A Comparison between Modified Mallampati Test, Thyromental Distance, Upper Lip Bite Test and Neck Circumference for Predicting Difficult Intubation

Nishtha Sharma¹, V. C. Chandak², Amol Singam³

¹Anaesthesiology/MD Resident/Jawaharlal Nehru Medical College, Sawangi (M), Wardha, Maharashtra, India
²Anaesthesiology/Associate Professor/Jawaharlal Nehru Medical College, Sawangi (M), Wardha, Maharashtra, India
³Anaesthesiology/Associate Professor/Jawaharlal Nehru Medical College, Sawangi (M), Wardha, Maharashtra, India

Abstract: Background: Various bedside screening tests are employed for the prediction of difficult intubation. In this study, we compared 4 bedside tests viz. the Modified Mallampati Test (MMT), the Thyromental Distance (TMD) measurement, the Upper Lip Bite Test (ULBT) and the Neck Circumference (NC) measurement in terms of Sensitivity, Specificity, Positive Predictive Value and Negative Predictive Value to determine which is the best individual predictor of difficult intubation, in apparently normal patients. Methods: 120 patients of ASA physical status I&II, aged 16-55years, of either gender, undergoing elective surgery under general anaesthesia were enrolled in this prospective study. MMT Classes III&IV, TMD<6.0cm, ULBT Class III and NC<23.5cm were considered to be predictive of difficult intubation. A Cormack-Lehane laryngoscopy grade 3 or 4 was considered a difficult intubation. Results: The incidence of difficult intubation was 6.67%. The sensitivity, specificity, positive and negative predictive values of the MMT were 62.50%, 86.61%, 25% and 97% respectively. The sensitivity, specificity, positive and negative predictive values of the TMD were 25%, 93.75%, 22.22% and 94.59% respectively. The sensitivity, specificity, positive and negative predictive values of the ULBT were 37.50%, 91.07%, 23.08% and 95.33% respectively. The sensitivity, specificity, positive and negative predictive values of the NC were 87.50%, 43.75%, 10% and 98% respectively. We also found a statistically significant correlation between the MMT and the CL grading and also between the ULBT and the CL grading. Conclusion: No single predictor is sufficient for prediction of difficult intubation on its own. All the studied bedside tests are poor to moderate predictors of difficult intubation. All the tests showed poor positive predictive values and high negative predictive values which suggests that they can be more useful predictors of easy intubation than difficult intubation.

Keywords: Modified Mallampati Test, Thyromental Distance, Upper Lip Bite Test, Neck Circumference, Difficult Intubation

1. Introduction

The most ultimate question, is how do we define difficult intubation? While some authors take into account the number of attempts¹, number of operators or special intubating aids ², ³, others⁴ argue that actual failed intubation is a very poor criterion for difficult intubation not only because of its very low incidence but also because of its dependence on the skill of the anaesthesiologist. These anaesthesiologists employ the Cormack-Lehane grading of the glottic view on direct laryngoscopy as the criterion for difficult intubation⁵.⁶.

The third definition is a modification of the above CL grading system. Wilson et al⁷ used a scale of 1-5 which was further transformed into what we now know as the Modified Cormack Lehane scale introduced by Yentis and Lee⁸ in 1998 whereas the 4th definition that we came across is that of the intubation difficulty scale (IDS) introduced by Adnet et al⁹ in 1997.

Despite all these definitions, the Cormack-Lehane grading still remains the most widely accepted criterion of difficult intubation.

A study pertaining to difficult intubation can be carried out in 2 ways. A prospective study or a retrospective study. The later seems like a better option since the incidence of difficult intubation is so low, then why not go for a retrospective study and just measure the features of patients with known difficult intubation? It is however, not that simple. The major drawback here would be that there is no control group- nothing to compare these parameters to. There is nothing to tell us about the easy patients. The percentage of false positives cannot be accurately assessed.

The better option then, is to study these features in a lot of routine patients and then identify those which are independent predictors of a difficult intubation.

Several authors have combined various predictors and developed different risk scoring systems⁵.¹⁰ Nonetheless, to develop a scoring system we must first identify which predictors are better than the others so that the best indicators can be combined for a more accurate risk-sum scoring system.

