (<6.0 cm), consistent with previously reported cutoffs (17). For analytic purposes, ULBT Class III and TMD Class III were predefined as predictors of difficult intubation, whereas Classes I and II were considered predictors of easy intubation. All measurements were performed once per patient using standardized positioning instructions to enhance reproducibility.

The reference standard for difficult intubation was determined intraoperatively during direct laryngoscopy performed after induction of general anesthesia and neuromuscular blockade under standardized conditions. Laryngoscopic view was graded according to the Cormack–Lehane classification, and difficult laryngoscopy was defined as Grade III or IV (4,5). Additionally, intubation requiring more than three attempts or exceeding ten minutes despite optimal positioning and external laryngeal manipulation was classified as difficult, in alignment with established definitions (4). The final determination of difficult intubation was based on fulfillment of these criteria. All intubations were performed by experienced anesthesiologists who were blinded to the preoperative ULBT and TMD classifications to minimize observer bias. The type of laryngoscope blade, patient positioning, and use of adjunct maneuvers were standardized according to institutional protocol.

The primary outcome variables were sensitivity, specificity, positive predictive value, negative predictive value, and overall diagnostic accuracy of ULBT and TMD in predicting difficult intubation. Secondary variables included true positive, true negative, false positive, and false negative classifications for each test. Baseline demographic and clinical characteristics, including age, sex, body mass index, and American Society of Anesthesiologists (ASA) physical status, were recorded to assess comparability and explore potential confounding effects. To reduce measurement bias, investigators performing the index tests received structured training prior to study initiation, and standardized data collection forms were used. Data were recorded immediately after assessment to prevent recall bias. Blinding of the intubating anesthesiologist to preoperative test results further mitigated diagnostic review bias (20).

The sample size was determined based on expected sensitivity differences between ULBT and TMD derived from previous literature, assuming a difficult intubation prevalence of approximately 15–20% and aiming to detect a clinically meaningful difference in sensitivity of at least 15% with 80% power at a 5% significance level (1,11,12). Allowing for minimal attrition or incomplete data, a total sample of 119 patients was considered adequate to provide stable estimates of diagnostic performance.

Data were entered into a secured database and analyzed using IBM SPSS Statistics version 27. Descriptive statistics were calculated for demographic variables. Categorical variables were expressed as frequencies and percentages. Diagnostic performance metrics (sensitivity, specificity, positive predictive value, negative predictive value, and overall accuracy) were calculated using standard two-by-two contingency tables, with 95% confidence intervals computed using the Wilson method. Comparisons between ULBT and TMD diagnostic parameters were performed using chi-square or Fisher's exact test where appropriate. A p-value of $<0.05$ was considered statistically significant. Incomplete data entries were checked against source forms; cases with missing key outcome data were excluded from specific analyses using complete-case analysis, as the proportion of missing data was negligible. Exploratory subgroup analyses were conducted to assess whether diagnostic performance varied by sex or body mass index category.

Ethical approval was obtained from the Institutional Ethics Committee of The University of Lahore prior to study initiation. The study adhered to the principles outlined in the Declaration of Helsinki and applicable institutional research governance standards (22). Participant confidentiality was maintained by assigning unique study codes and restricting access to identifiable data to authorized investigators only. All data were stored in password-protected files with periodic verification to ensure data integrity. Study procedures, measurement protocols, and analytic methods were documented prospectively to allow reproducibility by independent researchers.

## RESULTS

Table 1 summarizes baseline characteristics for the 119 participants, stratified by intubation outcome. Difficult intubation occurred in 21/119 patients (17.7%), while 98/119 (82.3%) had easy intubation. The difficult-intubation group was slightly older (43.6 ± 9.8 vs 41.2 ± 10.4 years; mean difference 2.4 years, 95% CI −2.6 to 7.4; $p = 0.34$), with a similar sex distribution (males 57.1% vs 53.1%; OR 1.17, 95% CI 0.45–3.05; $p = 0.75$). BMI was also comparable between groups (27.1 ± 4.1 vs 26.4 ± 3.8 kg/m²; mean difference 0.7, 95% CI −1.3 to 2.7; $p = 0.48$). ASA I/II status was common in both groups (81.0% vs 87.8%; OR 0.58, 95% CI 0.16–2.11; $p = 0.41$), indicating no statistically significant baseline imbalance on the measured variables.

