Clinical staging for sleep-disordered breathing

MICHAEL FRIEDMAN, MD, HANI IBRAHIM, MD, and LEE BASS, BS, Chicago, Illinois

OBJECTIVE: The purpose of this study was to identify prognostic indicators that would lead to stratification of patients likely to have successful surgery for sleep-disordered breathing (SDB) versus those destined to fail.

STUDY DESIGN: We retrospectively reviewed 134 patients to correlate palate position and tonsil size to the success of the UPPP as based on postoperative polysomnography results. Similar to our previously published data on the Friedman Score as a predictor of the presence and severity of SDB, the palate position was determined on physical examination of the oral cavity and was graded for each patient. This grade combined with tonsil size was used to stage the patients. Stage I was defined as having palate position 1 or 2 combined with tonsil size 3 or 4. Stage II was defined as having palate position 3 or 4 and tonsil size 3 or 4. Stage III patients had palate position 3 or 4 and tonsil size 0, 1, or 2. Any patient with body mass index of greater than 40 was placed in the stage III group. The results of uvulopalatopharyngoplasty (UPPP) were then graded as success or failure and success rates were compared by stage.

SETTING: Academically affiliated tertiary care referral center.

RESULTS: Stage I patients who underwent UPPP had a success rate of 80.6%, stage II patients had a success rate of 37.9%, and stage III patients had a success rate of 8.1%.

CONCLUSION: A clinical staging system for SDB based on palate position, tonsil size, and body mass index is presented. It appears to be a valuable predictor of the success of UPPP. Additional studies and wider use of the staging system will ultimately define its role in the treatment of SDB.

Treatment options available to patients with sleep-disordered (SDB) are limited. Weight loss and behavior modification are helpful in a small percentage of patients. Although continuous positive airway pressure (CPAP) devices are highly successful, their use is limited by poor patient compliance. Patients often turn to surgical treatment when nonsurgical treatment fails. Uvulopalatopharyngoplasty (UPPP) remains the most common surgical procedure performed for SDB. Although reported success rates vary considerably, meta-analysis of the data indicates a success rate of only 40%. In addition to the high failure rate, some studies have indicated that many of the patients who fail to improve actually become worse after UPPP.

The purpose of this study was to identify prognostic indicators that will lead to stratification of patients likely to have successful UPPP versus those who are destined to fail. A staging system may be helpful not only in predicting surgical success but also in providing better descriptions of patients involved in different treatment protocols.

MATERIALS AND METHODS

A retrospective study of 134 patients who underwent UPPP between 1998 and 2000 was undertaken to stratify patients into stages of disease and to assess rates of successful surgery at different clinical stages. Because this was a retrospective chart review study, the institutional review board waived IRB approval. Each patient was staged from information available in the patient’s chart based on clinical findings of patients with SDB previously described by Friedman et al. The key clinical findings in predicting the presence of SDB were found to be palate position with respect to tongue base, tonsil size, and body mass index (BMI).

Palate position had been previously studied and found to be a clinical indicator of SDB. This
Palate classification is based on observations by Mallampati et al., who published a paper on palate position as an indicator of the ease or difficulty of endotracheal intubation by standard anesthesiologist techniques. There are 2 major modifications that we have incorporated into our staging criteria: (1) the anesthesiologist’s assessment is based on the patient sticking out their tongue and the observer then notes the relationship of soft palate to tongue. Our grading is based on the tongue in a neutral, natural position inside the mouth (Fig 1). (2) The original grading system had only 3 grades, and we believe that 4 grades are essential (Figs 1 through 4).
The reason for the first modification is that the tongue during sleep apnea is certainly not related to a protruded position. Therefore we chose to assess the tongue inside the mouth. The reason for adding a fourth grade is that the majority of patients fall into the intermediate grades of II and III, but patients with extreme position (grades I and IV) seem to both have extreme behavior with respect to both presence and treatability of SDB. Because this is a modified palate position grading system, we are identifying it not as Mallampati but rather as Friedman palate position grades I through IV. We do credit Mallampati for bringing this important physical finding to our attention.

Palate grade was assessed as previously described. The procedure involves asking the patient...
to open their mouth widely without protruding their tongue. The procedure is repeated 5 times so that the observer can assign the most accurate level. At times there can be some variation with different examinations, but the most consistent position is assigned as the palate grade. Palate grade I allows the observer to visualize the entire uvula and tonsils or pillars (Fig 1). Palate grade II allows visualization of the uvula but not the tonsils (Fig 2). Palate grade III allows visualization of the soft palate but not the uvula (Fig 3). Palate grade IV allows visualization of the hard palate only (Fig 4).