The next concern which arises is the diagnostic accuracy of the various predictors of difficult intubation. The sensitivity, specificity, PPV and NPV of these tests varies from trial to trial. This inconsistency can be attributed to many reasons. Firstly, there are gross differences in the incidence of difficult intubation from one study to the next. Where Oates et al¹¹ found an incidence as low as 1.8%, Tse et al¹² encountered an incidence as high as 13% in their study.

The next concern which arises is the diagnostic accuracy of the various predictors of difficult intubation. The sensitivity, specificity, PPV and NPV of these tests varies from trial to trial. This inconsistency can be attributed to many reasons. Firstly, there are gross differences in the incidence of difficult intubation from one study to the next. Where Oates et al¹¹ found an incidence as low as 1.8%, Tse et al¹² encountered an incidence as high as 13% in their study.
The second reason for the varied performance of different predictors is of course the definition of difficult intubation used. While one study uses the CL grading as the gold standard of difficult intubation, another uses the IDS\(^2\).\(^3\)\(^4\).

The third reason is the different thresholds used for the different tests especially the quantitative tests. While Brodsky et al\(^1\) consider 46cm as the cut-off point of NC measurement, Gonzalez et al\(^5\) consider 43cm for the same whereas Shailaja et al\(^6\) take it to be 35cm. The same goes for the TMD measurement. While Frerk\(^2\) considers a cut-off point of <7.0cm, Khan et al\(^6\) considered <6.5cm whereas Ezri et al\(^7\) and Salimi et al\(^8\) considered <6.0cm and <4cm for prediction of difficult intubation respectively. While it is true that decreasing the threshold of a test will decrease the number of false positives, it is also very true that it will lead to a fall in the predicted number of difficult intubations as well.

The last reason to be considered here is the differences in the demographic profiles in different studies. While some studies only enroll apparently normal ASA physical status I & II patients, others consider the obese, the elderly, the obstetric patients as well as those with known anatomic abnormalities. The question here is whether the accuracy of these tests differs in special groups of patients, especially those in whom difficult intubation is a more dreaded possibility like the obese and the obstetric patients.

The important thing to remember, is that no bedside test is perfect. Therefore, becoming disheartened with the results of a particular test in a particular study and abandoning it all together is not the way to go.

We should find and agree upon standardization of not only the cut-off points of various quantitative tests, but also find one universal definition of difficult intubation to minimize the discrepancies in the outcome of different studies, to some extent.

This study was designed to compare the Modified Mallampati Test (MMT), the Thyromental Distance (TMD) measurement, the Upper Lip Bite Test (ULBT) and the Neck Circumference (NC) measurement in terms of Sensitivity, Specificity, Positive Predictive Value and Negative Predictive Value in order to determine which is the best individual predictor of difficult intubation, in apparently normal patients.

2. Material & Methods

This prospective study was conducted in 120 patients aged 16-55 years, of ASA physical status I & II, of either gender, after obtaining institutional ethical committee clearance and written informed consent from all the participants.

Patients with cervical spine abnormalities or trauma, edentulous patients, obese patients (BMI > 30kg/m\(^2\)), obstetric patients, patients with restricted mouth opening, oral pathology or facial injuries and patients with any thyroid pathology or neck mass were excluded from this study.

A thorough pre-anaesthetic evaluation was carried out in all the patients and the procedure was explained to them in detail. Airway assessment was done in all patients, using the Modified Mallampati Test (MMT), the Thyromental Distance (TMD) measurement, the Upper Lip Bite Test (ULBT) and the Neck Circumference (NC) measurement.

In our study, classification of the oropharyngeal view was done according to the Samsoon and Young’s modification of the Mallampati classification-the Modified Mallampati Test (MMT), wherein the seated patient was asked to open the mouth fully with the tongue protruded maximally without phonation. The test was performed twice in each patient to minimize the chances for erroneous observation. The patient was observed from eye level and the view was graded as below.