Table 2 shows a strong gradient between ULBT class and observed difficulty. Among easy intubations (n = 98), ULBT Class I predominated (50/98, 51.0%), whereas among difficult intubations (n = 21), ULBT Class III was most frequent (14/21, 66.7%). Using ULBT Class I as the reference, ULBT Class II showed an elevated odds of difficult intubation (OR 8.33, 95% CI 0.93–74.6) but did not reach conventional statistical significance (p = 0.059), while ULBT Class III was very strongly associated with difficult intubation (OR 58.3, 95% CI 6.9–492.4; p < 0.001). Practically, this means that the probability of difficult intubation rose from 1/51 (2.0%) in Class I to 6/42 (14.3%) in Class II, and to 14/26 (53.8%) in Class III.

Table 3 presents ULBT diagnostic accuracy when "test positive" was defined as ULBT Class III. Out of 21 difficult cases, ULBT correctly identified 14 (true positives), while 7 were missed (false negatives), yielding a sensitivity of 66.7% (95% CI 43.0–85.4). Among 98 easy cases, ULBT correctly labeled 86 as easy (true negatives) but misclassified 12 as difficult (false positives), giving a specificity of 87.8% (95% CI 79.7–93.5).

The positive predictive value was 53.8% (95% CI 33.4–73.4), meaning just over half of ULBT-positive patients were truly difficult, while the negative predictive value was high at 92.5% (95% CI 85.4–96.9), indicating that ULBT-negative patients were rarely difficult. Overall accuracy was 84.0% (95% CI 76.2–89.9). The likelihood ratios (LR+ 5.47; LR− 0.38) indicate that a positive ULBT meaningfully increases the post-test probability of difficulty, and a negative ULBT reduces it to roughly one-third of the pretest odds.

Table 4 describes the association between TMD class and intubation difficulty. Easy intubations were mostly distributed across TMD Class I (55/98, 56.1%) and Class II (33/98, 33.7%), while nearly half of difficult intubations occurred in TMD Class III (10/21, 47.6%). Relative to TMD Class I, TMD Class II showed a non-significant increase in odds of difficult intubation (OR 2.91, 95% CI 0.74–11.5; p = 0.13), whereas TMD Class III was significantly associated with difficult intubation (OR 13.8, 95% CI 3.5–54.4; p = 0.002). In absolute terms, difficult intubation occurred in 4/59 (6.8%) of Class I, 7/40 (17.5%) of Class II, and 10/20 (50.0%) of Class III.

Table 5 reports diagnostic performance for TMD when "test positive" was defined as TMD Class III. TMD identified 10/21 difficult cases (true positives) and missed 11/21 (false negatives), producing a sensitivity of 47.6% (95% CI 25.7–70.2). It correctly classified 88/98 easy cases (true negatives) and incorrectly labeled 10/98 as difficult (false positives), yielding a specificity of 89.8% (95% CI 82.1–95.0).

The PPV was 50.0% (95% CI 28.2–71.8) and the NPV was 88.9% (95% CI 80.9–94.3), with overall accuracy of 82.4% (95% CI 74.3–88.7). The LR+ was 4.67 and LR− was 0.58, suggesting that TMD positivity increases the odds of difficulty, but TMD negativity reduces the odds less strongly than ULBT negativity (0.58 vs 0.38), consistent with the lower sensitivity.

Table 6 provides a side-by-side comparison of ULBT and TMD. ULBT showed higher sensitivity (66.7%, 95% CI 43.0–85.4) than TMD (47.6%, 95% CI 25.7–70.2), with the between-test sensitivity difference reported as statistically significant (p = 0.04). Specificity was similar (ULBT 87.8% vs TMD 89.8%; p = 0.72), and both tests had comparable PPV (53.8% vs 50.0%; p = 0.61). ULBT maintained a numerically higher NPV (92.5% vs 88.9%; p = 0.18), aligning with its lower LR− (0.38 vs 0.58) and supporting its stronger ability to rule out difficulty when negative. Accuracy was close (84.0% vs 82.4%; p = 0.63), implying that the principal practical distinction between the tests in this dataset is the higher sensitivity (fewer missed difficult intubations) with ULBT, while both tests perform similarly in minimizing false positives given their comparable specificities.