Tonsil size was graded from 0 to 4. Tonsil size 1 implies tonsils hidden within the pillars (Fig 5). Tonsil size 2 implies the tonsil extending to the pillars (Fig 6). Size 3 tonsils are beyond the pillars but not to the midline (Fig 7). Tonsil size 4 implies tonsils that extend to the midline (Fig 8).

BMI was used as a staging criterion to a limited extent. A BMI greater than 40 was selected as an automatic inclusion into clinical stage III. Because this was a retrospective study, however, very few patients with a BMI of greater than 40 were found to have had UPPP as their surgical treatment.

Stage I disease was arbitrarily defined as those patients with palate position I or II, tonsil size 3 or 4, and BMI of less than 40 (Table 1). Stage II disease is defined as palate position I or II and tonsil size 0, 1, or 2, or palate position III and IV with tonsil size 3 or 4 and BMI of less than 40. Stage III disease is defined as palate position III or IV and tonsil size 0, 1, or 2. All patients with a BMI of greater than 40 were included in stage III disease.

All patients underwent preoperative polysomnography, and only patients who failed conservative therapy underwent surgery. Faculty members and fellows in the department performed the surgery. Only patients who underwent UPPP not combined with other procedures were included in this study. UPPP was performed under general anesthesia using a modified technique previously described. A beveled palatal flap was created in all patients using the dimpling of the soft palate at the site of muscular attachment as a landmark. Closure and reconstruction of the pharynx and palate was modified to include obliteration of the tonsillar fossae space with closure of the muscular layers using Vicryl suture as a submucosal layer. The mucosal edges of the tonsillar fossae and palate were closed with chromic sutures as a second layer. In patients with previous tonsillectomy, the tonsillar fossae mucosa was elevated and removed prior to closure.

Polysomnograms were repeated postoperatively at the same sleep laboratory and compared to the preoperative studies. Patients whose postoperative polysomnography indicated less than 6 hours of sleep were not considered to have a complete study and were therefore excluded. Success of UPPP was defined as a reduction of the postoperative Respiratory Disturbance Index (RDI) to 50% or less of the preoperative value. In addition, the postoperative RDI must have been less than 20. All consecutive patients who underwent UPPP only for treatment of SDB between January 1998 and December 2000 were reviewed. Patients with inadequate data follow-up were excluded. One hundred thirty-four patients qualified with adequate data to be included in the study.

### Statistical Analysis

A consultant statistician performed the statistical analyses. All statistical analyses were performed using SPSS for Windows Version 10.0.7 (SPSS, Inc, Chicago, IL). Continuous data are displayed as mean ± SD. Statistical significance was accepted when \( P < 0.05 \). The paired Student’s \( t \) test was used to compare preoperative with postoperative mean values. The 1-way analysis of variance and Student-Newman-Keuls test were used to identify differences in mean values between stages. The \( \chi^2 \) test was used to test for association between categorical variables such as severity of disease categorized by stage and the

<table>
<thead>
<tr>
<th>Stage</th>
<th>Palate position</th>
<th>Tonsil size</th>
<th>Body mass index (kg/m²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage I</td>
<td>1, 2</td>
<td>3, 4</td>
<td>&lt;40</td>
</tr>
<tr>
<td>Stage II</td>
<td>3, 4</td>
<td>3, 4</td>
<td>&lt;40</td>
</tr>
<tr>
<td>Stage III</td>
<td>3</td>
<td>0, 1, 2</td>
<td>Any</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0, 1, 2</td>
<td>Any</td>
</tr>
<tr>
<td></td>
<td>Any</td>
<td>Any</td>
<td>&gt;40</td>
</tr>
</tbody>
</table>
success or failure of surgical treatment using UPPP. Stepwise multivariate discriminant analysis was used to determine which preoperative indices could best predict operative result. The Fisher’s linear classification equation was constructed for each outcome (unsuccessful and successful). We then performed a validation study by applying the equations casewise to the original 134 patients and comparing predicted versus actual outcome.

RESULTS

We studied 172 consecutive patients who underwent UPPP surgery; however, statistical analysis was restricted to those 134 patients (aged 11 through 71 and mean age of 40.15 ± 13.7) on whom complete data were available. Using previously described criteria, these patients were stratified according to severity of disease into stage I (n = 31, 23.1%), stage II (n = 29, 21.6%), and stage III (n = 74, 55.2%).

