Systematic Review

Upper lip bite test for prediction of difficult airway: A systematic review

Elnaz Faramarzi1, Hassan Soleimanpour2, Zahid Hussain Khan3, Ata Mahmoodpoor4, Sarvin Sanaie5

ABSTRACT

Background and Objective: Upper lip bite test (ULBT) is one of the various bedside tests used for prediction of difficult laryngoscopic intubation. However, its usefulness is not still very clear, and there is controversy regarding its accuracy. The aim of this systematic review was to determine the accuracy of the ULBT for predicting difficult airway including difficult laryngoscopy or difficult tracheal intubation.

Methods: We searched the databases of PubMed, Scopus, and Google scholar for prospective studies published up until October 2016 assessing the accuracy of ULBT in comparison to Cormack-Lehane grading. The selected keywords were “upper lip bite test”, “upper lip catch test”, “prediction”, “difficult airway”, “difficult laryngoscopy”, “difficult intubation”. Inclusion criteria were studies assessing ULBT for prediction of difficult intubation, considering Cormack-Lehane grade III and IV as difficult airway, written in English, and reporting sensitivity, specificity, NPV, PPV, and accuracy. Exclusion criteria were studies not reporting accuracy or not having enough data for its calculation. Based on the mentioned criteria, 27 studies enrolling 18141 patients were included. This systematic review was performed based on the guidelines on conducting systematic reviews of diagnostic studies.

Results: Prevalence of airway difficulties according to the direct laryngoscopic view varied from 2.8% to 27% and according to the ULBT was from 2% to 21%. In 11 of the 27 studies, sensitivity of ULBT in prediction of difficult airway was more than 70%. All of the studies except one showed a high specificity for ULBT (>85%). Moreover, these studies indicated a high NPV. Accuracy of ULBT was >85% in 24 out of 27 studies.

Conclusion: It appears that ULBT is a useful bedside test for evaluation of patient airway before the general anesthesia.

KEYWORDS: upper lip bite test, upper lip catch test, prediction, difficult airway, difficult laryngoscopy, difficult intubation.

doi: https://doi.org/10.12669/pjms.344.15364


This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

Difficult laryngoscopy and difficult tracheal intubation occur in 1.5% to 13% of patients undergoing general anesthesia and have always been a concern for anesthesiologists. Different method has been introduced by physician for management of difficult airway. However, the important note is the early and accurate detection of difficult airway for its safe management because failed intubation can have serious consequences and lead to high morbidity and mortality of the patients. Various bedside tests have been used for prediction of difficult laryngoscopy and intubation; of which, upper lip bite test (ULBT) has been proposed by Khan ZH et al as a good predictor for difficult laryngoscopic intubation. However, its
usefulness is not still very clear, as various studies have demonstrated different results regarding its diagnostic accuracy. In a prospective blinded study comparing the ULBT with modified Mallampati test (MMT), ULBT significantly showed higher accuracy and specificity than MMT ($P < 0.001$). However, there were no significant differences in sensitivity, positive and negative predictive values between two tests ($P > 0.05$).

In another study, comparing ULBT with measurement of sternomental distance (SMD), thyromental distance (TMD), and interincisor distance (IID), it was revealed that the specificity and accuracy of the ULBT is significantly higher than the older tests. Also, ULBT, when combined with SMD, showed the highest sensitivity.\(^5\) A study evaluated the role of ULBT, MMT and TMD individually and also in various combinations in prediction of difficult laryngoscopy. Unlike the previous studies, this study showed that none of these three tests is a suitable predictive test when it is used alone. However, higher diagnostic value is achieved when they are combined together.\(^5\) Furthermore, the accuracy and reliability of the ULBT may vary according to patients’ sex and ethnic group; as lip size varies among different ethnicities. In addition, patients with collagen lip injections might show false positives or false negative results.\(^7\)

The aim of this systematic review was to determine the accuracy of the ULBT for predicting difficult airway including difficult laryngoscopy or difficult tracheal intubation. The null hypothesis was that ULBT had poor accuracy for identifying difficult airway.

**METHODS**

*Data sources:* We searched the databases of PubMed, Scopus, and Google scholar for articles published up until October 2016. Key words were selected based on Mesh terms and included “upper lip bite test”, “upper lip catch test”, “prediction”, “difficult airway”, “difficult laryngoscopy”, “difficult intubation”. The manual search of the references of eligible articles for additional studies which were not identified by the electronic search was performed.