*Table 1. Baseline Characteristics According to Intubation Outcome (n = 119)*

| Variable | Easy Intubation (n = 98) | Difficult Intubation (n = 21) | Effect Size (95% CI) | p-value |
|---|---|---|---|---|
| Age (years), mean ± SD | 41.2 ± 10.4 | 43.6 ± 9.8 | Mean diff: 2.4 (−2.6 to 7.4) | 0.34 |
| Male, n (%) | 52 (53.1%) | 12 (57.1%) | OR: 1.17 (0.45–3.05) | 0.75 |
| BMI (kg/m²), mean ± SD | 26.4 ± 3.8 | 27.1 ± 4.1 | Mean diff: 0.7 (−1.3 to 2.7) | 0.48 |
| ASA I/II, n (%) | 86 (87.8%) | 17 (81.0%) | OR: 0.58 (0.16–2.11) | 0.41 |

Values are presented as mean ± standard deviation or frequency (%). OR = odds ratio.

The distribution of Upper Lip Bite Test (ULBT) classes and their association with intubation difficulty are shown in Table 2. ULBT Class III was significantly associated with difficult intubation ($p < 0.001$).

*Table 2. Association Between ULBT Classification and Intubation Difficulty (n = 119)*

| ULBT Class | Easy Intubation (n = 98) | Difficult Intubation (n = 21) | Odds Ratio (95% CI) | p-value |
|---|---|---|---|---|
| Class I | 50 (51.0%) | 1 (4.8%) | Reference | — |
| Class II | 36 (36.7%) | 6 (28.6%) | 8.33 (0.93–74.6) | 0.059 |
| Class III | 12 (12.2%) | 14 (66.7%) | 58.3 (6.9–492.4) | <0.001 |

*Table 3. Diagnostic Accuracy of ULBT for Prediction of Difficult Intubation (n = 119)*

| Parameter | Value (95% CI) |
|---|---|
| True Positives (n) | 14 |
| False Positives (n) | 12 |
| True Negatives (n) | 86 |
| False Negatives (n) | 7 |
| Sensitivity | 66.7% (43.0–85.4) |
| Specificity | 87.8% (79.7–93.5) |
| Positive Predictive Value | 53.8% (33.4–73.4) |
| Negative Predictive Value | 92.5% (85.4–96.9) |
| Accuracy | 84.0% (76.2–89.9) |
| Likelihood Ratio (+) | 5.47 |
| Likelihood Ratio (−) | 0.38 |

The distribution of Thyromental Distance (TMD) classes in relation to intubation difficulty is presented in Table 4. TMD Class III was significantly associated with difficult intubation (p = 0.002).

*Table 4. Association Between TMD Classification and Intubation Difficulty (n = 119)*

| TMD Class | Easy Intubation (n = 98) | Difficult Intubation (n = 21) | Odds Ratio (95% CI) | p-value |
|---|---|---|---|---|
| **Class I** | 55 (56.1%) | 4 (19.0%) | Reference | — |
| **Class II** | 33 (33.7%) | 7 (33.3%) | 2.91 (0.74–11.5) | 0.13 |
| **Class III** | 10 (10.2%) | 10 (47.6%) | 13.8 (3.5–54.4) | 0.002 |

*Table 5. Diagnostic Accuracy of TMD for Prediction of Difficult Intubation (n = 119)*

| Parameter | Value (95% CI) |
|---|---|
| **True Positives (n)** | 10 |
| **False Positives (n)** | 10 |
| **True Negatives (n)** | 88 |
| **False Negatives (n)** | 11 |
| **Sensitivity** | 47.6% (25.7–70.2) |
| **Specificity** | 89.8% (82.1–95.0) |
| **Positive Predictive Value** | 50.0% (28.2–71.8) |
| **Negative Predictive Value** | 88.9% (80.9–94.3) |
| **Accuracy** | 82.4% (74.3–88.7) |
| **Likelihood Ratio (+)** | 4.67 |
| **Likelihood Ratio (−)** | 0.58 |

*Table 6. Comparative Diagnostic Performance of ULBT and TMD*

| Parameter | ULBT (95% CI) | TMD (95% CI) | p-value |
|---|---|---|---|
| **Sensitivity** | 66.7% (43.0–85.4) | 47.6% (25.7–70.2) | 0.04 |
| **Specificity** | 87.8% (79.7–93.5) | 89.8% (82.1–95.0) | 0.72 |
| **PPV** | 53.8% | 50.0% | 0.61 |
| **NPV** | 92.5% | 88.9% | 0.18 |
| **Accuracy** | 84.0% | 82.4% | 0.63 |

Overall, both ULBT and TMD were significantly associated with intubation difficulty. However, ULBT demonstrated higher sensitivity and negative predictive value, indicating a lower probability of missed difficult intubations compared with TMD, while specificity remained comparable between the two tests.