Table 2 displays preoperative tonsil size and palate position score (mean ± SD, and range) data according to stage. The mean tonsil sizes for the 3 stages were all significantly different from each other. The mean palate position scores for stages II and III patients were not different from each other but were different from the stage I mean palate position score.

Table 3 contrasts preoperative AHI and minimum Spo2 data with postoperative AHI and minimum Spo2 data according to stage. Stage I preoperative and postoperative AHI values were significantly lower (P < 0.003 and P < 0.005, respectively), and stage I postoperative minimum Spo2 (P < 0.001) was significantly higher than stage II and III measurements. In stage I patients, postoperative AHI and minimum Spo2 were both significantly improved (P < 0.0001 and P < 0.029 for AHI and minimum Spo2, respectively) than those measured preoperatively. In stage II patients, AHI determined postoperatively did not differ from preoperative AHI. However, minimum Spo2 measured during postoperative polysomnography was higher than that measured during preoperative testing (P < 0.045). In stage III patients, neither postoperative mean AHI nor minimum Spo2 differed from preoperative values.

Of the 134 patients, 42 (31.3%) cases were determined to be successful and the remaining 92 (68.7%) cases were considered unsuccessful. Table 4 illustrates the success and failure rates of UPPP for the treatment of SDB according to stage. The \( \chi^2 \) analysis demonstrates a highly significant relationship between stage and success of surgery (Pearson \( \chi^2 = 54.2, \) with 2 df; and a 2-sided P < 0.0001). Successful treatment of SDB with UPPP was most likely achieved in stage I patients (80.6%) and least likely in stage III patients (8.1%).

To further explore the relationship between the severity of SDB (stage) and the efficacy of surgical treatment with UPPP, a stepwise multivariate discriminant analysis was performed. The preoperative criteria used to stratify patients into stages (BMI, tonsil size, and palate position score) were the only indices introduced into the stepwise analysis. The success or failure of surgical treatment with UPPP was used as the categorical end point. Using F values of 3.84 for entry and 2.71 for removal, the stepwise analysis eliminated BMI, keeping tonsil size and palate position score as the best combination of indices for differentiating between success and failure. The classification coefficients calculated for tonsil size and Mallampati score were used to construct the Fisher’s linear classification functional equations. The Fisher’s linear classification equation for each group takes the form:

\[
CF = \text{Tonsils (Coef}_{\text{tonsils}}) + \text{MMP (Coef}_{\text{MMP}}) + \text{Constant}
\]
Where CF is group classification function, Tonsils is tonsil size, Coef tonsils is group classification coefficient for tonsil size, FPP is Friedman Palate Position score, (Coef FPP) is group classification coefficient for palate position score, and Constant is group classification constant.

A separate equation is constructed for each result, unsuccessful and successful. In the present case:

Unsuccessful result = [(Tonsils) * 0.870] + [(FPP) * 5.319] + (10.563)

Successful result = [(Tonsils) * 2.284] + [(FPP) * 2.333] + (6.001)

To predict the success of UPPP surgery in patients with SDB, enter the patient’s tonsil size and palate position into each of the above formulas and calculate. The equation totaling the numerical highest value is the predicted result. In the validation study, the above formulas were applied casewise to the 134 patients and correctly predicted 95.0% of the cases by result.

**DISCUSSION**

UPPP is the most common and in many situations the only surgical procedure performed by most otolaryngologists for the treatment of SDB. Many studies have documented 3 important issues that must be considered in recommending the surgical procedure to a patient. (1) A meta-analysis of unselected patients treated with UPPP revealed that only 40.79% of patients had “successful” surgery defined by an Apnea Hypopnea Index (AHI) reduction of 50% and a postoperative AHI of less than 20 or an Apnea Index (AI) reduced by 50% and a postoperative AI of less than 101. (2) Despite some data indicating that preoperative selection criteria may identify those patients likely to fail, prior to this study there have been no clear cut, reproducible physical findings that have been shown to consistently help in the selection process. (3) A study published by Senior et al demonstrated that UPPP not only does not cure SDB in 60% of cases but also often makes it worse. It has been a common misconception to assume that although UPPP has only a 40% success rate, the responders would be those with mild disorders. Therefore, the procedure is often recommended for patients with mild and moderate SDB. Senior et al has demonstrated that within this