*Study selection:* Our inclusion criteria were as followings: prospective observational studies assessing preoperative ULBT to predict difficult intubation in patients undergoing general anesthesia, articles in English language, and studies reporting sensitivity, specificity, NPV, PPV, and accuracy. Albeit, some studies had not reported the accuracy; so, we calculated the accuracy based on the given results, where possible. If there were not enough data for its calculation, the study was not included to the present review. For all studies, Cormack-Lehane grade III and IV was considered as the gold standard. Difficult airway was defined by grade III score in the ULBT and the studies that reported grade II and III as the difficult airway was excluded from the study. The flow diagram of the study is presented in Fig.1.

*Data extraction:* This systematic review was performed based on the guidelines on conducting systematic reviews of diagnostic studies.\(^8\)

**RESULTS**

The results of our search involved 27 studies (9-33) based on our inclusion criteria, as shown in Fig.1. Accuracy, sensitivity, specificity, PPV, and NPV of ULBT for each study are presented in Table-I. The total number of patients included in this systematic review is 18141 aging ≥15 years. Prevalence of airway difficulties in the reviewed studies according to Cormack-Lehane grading and also ULBT grading is presented in Table-I. In 11 of the 27 studies, sensitivity of ULBT in prediction of difficult airway compared to the gold standard was more than 70%. All of the studies except one showed...
<table>
<thead>
<tr>
<th>Study (year)</th>
<th>Number of patients</th>
<th>Prevalence of airway to reference standard (%)</th>
<th>Prevalence of airway to ULBT (%)</th>
<th>Accuracy</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPV</th>
<th>NPV%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Khan et al. (2003)</td>
<td>300</td>
<td>17 (5.7)</td>
<td>45 (15)</td>
<td>88</td>
<td>76.5</td>
<td>88.7</td>
<td>28.9</td>
<td>98.4</td>
</tr>
<tr>
<td>Chohedri et al. (2005)</td>
<td>500</td>
<td>14 (2.8)</td>
<td>10 (2)</td>
<td>96</td>
<td>14.2</td>
<td>98.3</td>
<td>20</td>
<td>97.5</td>
</tr>
<tr>
<td>Hester et al. (2007)</td>
<td>50</td>
<td>9 (18)</td>
<td>6 (12)</td>
<td>90</td>
<td>55</td>
<td>97</td>
<td>83</td>
<td>90</td>
</tr>
<tr>
<td>Allahyary et al. (2008)</td>
<td>200</td>
<td>37 (18.2)</td>
<td>39 (19.2)</td>
<td>97</td>
<td>94.6</td>
<td>97.6</td>
<td>89.7</td>
<td>98.8</td>
</tr>
<tr>
<td>Honarmand et al. (2008)</td>
<td>400</td>
<td>35 (8.75)</td>
<td>17 (4.25)</td>
<td>90</td>
<td>17.1</td>
<td>96.9</td>
<td>35.3</td>
<td>92.2</td>
</tr>
<tr>
<td>Khan et al. (2009)</td>
<td>380</td>
<td>19 (5)</td>
<td>45 (11.8)</td>
<td>91</td>
<td>78.9</td>
<td>91.9</td>
<td>33</td>
<td>98.8</td>
</tr>
<tr>
<td>Myneni et al. (2010)</td>
<td>5999</td>
<td>173 (2.88)</td>
<td>171 (2.85)</td>
<td>94.73</td>
<td>8.1</td>
<td>97.6</td>
<td>8.2</td>
<td>97.6</td>
</tr>
<tr>
<td>Karnjanawanichkul et al. (2010)</td>
<td>400</td>
<td>55 (13.75)</td>
<td>9 (2.25)</td>
<td>85.15</td>