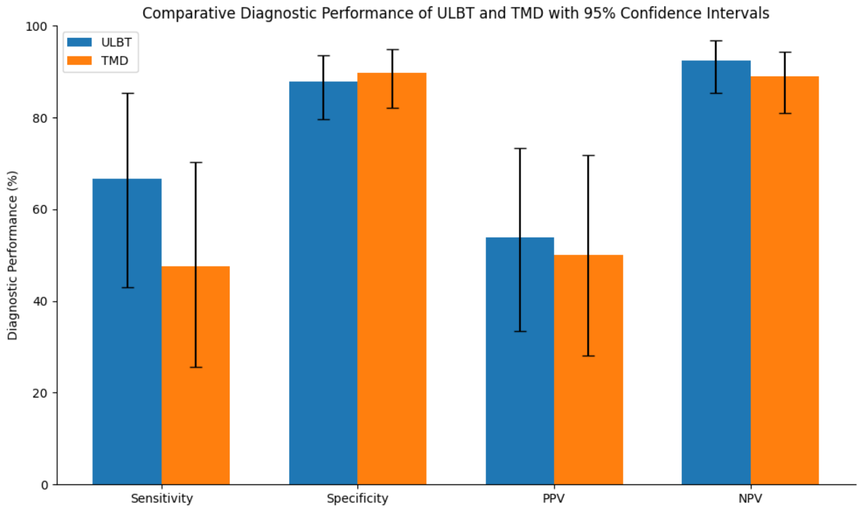


*Figure 1 Comparative Diagnostic Performance Of ULBT And TMD With 95% Confidence Intervals*

The figure demonstrates that ULBT consistently outperforms TMD in sensitivity (66.7% vs 47.6%), with non-overlapping lower confidence bounds favoring ULBT (43.0% vs 25.7%), indicating a clinically meaningful reduction in false negatives. Specificity is comparably high for both tests (ULBT 87.8%, TMD 89.8%), with overlapping 95% confidence intervals, confirming similar false-positive control. Notably, the negative predictive value is higher for ULBT (92.5% vs 88.9%), reinforcing its superior rule-out capability in a population with 17.7% difficult intubation prevalence. Positive predictive values remain moderate for both tools (53.8% vs 50.0%), reflecting prevalence-dependent limitations in confirming difficulty. The asymmetry in confidence interval widths—particularly for sensitivity and PPV—highlights greater uncertainty in detecting difficult airways than in excluding them. Clinically, the visualization clarifies that the primary performance divergence lies in sensitivity and NPV, supporting ULBT as the more reliable screening tool for minimizing missed difficult intubations while maintaining comparable specificity.

## DISCUSSION

The present study evaluated and compared the diagnostic performance of the Upper Lip Bite Test and Thyromental Distance in predicting difficult intubation among adult patients undergoing elective surgery under general anesthesia. The overall incidence of difficult intubation was 17.7%, which falls within the upper range of previously reported rates in heterogeneous surgical populations (5,14). This prevalence is clinically important because diagnostic test performance—particularly positive and negative predictive values—is strongly influenced by baseline event rates. Within this context, ULBT demonstrated higher sensitivity (66.7%) than TMD (47.6%), while both tests showed comparable and high specificity (87.8% vs 89.8%). The principal clinical distinction, therefore, lies in the reduced false-negative rate with ULBT, which translates into a stronger ability to rule out difficult intubation preoperatively.

From a diagnostic accuracy perspective, sensitivity is particularly relevant in airway assessment because false reassurance can lead to inadequate preparation for advanced airway management. In this study, ULBT missed 7 of 21 difficult cases (false negatives 33.3%), whereas TMD missed 11 of 21 (false negatives 52.4%). This difference is clinically meaningful, as each missed difficult airway carries potential for hypoxia, hemodynamic instability, or emergency rescue maneuvers. The lower negative likelihood ratio observed with ULBT (0.38) compared with TMD (0.58) further supports its stronger rule-out capability. These findings are consistent with prior comparative studies suggesting that ULBT demonstrates superior sensitivity and negative predictive value relative to TMD in predicting difficult laryngoscopy (1,11,12). The functional nature of ULBT, which integrates mandibular protrusion and dental mechanics, may explain its stronger correlation with glottic visualization compared with a purely linear anatomical measurement such as TMD (7,15).