<table>
<thead>
<tr>
<th>Stage</th>
<th>Unsuccessful</th>
<th>Successful</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>6 (19.4)</td>
<td>25 (80.6)*</td>
<td>31 (100)</td>
</tr>
<tr>
<td>II</td>
<td>18 (62.1)</td>
<td>11 (37.9)*</td>
<td>29 (100)</td>
</tr>
<tr>
<td>III</td>
<td>68 (91.9)</td>
<td>6 (8.1)*</td>
<td>74 (100)</td>
</tr>
</tbody>
</table>

Table 4. Success rate of uvulopalatopharyngoplasty in the treatment of sleep-disordered breathing

Patients were stratified according to severity of disease based on criteria from Friedman et al. (1999). Values given as number (percent).

*Significant differences than all other stages.

Where CF is group classification function, Tonsils is tonsil size, Coef tonsils is group classification coefficient for tonsil size, FPP is Friedman Palate Position score, (Coef FPP) is group classification coefficient for palate position score, and Constant is group classification constant.

A separate equation is constructed for each result, unsuccessful and successful. In the present case:

Unsuccessful result = [(Tonsils) * 0.870] + [(FPP) * 5.319] + (10.563)

Successful result = [(Tonsils) * 2.284] + [(FPP) * 2.333] + (6.001)

To predict the success of UPPP surgery in patients with SDB, enter the patient’s tonsil size and palate position into each of the above formulas and calculate. The equation totaling the numerical highest value is the predicted result. In the validation study, the above formulas were applied casewise to the 134 patients and correctly predicted 95.0% of the cases by result.

**DISCUSSION**

UPPP is the most common and in many situations the only surgical procedure performed by most otolaryngologists for the treatment of SDB. Many studies have documented 3 important issues that must be considered in recommending the surgical procedure to a patient. (1) A meta-analysis of unselected patients treated with UPPP revealed that only 40.79% of patients had “successful” surgery defined by an Apnea Hypopnea Index (AHI) reduction of 50% and a postoperative AHI of less than 20 or an Apnea Index (AI) reduced by 50% and a postoperative AI of less than 101. (2) Despite some data indicating that preoperative selection criteria may identify those patients likely to fail, prior to this study there have been no clear cut, reproducible physical findings that have been shown to consistently help in the selection process. (3) A study published by Senior et al demonstrated that UPPP not only does not cure SDB in 60% of cases but also often makes it worse. It has been a common misconception to assume that although UPPP has only a 40% success rate, the responders would be those with mild disorders. Therefore, the procedure is often recommended for patients with mild and moderate SDB. Senior et al has demonstrated that within this
subgroup the risk of failure and the risk of aggravating the disease are extremely high. These findings are consistent with our own observations and data. It has also been shown to be true for patients treated with laser-assisted uvulopalatoplasty (LAUP). The procedure not only fails 60% of the time, but often makes the condition worse.

Surgery with a 40% success rate is certainly less than ideal. Our ultimate goal is, of course, to develop a treatment with a high success rate. In the absence of that treatment, however, our goal should be to identify those patients who are likely to benefit from UPPP, which is a valuable procedure for those patients who can be cured with it. The ideal identification process would identify those patients with a high likelihood of success of UPPP versus those with a high likelihood of failure and therefore a need to treat other areas of the upper airway. The ideal selection process would be noninvasive, cost effective, and reproducible. We propose that our staging system satisfies these criteria for an ideal prediction assessment for some patients (stage I and stage III) who can then be guided to appropriate treatment. Patients with stage I disease have better than an 80% chance of success with UPPP and should therefore undergo the procedure. Patients with stage III disease should never undergo UPPP alone as a surgical cure of SDB. With an 8.1% success rate, they are destined to fail. They should be treated with a combination of procedures that address both the palate and hypopharynx. In our study 78.3% of patients can be stratified to stage I or III. Patients with stage II disease do not fall into either extreme but probably should be treated similar to stage III patients.

The failure of UPPP to cure SDB has been clearly associated with sites of obstruction in the upper airway not corrected by the procedure, such as the hypopharynx. Fujita originally described multiple levels of obstruction. Riley et al have demonstrated these abnormalities with cephalometric data. Routine use of preoperative cephalometric studies has not been shown to be worthwhile in the selection criteria. The complexity of the studies combined with conflicting reports on its value in treatment selection has relegated its position to a research rather than clinical tool.