Class I - Soft palate, fauces, uvula, and pillars are seen
Class II - Soft palate, fauces, and uvula are seen
Class III - Soft palate and base of uvula
Class IV - Soft palate not visible

The TMD was measured in centimeters (cm) from the thyroid notch to the bony point of the chin with the head fully extended on the neck with the mouth closed. A ruler was used and the distance was approximated to the nearest 0.5cm. It was classified as below:
- Class I: >6.5cm = normal value
- Class II: 6.0-6.5cm = laryngoscopy and intubation are difficult but usually possible
- Class III: <6.0cm = laryngoscopy may be impossible

The ULBT is classified according to the ability of the lower teeth to bite the upper lip. The observer demonstrated the test to the patient before asking them to perform it. It was assessed and classified as under:
- Class I = lower incisors can bite the upper lip above the vermillion line
- Class II = lower incisors can bite the upper lip below the vermillion line
- Class III = lower incisors cannot bite the upper lip

The NC was measured in centimeters (cm) using a measuring tape at the level of the thyroid cartilage.

MMT Classes III&IV, TMD <6.0cm\(^9\),\(^10\),\(^15\), ULBT Class III and NC > 35cm\(^13\) were taken as predictors of difficult intubation. Difficult intubation was defined as Cormack-Lehane grades 3&4 of glottic view upon direct laryngoscopy.

The sensitivity, specificity, positive predictive value and negative predictive value of all these tests was calculated and compared. The correlation, if any, between each test and the CL grading was also assessed.

3. Statistical Analysis

Statistical analysis was done by using descriptive and inferential statistics with the help of Chi-square test and Binary classification. Software used in the analysis was
SPSS 17.0 version and Graph Pad prism 6.0 version. The p-value <0.05 was considered to be significant.

4. Results

The incidence of difficult intubation in our study was 6.67%. There was no failed intubation in our study.

Table 1 shows the distribution of patients according to their demographic profile. In our study, the mean Age of the patients was 34.34 ± 9.15; the mean Height was 1.65 ± 0.05; the mean Weight was 57.88 ± 5.67 and the mean BMI was 21.02 ± 1.49.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Mean ± SD</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (yrs)</td>
<td>34.34 ± 9.15</td>
<td>34.50</td>
</tr>
<tr>
<td>Height (m)</td>
<td>1.65 ± 0.05</td>
<td>1.65</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>57.88 ± 5.67</td>
<td>57</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>21.02 ± 1.49</td>
<td>20.68</td>
</tr>
</tbody>
</table>

BMI = Body Mass Index

Table 2 shows the distribution of patients according to the MMT grades. In our study, 68 (56.67%) patients were in Class I, 32 (26.67%) patients in Class II, 20 (16.66%) patients in Class III and 0 patients in Class IV respectively.

<table>
<thead>
<tr>
<th>Class</th>
<th>No of patients</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class I</td>
<td>68</td>
<td>56.67</td>
</tr>
<tr>
<td>Class II</td>
<td>32</td>
<td>26.67</td>
</tr>
<tr>
<td>Class III</td>
<td>20</td>
<td>16.66</td>
</tr>
<tr>
<td>Class IV</td>
<td>0</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Table 3 shows the distribution of patients according to the TMD measurement. There were 84 (70%) patients in Class I (>35cm); 27 (22.5%) patients in Class II (6.0-6.5cms) and 9 (7.5%) patients in Class III (<6.0cm) respectively.

<table>
<thead>
<tr>
<th>Class</th>
<th>No of patients</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class I</td>
<td>84</td>
<td>70.00</td>
</tr>
<tr>
<td>Class II</td>
<td>27</td>
<td>22.50</td>
</tr>
<tr>
<td>Class III</td>
<td>9</td>
<td>7.50</td>
</tr>
<tr>
<td>TOTAL</td>
<td>120</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Table 4 shows the distribution of patients according to the ULBT. There were 72 (60%) patients in Class I; 35 (29.17%) patients in Class II and 13 (10.83%) patients in Class III respectively.

<table>
<thead>
<tr>
<th>Class</th>
<th>No. of patients</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class I</td>
<td>72</td>
<td>60.00</td>
</tr>
<tr>
<td>Class II</td>
<td>35</td>
<td>29.17</td>
</tr>
<tr>
<td>Class III</td>
<td>13</td>
<td>10.83</td>
</tr>
<tr>
<td>TOTAL</td>
<td>120</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Table 5 shows the distribution of patients according to the NC measurement. The <35cm group had 50 (41.67%) patients and the ≥ 35cm group had 70 (58.33%) patients.