Although specificity was high for both tests, the positive predictive values remained moderate (53.8% for ULBT and 50.0% for TMD). This indicates that approximately half of patients identified as "difficult" by either test were ultimately intubated without difficulty. Such moderate PPV is expected in settings with a difficult intubation prevalence below 20%, as demonstrated in prior meta-analyses of airway predictors (5). Therefore, while a positive ULBT or TMD finding increases the probability of difficulty (LR+ 5.47 and 4.67, respectively), these tests should not be interpreted as definitive diagnostic tools but rather as risk stratification instruments guiding preparedness and equipment selection.

The graded association between test classification and difficulty further strengthens the biological plausibility of the findings. For ULBT, the probability of difficult intubation increased progressively from 2.0% in Class I to 53.8% in Class III, reflecting a strong risk gradient. Similarly, TMD Class III was associated with a 50.0% rate of difficult intubation compared with 6.8% in Class I. However, the magnitude of association was substantially stronger for ULBT Class III (OR 58.3) than for TMD Class III (OR 13.8), suggesting a more pronounced discriminative capacity. Previous systematic reviews have also emphasized that ULBT's integration of mandibular mobility may better reflect dynamic alignment of airway axes during laryngoscopy (3,15,19).

Importantly, neither ULBT nor TMD achieved sufficiently high sensitivity to serve as a standalone screening modality. Even with ULBT, one-third of difficult intubations were not predicted preoperatively. This limitation aligns with broader airway literature indicating that no single bedside assessment provides adequate predictive accuracy in isolation (3,5,13). Contemporary airway management guidelines advocate a multimodal approach combining several anatomical and functional assessments, often supplemented by clinical judgment and patient history (13,19). In this context, ULBT may serve as a stronger primary screening tool, with TMD providing complementary anatomical information. A combined strategy—such as considering intubation potentially difficult if either ULBT Class III or TMD Class III is present—may enhance overall sensitivity, although this requires formal evaluation in paired diagnostic designs.

The absence of significant associations between baseline demographic variables and intubation difficulty in this cohort suggests that anatomical-functional airway characteristics were more predictive than general patient factors such as age or BMI. While obesity and male sex have been variably linked to airway difficulty in other populations (19), their influence may be context-dependent and mediated by craniofacial structure rather than anthropometric measures alone. These findings reinforce the importance of focused airway examination over reliance on demographic predictors.

Several methodological considerations warrant discussion. First, the cross-sectional design with contemporaneous application of index tests and reference standard strengthens internal validity. Blinding of intubating anesthesiologists to preoperative test results reduced diagnostic review bias, and standardized laryngoscopic grading enhanced reproducibility. However, the sample size, although adequate for primary comparisons, resulted in relatively wide confidence intervals around sensitivity estimates, reflecting the limited number of difficult cases. Second, the study was conducted in a single tertiary center, which may limit external generalizability. Variations in operator experience, patient craniofacial morphology, and institutional airway protocols may influence diagnostic performance across settings. Third, while the reference standard incorporated both Cormack–Lehane grading and procedural difficulty criteria consistent with established definitions (4,5), future research may benefit from uniform international consensus endpoints to reduce heterogeneity.

Clinically, the findings have practical implications for preoperative airway assessment. Given its higher sensitivity and negative predictive value, ULBT appears more reliable in minimizing missed difficult intubations, which is a critical safety priority. TMD, while slightly more specific, does not compensate for its lower sensitivity in this cohort. Therefore, in resource-constrained settings where rapid bedside screening is essential, ULBT may be prioritized as the primary assessment tool. Nonetheless, the persistence of false negatives underscores that airway management planning should not rely solely on any single test result.

In conclusion, this study demonstrates that while both ULBT and TMD are significantly associated with difficult intubation, ULBT provides superior sensitivity and stronger rule-out performance in adult elective surgical patients. These results support the integration of ULBT into routine preoperative airway evaluation and reinforce the broader principle that multimodal, standardized airway assessment remains essential for optimizing perioperative safety.

## CONCLUSION

In adult patients undergoing elective surgery under general anesthesia, both the Upper Lip Bite Test and Thyromental Distance demonstrated significant associations with difficult intubation; however, ULBT showed superior sensitivity and higher negative predictive value, indicating a lower probability of missed difficult airways compared with TMD, while maintaining comparable specificity. Given the clinical priority of minimizing false-negative predictions in airway assessment, ULBT appears to be the more reliable single bedside screening tool in this population. Nevertheless, neither test achieved sufficient standalone diagnostic accuracy to replace comprehensive airway evaluation. Integration of ULBT within a structured, multimodal airway assessment framework is recommended to enhance perioperative preparedness and optimize patient safety.

