A Proposed Grading System for Post-Intubation Tracheal Stenosis

Ali Ghorbani 1, Azizollah Abbasi Dezfooli 1, Mohammad Behgam Shadmehr 2, Saviz Pejhan 2, Seyed Reza Saghebi 3, Aba-saad Ghare-Daghi 1, Roya Farzanegan 1, Nahid Jahanshahi 1

1 Lung Transplantation Research Center, National Research Institute of Tuberculosis and Lung Diseases (NRITLD), Shahid Beheshti University of Medical Sciences, Tehran, Iran, 2 Tracheal Diseases Research Center, National Research Institute of Tuberculosis and Lung Diseases (NRITLD), Shahid Beheshti University of Medical Sciences, Tehran, Iran, 3 Chronic Respiratory Diseases Research Center, National Research Institute of Tuberculosis and Lung Diseases (NRITLD), Shahid Beheshti University of Medical Sciences, Tehran, Iran

Background: Tracheal stenosis is still a serious consequence of endotracheal intubation. Previous classification systems are commonly descriptive and are not intended to deal with management approach. The aim of this study was to present a classification system for post intubation tracheal stenosis and evaluate its efficacy in distinguishing critically ill patients who need surgical intervention.

Materials and Methods: This classification system was developed based on size and type of stenosis and associated clinical signs and symptoms. Stenosis was graded based on the results of clinical examination and rigid bronchoscopy. All patients received surgical or conservative treatment based on the judgment of a surgeon experienced in management of post-intubation tracheal stenosis without considering their score. ROC curve analysis was done and cut-off point was established based on the greatest Youden index.

Results: Sixty patients were studied. Resection and anastomosis were done for 49 patients. The mean score for all samples was 9.18 (range 8.77-9.45). Chosen cutoff point was 8.5 and calculated sensitivity and specificity were 89% and 42%, respectively. Positive and negative predictive values were 83.7% and 54.5%, respectively. A reasonable agreement between the estimated score and surgeon’s clinical judgment (kappa=0.78) was observed. A statistically significant relationship was observed between scores greater than 8.5 and need for surgical intervention (P=0.007).

Conclusion: We presented a scoring system for post-intubation and tracheostomy tracheal stenosis using main factors influencing diagnosis and treatment and its efficacy was evaluated prospectively. It seems that this system would be capable of assimilating the treatment interventions and comparing them.

Key words: Post-intubation stenosis, Tracheostomy, Classification system

INTRODUCTION

Despite technical improvements and specialized care for the management of critically ill patients, tracheal stenosis is still an important clinical complication following tracheal intubation and tracheostomy (1-3). Patients' quality of life is negatively affected due to dyspnea and other problems like stridor, and respiratory distress; respiratory failure may even occur (4). Various treatment options have been proposed for tracheal stenosis following intubation such as surgery, endoscopic treatments, and
stenoting (5, 6). Characteristics of the stenosis i.e. its location, type, and severity are important factors that should be considered when selecting the treatment method (7). On the other hand, comparison of various treatments without considering the lesion’s characteristics is not possible. Therefore, with respect to all the above, a standard system for description and classification of tracheal stenosis seems necessary. Such classification system will facilitate the selection of treatment method, make the comparison of clinical trials possible and eventually improves patients' quality of life. A limited number of studies have tried to design and present such classification system. In the first study conducted in 1979 classification of tracheal stenosis was performed using Tantalum Tracheography and flow-volume loop. By using this classification system, site of stenosis and its adjacent compliance would be determined and the obtained results could be compared with bronchoscopy findings. Another study in 1983 used acoustic reflection technique along with flow-volume loop and tracheal tomogram for determining the site of stenosis and its dimensions. In 1984, for the first time a system for classification and treatment of tracheal stenosis based on the level of obstruction and percentage of stenosis was proposed.

In 1986 a study evaluated hospital charts and articles and prepared a proposal for a standard reporting system. The 5th study was conducted in 1992 entitled “classification of laryngotracheal stenosis” and evaluated factors affecting treatment. The 6th study was done in 1994 and designed a scoring system for subglottic stenosis using intratracheal tubes. Finally, in a study in 2007 classification of central airway stenosis was done using bronchoscopy. However, none of these studies addressed the efficacy of these classification systems for in-time diagnosis and selection of an appropriate treatment method for patients (7-13).

The present study aimed at presenting a classification system for patients suffering from tracheal stenosis after intubation and evaluating its efficacy in diagnosing critically ill patients requiring surgery.

**MATERIALS AND METHODS**

For the classification of post-intubation tracheal stenoses 3 factors including diameter of the stricture, type of stricture and associated clinical signs and symptoms were used as follows:

A) Diameter of stricture: reduction of tracheal diameter compared to its normal size presented as percentage:
- Score 0: Stenosis rate between 0-25%
- Score 1: Stenosis rate between 26-50%
- Score 2: Stenosis rate between 51-75%
- Score 3: Stenosis rate between 76-90%
- Score 4: Stenosis rate over 91%

Rigid bronchoscope was passed contiguously through the lumen of stricture (determining the tracheal diameter at its narrowest location) to determine the stenosis rate.