<td>7.14</td>
<td>98.54</td>
<td>44.4</td>
<td>86.7</td>
</tr>
<tr>
<td>Sharma et al. (2010)</td>
<td>62 Acromegaly</td>
<td>15 (24)</td>
<td>9 (14)</td>
<td>74.2</td>
<td>26.7</td>
<td>89.4</td>
<td>44.4</td>
<td>79.2</td>
</tr>
<tr>
<td>Ali et al. (2012)</td>
<td>324</td>
<td>56 (17.3)</td>
<td>68 (21)</td>
<td>91.97</td>
<td>87.5</td>
<td>92.9</td>
<td>71.6</td>
<td>97.3</td>
</tr>
<tr>
<td>Khan et al. (2013)</td>
<td>4500</td>
<td>265 (58.8)</td>
<td>576 (12.8)</td>
<td>90.91</td>
<td>81.5</td>
<td>91.4</td>
<td>37.5</td>
<td>98.7</td>
</tr>
<tr>
<td>Salimi et al. (2008)</td>
<td>350</td>
<td>20 (5.7)</td>
<td>36 (10.3)</td>
<td>92.6</td>
<td>70</td>
<td>93.3</td>
<td>39</td>
<td>98.1</td>
</tr>
<tr>
<td>Shah et al. (2013)</td>
<td>480</td>
<td>67 (13.95)</td>
<td>85 (17.7)</td>
<td>89.16</td>
<td>74.63</td>
<td>91.53</td>
<td>58.82</td>
<td>95.7</td>
</tr>
<tr>
<td>Mohan et al. (2013)</td>
<td>140</td>
<td>15 (10.71)</td>
<td>9 (6.42)</td>
<td>91.42</td>
<td>40</td>
<td>97.6</td>
<td>66.7</td>
<td>93.12</td>
</tr>
<tr>
<td>Shah et al. (2014)</td>
<td>450</td>
<td>47 (10.4)</td>
<td>43 (9.55)</td>
<td>95.5</td>
<td>91.5</td>
<td>96</td>
<td>72.8</td>
<td>98.9</td>
</tr>
<tr>
<td>Srinivasa et al. (2014)</td>
<td>100</td>
<td>NS*</td>
<td>NS</td>
<td>91</td>
<td>77.14</td>
<td>98.46</td>
<td>96.43</td>
<td>88.89</td>
</tr>
<tr>
<td>Mehta et al. (2014)</td>
<td>450</td>
<td>32 (7)</td>
<td>16 (3.5)</td>
<td>95.1</td>
<td>50</td>
<td>98.56</td>
<td>72.77</td>
<td>92.26</td>
</tr>
<tr>
<td>Kolarkar et al. (2015)</td>
<td>300</td>
<td>18 (6)</td>
<td>18 (6)</td>
<td>92.67</td>
<td>100</td>
<td>45</td>
<td>92.2</td>
<td>100</td>
</tr>
<tr>
<td>Honarmand et al. (2015)</td>
<td>600</td>
<td>88 (14.5)</td>
<td>46 (8)</td>
<td>92</td>
<td>48.86</td>
<td>99.41</td>
<td>93.5</td>
<td>91.9</td>
</tr>
<tr>
<td>Javaherforoosh et al. (2015)</td>
<td>448</td>
<td>38 (8.4)</td>
<td>29 (6.47)</td>
<td>89.95</td>
<td>28.9</td>
<td>95.6</td>
<td>37.9</td>
<td>93.5</td>
</tr>
<tr>
<td>Sharma et al. (2015)</td>
<td>150</td>
<td>8 (5.33)</td>
<td>4 (2.66)</td>
<td>94.66</td>
<td>25</td>
<td>98.6</td>
<td>50</td>
<td>95.9</td>
</tr>
<tr>
<td>Vallem et al. (2015)</td>
<td>200</td>
<td>54 (27)</td>
<td>6 (3)</td>
<td>73.5</td>
<td>5.66</td>
<td>97.9</td>
<td>50</td>
<td>74.14</td>
</tr>
<tr>
<td>Min et al. (2016)</td>
<td>243</td>
<td>35 (14.4)</td>
<td>18 (7.4)</td>
<td>84.77</td>
<td>22.9</td>
<td>95.2</td>
<td>44.4</td>
<td>88</td>
</tr>
<tr>
<td>Aswar et al. (2016)</td>
<td>200</td>
<td>16 (8)</td>
<td>13 (6.5)</td>
<td>89.5</td>
<td>25</td>
<td>95.11</td>
<td>30.77</td>
<td>93.58</td>
</tr>
<tr>
<td>Varghese et al. (2017)</td>
<td>199</td>
<td>16 (8)</td>
<td>5 (2.5)</td>
<td>92</td>
<td>18.8</td>
<td>98.9</td>
<td>60</td>
<td>93.2</td>
</tr>
<tr>
<td>Sangeeta et al. (2016)</td>
<td>350</td>
<td>30 (8.57)</td>
<td>36 (10.29)</td>
<td>94.85</td>
<td>80</td>
<td>96.25</td>
<td>66.66</td>
<td>98.08</td>
</tr>
</tbody>
</table>

*NS
a high specificity for ULBT (>85%). Moreover, these studies indicated a high NPV. Accuracy of ULBT was >85% in 24 out of 27 studies.