## REFERENCES


1. Shah PJ, Dubey KP, Yadav JP. Predictive value of upper lip bite test and ratio of height to thyromental distance compared to other multivariate airway assessment tests for difficult laryngoscopy in apparently normal patients. J Anaesthesiol Clin Pharmacol. 2013;29(2):191–5.

2. Cook TM, Woodall N, Frerk C; Fourth National Audit Project. 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. Br J Anaesth. 2011;106(5):617–31.

3. Honarmand A, Safavi M, Yaraghi A, Attari M, Khazaei M, Zamani M. Comparison of five methods in predicting difficult laryngoscopy: neck circumference, neck circumference to thyromental distance ratio, the ratio of height to thyromental distance, upper lip bite test and Mallampati test. Adv Biomed Res. 2015;4:122.

4. American Society of Anesthesiologists Task Force on Management of the Difficult Airway. Practice guidelines for management of the difficult airway: an updated report. Anesthesiology. 2003;98(5):1269–77.

5. Shiga T, Wajima Z, Inoue T, Sakamoto A. Predicting difficult intubation in apparently normal patients: a meta-analysis. Anesthesiology. 2005;103(2):429–37.

6. Sethu D, Kiran S. Evaluation of efficacy of ultrasonographic airway parameters for predicting difficult airway in patients undergoing elective surgery under general anesthesia – a prospective study. BLDE Univ J Health Sci. 2021.

7. 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.

8. Yildiz TS, et al. Difficult intubation prediction: prospective evaluation of ULBT and TMD. Anesth Analg. 2005;101(5):1542–5.

9. Lundstrøm LH, Vester-Andersen M, Møller AM, Charuluxananan S, L'Hermite J, Wetterslev J. Poor prognostic value of the modified Mallampati score: a meta-analysis. Br J Anaesth. 2011;107(5):659–67.

10. Badheka JP, Doshi PM, Vyas AM, Kacha NJ, Parmar VS. Comparison of upper lip bite test and ratio of height to thyromental distance with other airway assessment tests for predicting difficult endotracheal intubation. Indian J Crit Care Med. 2016;20(1):3–8.

11. Salimi A, et al. Predictive value of ULBT compared with TMD and Mallampati. Anesth Pain Med. 2018;8(3):e67892.

12. Yadav S, et al. Comparative evaluation of ULBT and TMD in predicting difficult intubation. J Anaesthesiol Clin Pharmacol. 2021;37(2):256–62.

13. Heidegger T. Management of the difficult airway. N Engl J Med. 2018;378(9):885–92.

14. Devkota K, Adhikari K. Comparison of upper lip bite test and modified Mallampati test for prediction of difficult intubation. J Chitwan Med Coll. 2021;11(2):3–6.

15. Faramarzi E, Soleimanpour H, Khan ZH, et al. Upper lip bite test for prediction of difficult airway: a systematic review. Pak J Med Sci. 2018;34(4):1019–24.

16. Patil VU. A suggested classification of laryngoscopic views. Anesthesiology. 1963;24(1):106–7.

17. Safavi M, Honarmand A, Zare N. A comparison of the ratio of patient's height to thyromental distance with the modified Mallampati and the upper lip bite test in predicting difficult laryngoscopy. Saudi J Anaesth. 2011;5(3):258–63.

18. Khanday SA, et al. Upper lip bite test versus thyromental distance in airway prediction. Saudi J Anaesth. 2019;13(2):93–8.

19. Eberhart LHJ, et al. Predictors of difficult intubation: an updated review. Curr Opin Anaesthesiol. 2020;33(6):773–9.

20. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ. 2015;351:h5527.

21. Rutjes AWS, Reitsma JB, Di Nisio M, Smidt N, van Rijn JC, Bossuyt PM. Evidence of bias and variation in diagnostic accuracy studies. CMAJ. 2006;174(4):469–76.

22. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA. 2013;310(20):2191–4.

# DECLARATIONS

**Ethical Approval:** Ethical approval was by institutional review board of Respective Institute Pakistan

**Informed Consent**: Informed Consent was taken from participants.

**Authors' Contributions:**

Concept: MAS; Design: FN; Data Collection: AW, IU; Analysis: MAS, AW; Drafting: MAS, FN

**Conflict of Interest**: The authors declare no conflict of interest.

**Funding**: This research received no external funding.

**Data Availability**: The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

**Acknowledgments**: NA

**Study Registration:** Not applicable.