B) Type of stenosis: was determined based on the type of tissue of the lesion and included:
- Score 1: Granulation tissue
- Score 2: Granulation tissue, fibrosis and inflammation
- Score 3: Fibrosis
- Score 4: Malacia

Type of lesion was determined by observing it through the rigid bronchoscope’s lumen.

C) Clinical symptoms: were determined through taking a history and physical examination by the clinician:
- Score 1: Dyspnea only during intense activity
- Score 2: Dyspnea during normal activity but physical examination was normal
- Score 3: Long inhalation and exhalation but with no stridor or retraction
- Score 4: Presence of stridor and retraction

Based on this scoring system, each patient received score between score 2 to total score 12.

Patients with idiopathic tracheal stenosis, recurrent tracheal stenosis, direct tracheal trauma and subglottic stenosis were not enrolled.

All patients with tracheal stenosis presenting to Masih Daneshvari Hospital that did not meet the exclusion criteria were entered the study. After obtaining a history
and performing clinical examination by the surgeon, patients underwent rigid bronchoscopy and scoring was done considering all the findings. Medical or surgical treatment was selected and performed for patients based on the physician’s judgment.

Surgical treatment included resection of the lesion and anastomosing the intact free ends to each other.

Maintenance therapy included a course of treatment with anti-inflammatory drugs, performing another rigid bronchoscopy if required and other measures like cold vapor and administration of mucolytic agents.

**Scoring system:** Based on the type and severity of symptoms each patient received a score between 2 to 12. Higher scores were indicative of a worse patient’s condition and possible need for surgery.

**Data analysis:** In order to assess the diagnostic value of the obtained score in determining the need for surgery, clinical judgment of a surgeon expert in the field of airway stenosis was set as the gold standard. ROC curve was drawn for various scores and area under the curve was calculated.

Youden index (sensitivity + specificity) was calculated for all points in the curve and the point with the highest index was accepted as the cut-off point. Prediction reliability of this system was determined based on the reliability index (ratio of correct predictions to all predictions). Kappa index was used to compare the agreement between prediction based on the score and physician’s diagnosis. Quantitative variables were compared using t-test and Mann Whitney U test if needed while qualitative variables were compared using chi square test and Fisher’s exact test.

**RESULTS**

From 2009 to 2010, 60 patients with the mentioned criteria entered the study. Table 1 demonstrates distribution of understudy population based on their demographic and stricture’s characteristics. Based on the physician’s diagnosis, 49 patients required surgery and underwent resection and anastomosis of the lesion.

<table>
<thead>
<tr>
<th>Age(yrs)(mean ± SD)</th>
<th>Total (n=60)</th>
<th>Surgical (n=49)</th>
<th>Medical (n=11)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>41(68.3)</td>
<td>35(71.4)</td>
<td>6(54.5)</td>
<td>0.14</td>
</tr>
<tr>
<td>Female</td>
<td>19(31.7)</td>
<td>14(28.6)</td>
<td>5(45.5)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Decrease in tracheal diameter</th>
<th>Total (n=60)</th>
<th>Surgical (n=49)</th>
<th>Medical (n=11)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;25%</td>
<td>4(6.7)</td>
<td>0(0)</td>
<td>4(36.4)</td>
<td></td>
</tr>
<tr>
<td>26-50%</td>
<td>1(1.7)</td>
<td>1(2)</td>
<td>0(0)</td>
<td></td>
</tr>
<tr>
<td>51-75%</td>
<td>1(1.7)</td>
<td>0(0)</td>
<td>1(9.1)</td>
<td>0.045</td>
</tr>
<tr>
<td>76-90%</td>
<td>36(60)</td>
<td>36(73.5)</td>
<td>3(27.3)</td>
<td></td>
</tr>
<tr>
<td>&gt;90%</td>
<td>18(30)</td>
<td>15(30.6)</td>
<td>3(27.3)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type of lesion</th>
<th>Total (n=60)</th>
<th>Surgical (n=49)</th>
<th>Medical (n=11)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Granulation</td>
<td>1(1.7)</td>
<td>1(2)</td>
<td>0(0)</td>
<td></td>
</tr>
<tr>
<td>Granulation, fibrosis and inflammation</td>
<td>2(3.3)</td>
<td>1(2)</td>
<td>1(9.1)</td>
<td>0.82</td>
</tr>
<tr>
<td>Fibrosis</td>
<td>32(53.3)</td>
<td>25(51)</td>
<td>7(63.6)</td>
<td></td>
</tr>
<tr>
<td>Malacia</td>
<td>5(8.3)</td>
<td>4(8.2)</td>
<td>1(9.1)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Clinical signs and symptoms</th>
<th>Total (n=60)</th>
<th>Surgical (n=49)</th>
<th>Medical (n=11)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dyspnea only during severe activity</td>
<td>5(8.3)</td>
<td>2(4.1)</td>
<td>3(27.3)</td>
<td></td>
</tr>
<tr>
<td>Exertional dyspnea with examination normal</td>
<td>4(6.7)</td>
<td>4(8.2)</td>
<td>0(0)</td>
<td></td>
</tr>
<tr>
<td>Long inhalation and exhalation but with no stridor or retraction</td>
<td>10(16.7)</td>
<td>6(12.2)</td>
<td>4(36.4)</td>
<td>0.009</td>
</tr>
<tr>
<td>Patient has stridor and retraction</td>
<td>41(68.3)</td>
<td>37(75.5)</td>
<td>4(36.4)</td>
<td></td>
</tr>
</tbody>
</table>