<table>
<thead>
<tr>
<th>NC (cm)</th>
<th>No of patients</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;35</td>
<td>50</td>
<td>41.67</td>
</tr>
<tr>
<td>≥ 35</td>
<td>70</td>
<td>58.33</td>
</tr>
<tr>
<td>TOTAL</td>
<td>120</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 6 shows the distribution of patients according to Cormack-Lehane Classification of the view on direct laryngoscopy. There were 81 (67.5%) patients with Grade 1; 31 (25.83%) patients had Grade 2; 8 (6.67%) patients had Grade 3 and 0 (0%) patients had Grade 4 respectively.

<table>
<thead>
<tr>
<th>Grade</th>
<th>No of patients</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade 1</td>
<td>81</td>
<td>67.50</td>
</tr>
<tr>
<td>Grade 2</td>
<td>31</td>
<td>25.83</td>
</tr>
<tr>
<td>Grade 3</td>
<td>8</td>
<td>6.67</td>
</tr>
<tr>
<td>Grade 4</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>TOTAL</td>
<td>120</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 7 shows the comparison between the sensitivity, specificity, PPV, NPV and Accuracy of the MMT, the TMD, the ULBT and the NC.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>MMT</th>
<th>TMD</th>
<th>ULBT</th>
<th>NC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>62.50%</td>
<td>25%</td>
<td>37.50%</td>
<td>87.50%</td>
</tr>
<tr>
<td>Specificity</td>
<td>86.61%</td>
<td>93.75%</td>
<td>91.07%</td>
<td>43.75%</td>
</tr>
<tr>
<td>PPV</td>
<td>25%</td>
<td>22.22%</td>
<td>23.08%</td>
<td>10%</td>
</tr>
<tr>
<td>NPV</td>
<td>97%</td>
<td>94.59%</td>
<td>95.33%</td>
<td>98%</td>
</tr>
<tr>
<td>Accuracy</td>
<td>85%</td>
<td>89.16%</td>
<td>87.50%</td>
<td>46.67%</td>
</tr>
<tr>
<td>p-value</td>
<td>8.24, p=0.004,S</td>
<td>3.78, p=0.05,NS</td>
<td>6.31, p=0.012,S</td>
<td>3.00, p=0.08, NS</td>
</tr>
</tbody>
</table>

PPV = Positive Predictive Value; NPV = Negative Predictive Value

The sensitivity, specificity, positive and negative predictive values of the MMT were 62.50%, 86.61%, 25% and 97% respectively whereas the sensitivity, specificity, positive and negative predictive values of the TMD were 25%, 93.75%, 22.22% and 94.59% respectively. The sensitivity, specificity, positive and negative predictive values of the ULBT were 37.50%, 91.07%, 23.08% and 95.33% respectively whereas the sensitivity, specificity, positive and negative predictive values of the NC were 87.50%, 43.75%, 10% and 98% respectively.

We also found a statistically significant correlation between the MMT and the CL grading and also between the ULBT and the CL grading.

5. Discussion

In apparently normal patients with no clinically obvious anatomical abnormalities, unanticipated difficult intubation is not only a challenge to the skill of the anaesthesiologist but also carries the risk of significant morbidity and mortality for the patient.

The various bedside screening tests employed for prediction of difficult intubation should not only have the benefit of
The Modified Mallampati Test has been criticized over the years for various limitations. It has been mentioned that the demarcation between classes II & III is unclear[17-19]. It has also been pointed out that this test has a high inter-observer variability and a large number of false positives[16]. Another limitation is that many patients involuntarily tend to phonate which affects the classification[18,19]. Despite all its shortcomings, the MMT is still the most routinely used bedside test for the prediction of difficult intubation.