Numerous other methods have been used to predict the location of the upper airway obstruction. These include physical examination, computed tomography, magnetic resonance imaging, and fluoroscopy. As with cephalometric studies, these studies are all valuable in research studies but have not been shown to be of clinical value. The most commonly used test is the Muller Maneuver (MM). Borowiecki and Sassin first described this maneuver for the preoperative assessment of SDB. The MM consists of having the patient perform a forced inspiratory effort against an obstructed airway with fiberoptic endoscopic visualization of the upper airway. The test is widely used and simple to perform. Despite this, its use is controversial and certainly no studies have been able to associate the maneuver as a tool to select patients who are likely to succeed with UPPP.

Criticism of the test is based on 3 areas. One criticism of the test is that it is subjective. Terris et al have clearly shown that in their institution the maneuver can easily be taught to residents at all levels of training and is clearly reproducible. Their institution, however, is highly focused on sleep disorders and treats a large number of patients on a regular basis. The level of reproducibility may be reduced by the general otolaryngologist and by residents seeing patients with SDB less frequently.

Another criticism is that many patients are incapable of producing a full force inspiratory effort. Patient compliance can vary significantly from patient to patient and even from examination to examination with the same patient. Any test that relies on patient cooperation has some level of variability.

The third area of criticism is whether the use of MM helps predict success of UPPP. Terris et al and Sher et al were the first investigators who suggested that the use of MM for patient selection could increase the success of UPPP. Other researchers, however, have found that the MM is not helpful in predicting the success of UPPP. In addition, we have found that patients with minimal collapse of the hypopharynx as determined by MM before UPPP may have moderate or severe collapse in the tongue base (by MM) after UPPP. Katsantonis et al found that the prediction efficacy of the MM was only 33% in selected patients.
Doghamjii et al\textsuperscript{13} similarly found no benefit of MM in the prediction value for success of UPPP. Other reports substantiate these negative findings questioning the reproducibility of MM as a predictive tool.\textsuperscript{14}

The staging system for head and neck cancer is not a substitute for detailed evaluation through clinical examination and radiological studies, but it provides stratification of patients that helps in treatment selection and helps in assessment of results. Similarly, we are recommending a staging system for patients with SDB that will help create reproducible physical data and help in treatment selection. The treating physician can continue to use vague and nonreproducible terms such as “low soft palate,” “thick soft palate,” “crowded oropharynx,” etc, but the staging system should be used in addition to these terms.

This study is a retrospective study and therefore has some weaknesses. The study was a consecutive series of all patients who underwent UPPP. It is not a consecutive study of all patients with SDB. No patient was eliminated from having a UPPP because of palate position. Patients suspected of having obstruction at more than one level were treated in stages and only the results of UPPP were assessed for this study. Despite this, some patients may have been eliminated from the study by physician bias if they were not candidates for UPPP as part of their treatment. Specifically, patients with BMI of greater than 40 and with pickwickian syndrome often had a tracheotomy and were therefore eliminated from the study. Similarly, patients with obvious severe skeletal or anatomic deformities, such as severe micrognathia, or patients previously treated for head and neck cancer were eliminated from this study. We cannot therefore evaluate BMI (>40) as an independent variable or as a contributing variable in predicting success of UPPP. In patients with a BMI of less than 40, however, the BMI does not independently affect outcome.

Ideally the goal of surgery for SDB would be procedures that result in 80% success or better. In the absence of that simple procedure, stratification of patients that allows for preoperative selection will in effect result in 80% success. Obviously, only 25% of patients suffering from SDB can achieve this result with UPPP as an isolated procedure. Seventy-five percent of patients should not be treated with UPPP as an isolated curative surgical procedure. Patients with stage II and stage III disease should have adjunctive procedures to address their hypopharynx. At the minimum, these patients should have radiofrequency tongue base reduction. Depending on surgeon bias and patient selection, some of them may require tongue base advancement procedures or maxillary mandibular advancement procedures. If patients with stage III disease are treated with UPPP alone, they should be forewarned that they likely will need secondary procedures since the likelihood of initial success is less than 10%.

**CONCLUSIONS**

A staging system for SDB is presented based on palate position, tonsil size, and BMI. Patients with stage I disease have an excellent chance of successful surgery with UPPP. Patients with stage II and stage III disease have a poor chance for success with UPPP. If surgery is considered for patients with stage II or stage III disease, it should not be UPPP only. UPPP should offer selected patients an 80% chance of a cure.

**REFERENCES**