DISCUSSION

Incidence of a difficult laryngoscopy or endotracheal intubation is reported to vary from 1.5% to 13.1. Difficult or failed intubation is a major cause of related anesthesia mortality.1,3 Therefore, airway management is a considerable challenge in anesthesia and preoperative airway assessment facilitates a very important role in prediction of difficult laryngoscopy. There are many preoperative tests for prediction of difficult intubation. The most common are the Mallampati classification, TMD, SMD, IID and maximum mouth opening test; none of them being ideal compared to direct laryngoscopic view (Gold standard). Due to important roles of the range of freedom of the mandibular movement and the architecture of the teeth in facilitating laryngoscopic intubation, ULBT was introduced by Khan et al as a good predictor for difficult laryngoscopic intubation.4 Taking into account that an ideal test for prediction of difficult airway is the one with high sensitivity and specificity, few false positive predictions and of course, easy to perfomce, different studies have evaluated the diagnostic value of ULBT. The results of these studies are inconsistent. Therefore, we evaluated the accuracy of ULBT for the prediction of difficult airway in this systematic review. The 27 included studies described 18141 patients in whom difficult airway is evaluated by ULBT. The reference test was Cormack-Lehane grading system in all of the studies.11

Prevalence of airway difficulties according to the reference standard varied from 2.8% to 27% and according to the ULBT was from 2% to 21%. Significant variability in sensitivity and specificity was reported by the studies. However, ULBT had an overall high specificity and moderate level of sensitivity in these studies. In 11 out of 27 studies, sensitivity of ULBT in prediction of difficult airway compared to the gold standard was more than 70%.4,5,12-17 The moderate sensitivity of ULBT means that this test will not identify several patients who present with difficult intubation in Cormack-Lehane grading (smaller number of patients with true positive and larger numbers with false negative in ULBT). All studies except one of them showed high specificity for ULBT (>85%). Moreover, these studies indicated a high NPV. These findings is due to high true negative number; indicating high ability of this test to diagnose the patients who do not have difficult airway and therefore is a good test for detection of ease of laryngoscopy. Based on the formula used for accuracy calculation which involves true positive and true negative of patients with difficult airway, a test with high accuracy is an optimal test for prediction of difficult laryngoscopy. We observed a high accuracy of ULBT (>85%) in 24 of 27 studies meaning that ULBT has an optimal diagnostic value in preoperative assessment of patients candidate for general anesthesia.4,5,12-14,17-25

Strength and limitation of the study: The strength of this study is that we reported the findings of studies that compared ULBT with Cormack-Lehane grading, not the ones comparing ULBT with other predictive tests. In addition, we evaluated the accuracy of ULBT used as a single test to achieve precise results; as ULBT has been assessed in combination with other tests in some studies. In these cases, it is not possible to attribute the results to ULBT alone. Because of heterogeneity of the studies, we were not able to conduct a meta-analysis on our findings which is the limitation of the present study.

CONCLUSION

ULBT has moderate sensitivity and PPV, and high specificity, NPV and accuracy. So, it appears that ULBT is a useful bedside test for evaluation of patient airway before the general anesthesia. However, we suggest performing further studies with homogenous patients to achieve more clear results and to carry out a meta-analysis on the results.

Acknowledgment: The authors are grateful for the financial support of Research Vice-Chancellor, Tabriz University of Medical Sciences.

Source of funding: Research vice-chancellor, Tabriz University of Medical Sciences.

Declaration of Interests: None.

REFERENCES


Johnson C. 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. Surv Anaesthesiol. 2004;48:59.


**Author’s Contribution:**

**EF, ZHK & SS:** Conceived, designed and did statistical analysis & editing of manuscript.

**HS & AM:** Did data collection and manuscript writing.

**SS, AM & EF:** Takes the responsibility and is accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

**Authors:**

1. Elnaz Faramarzi
   Assistant Professor,
   Liver and Gastrointestinal Disease Research Center,

2. Hassan Soleimanpour
   Professor of Anesthesiology,
   Department of Emergency Medicine, Faculty of Medicine,

3. Zahid Hussain Khan
   Professor of Anesthesiology, Department of Anesthesiology,
   Faculty of Medicine, Tehran University of Medical Sciences,
   Tehran, Iran.

4. Ata Mahmoodpoor
   Associate Professor of Anesthesiology, Fellowship of Critical Care Medicine, Department of Anesthesiology,
   Faculty of Medicine,

5. Sarvin Saniae
   Assistant Professor, Tuberculosis and Lung Disease Research Center,

1,2,4,5: Tabriz University of Medical Sciences, Tabriz, Iran.