In our study we found the sensitivity of the MMT to be 62.5% which is similar to the results found by Isaac and Sharma[20] (62.5%). The specificity of the MMT in our study was 86.6% which correlates with the findings of Shiga et al[21] (86%) and Gonzalez et al[22] (87%). The positive predictive value of the MMT in our study is 25% which is close to the results of Tse et al[7] (22%). The negative predictive value of the MMT in our study is 97% which is similar to the studies of Khan et al[23] (98.4%) and Gonzalez et al[12] (96%). The number of false positive cases in our study were 15.

We also found a statistically significant (p-value = 0.002) correlation between the MMT and the CL grading which was similar to the results of Mallampati et al[22] and Ezri et al[15].

The measurement of the TMD originated as a quantitative assessment of the receding jaw. Even though some authors[23] have questioned its relevance as a predictor of difficult intubation, it is a popular bedside test because of its ability to assess not only a receding mandible but also the range of head extension. While the MMT predicts difficult intubation by assessing the amount of soft tissue that would be deformed during laryngoscopy, the TMD measurement does so by assessing the size of the region to which the soft tissue has to be deformed[24].

In our study we found the sensitivity of the TMD to be 25% which is slightly more than that found by Shiga et al[21] (20%). The specificity of the TMD in our study is 93.75% which correlates with the results of Shiga et al[21] (94%). The positive predictive value of the TMD in our study is 22.22% which correlates with the findings of Salimiet al[16] (22%) and Tse et al[7] (20%). The negative predictive value of the TMD in our study is 94.59% is similar to that found by Mehta et al[25] (93.79%).

There was no correlation between the TMD and the CL grading similar to the results found by Brodsky et al[1].

The ULBT is classified according to the ability of the lower teeth to bite the upper lip. The ULBT not only takes into account the presence of buck teeth but also the range and freedom of movements of the mandible. Moreover, the 3 classes are clearly demarcated, making inter-observer variability highly unlikely.

he ULBT has an important limitation that it cannot be used in edentulous patients and also some patients have difficulty in understanding how to perform the test. Since edentulous patients are excluded from our study, we did not encounter this problem and also the observer in our study demonstrated the test to the patient making it easier to perform.

In our study we found the sensitivity of the ULBT to be 37.50% which was close to the findings of Mohan and Mohana[26] (40%). The specificity of the ULBT in our study is 91.07% which correlates with the findings of Khan et al[14] (91.9%), Shah et al[27] (91.53%), Eberhart et al[28] (92.5%) and Ali et al[29] (92.9%). The positive predictive value of the ULBT in our study is 23.08% which is slightly less than the original findings of Khan et al[6] (28.9%). The negative predictive value of the ULBT in our study is 95.33% which correlates with the results of Shah et al[27] (95.70%) and Ali et al[29] (97.3%).

There was a statistically significant (p-value = 0.009) correlation between the ULBT and the CL grading similar to the results reported by Hester et al[30].

The neck circumference has been regarded a good predictor of difficult intubation in terms of sensitivity[12] and inter-observer reliability[31].

In our study we found the sensitivity of the NC to be 87.50% which is lower than that reported by Gonzalez et al[12] (92%). The specificity of the NC in our study is 43.75% which correlates with that found by Chara et al[32] (44.6%). The positive predictive value of the NC in our study is 10% which is lower than that reported by Chara et al[32] (15.4%). The negative predictive value of the NC in our study is 98% which is similar to that of Gonzalez et al[12] (99%).

There was no correlation between the NC and the CL grading similar to the results reported by Karc et al[33].

Combining all the results, the TMD shows the highest accuracy but it also has the poorest sensitivity. The NC has the highest sensitivity but a very poor specificity and accuracy.

All the tests showed poor positive predictive values and high negative predictive values which suggests that they can be more useful predictors of easy intubation than difficult intubation.

The limitations of our study were a strictly limited demographic profile. This study can be conducted in a much wider demographic scenario. We did not assess these tests in combination, which was another limitation of our study. Various combinations of these tests should be assessed and applied for better prediction of difficult intubation.

In conclusion, no single predictor is sufficient for prediction of difficult intubation on its own. All the studied bedside tests are poor to moderate predictors of difficult intubation. All the tests showed poor positive predictive values and high negative predictive values which suggests that they can be more useful predictors of easy intubation than difficult intubation.
References