The mean age of patients was 27.4±11.4 yrs. A total of 68% were males and 90% of patients had over 75% stenosis. Fibrosis was the most common cause of stenosis. A total of 68% of patients had presented with stridor and retraction of respiratory muscles.

The mean score obtained for all patients was 9.18 (range 8.77-9.58). According to ROC curve analysis the best cut-off point (the minimum number of false positive rate observed) was 8.5 with the sensitivity and specificity of 89% and 42%, respectively. Positive and predictive values for this cut-off point were 83.7% and 54.5%, respectively. The agreement between this system and surgeon’s opinion regarding the need for surgery was 78.3%. Chi square test...
showed a significant association between gaining a score over 8.5 and need for surgery (P=0.007).

**DISCUSSION**

Findings presented here are part of a large study aiming at designing a scoring system to unify the clinical judgment of surgeons, easy reporting of the performed procedures and also comparing different treatment modalities and their results for post-intubation tracheal stenoses.

This study used 3 factors including diameter of the stricture, type of stricture and associated clinical signs and symptoms for classification of these patients. However, Freitag et al. and McCaffrey et al. have included other factors like site of stricture in their scoring system (7, 10).

In Freitag et al. (7) study all types of tracheal stenoses including structural and dynamic types were included and location of stricture and multiple sites were also determined. Classification and analysis were done using simple numerical scoring system. In our study, we focused on post-intubation tracheal stenoses to reach a correct therapeutic judgment.

McCaffrey (10) in his retrospective study evaluated the factors affecting prognosis. Whereas, our study was prospective and aimed at designing an appropriate classification system for choosing the best treatment method in the shortest time period possible.

Sub-glottic and bronchial stenoses were not included in our study because the diagnostic and therapeutic procedures for such strictures are different from the ones indicated for main trachea.

However, Grundfast et al. (9) in their retrospective study on hospital charts and journals evaluated sub-glottic stenoses and offered a proposal for a standard reporting system. In some other systems, length of stricture has also been evaluated (9, 10). However, in the present study, length of stricture was not evaluated despite its importance in prognosis and choosing the treatment of choice because the authors wanted to focus on the simplicity of the system and make its application easy at patient’s bed side. These two factors were not used in the first classification system proposed for tracheal stenoses (8). Number of strictures has also been included in some other systems. In Anand et al. study, number of strictures was included in the classification system as well (11). In our study, in cases with multiple stenoses, in the first phase the most severe one would be selected as our reference and selection of the treatment option was based on that one.

Myer et al. in 1994 used endotracheal tubes of different sizes for determination of the stricture’s diameter (13). Considering the standard sizes of rigid bronchoscope tubes, it seems that application of these tubes in critically ill patients will only waste patient’s time. We used these tubes to determine the diameter of stricture in our study. All the above mentioned studies were retrospective and aimed at comparing the outcome of procedures performed to manage the stricture. Whereas, our study was prospective and aimed at designing a precise scoring system with easy application that can be used at patient’s bed side rapidly and can be helpful for decision making regarding therapeutic procedures. Only one study in 2007 designed such system prospectively and evaluated its validity. However, the mentioned study only evaluated the acceptance of the newly designed system and did not assess its diagnostic value (7).

The present study used all the indices suggested in other studies and has the advantage of evaluating the diagnostic value of the obtained score for detecting patients requiring surgery. Also, it has the potential of being adjusted and assessed in the future for other stenotic areas like the main bronchi and subglottic stenoses.

The cut-off point selected for making the decision of surgery based on our study findings was the score of 8.5 out of 12. With the sensitivity of 89% and specificity of 54%, this point had the ability to differentiate patients requiring surgery from those that can be managed with conservative treatments (P=0.007).

Future studies on a larger sample size can definitely achieve a cut-off point with better diagnostic value. Researchers hoped that the present system can be effective
in unifying the clinical judgments and make the comparison of different treatment modalities possible.

REFERENCES


