Physiology-Based Modeling May Predict Surgical Treatment Outcome for Obstructive Sleep Apnea

Yanru Li, MD; Jingying Ye, MD; Demin Han, MD, PhD; Xin Cao, MD; Xiufeng Ding; Yuhuan Zhang; Wen Xu, MD; Jeremy Orr, MD; Rachel Jen, MD; Scott Sands, PhD; Atul Malhotra, MD; Robert Owens, MD

1Department of Otolaryngology Head and Neck Surgery, Beijing Tongren Hospital, Capital Medical University; Key Laboratory of Otolaryngology Head and Neck Surgery (Ministry of Education of China), Beijing, China; 1Pulmonary and Sleep Division, University of California at San Diego, California; 2Department of Otolaryngology Head and Neck Surgery, Beijing Tsinghua Changgung Hospital, Tsinghua University, Beijing, China; Division of Sleep and Circadian Disorders, Brigham and Women’s Hospital and Harvard Medical School, Boston, Massachusetts; 2Department of Allergy Immunology and Respiratory Medicine and Central Clinical School, The Alfred and Monash University, Melbourne, Australia

Study Objectives: To test whether the integration of both anatomical and nonanatomical parameters (ventilatory control, arousal threshold, muscle responsiveness) in a physiology-based model will improve the ability to predict outcomes after upper airway surgery for obstructive sleep apnea (OSA).

Methods: In 31 patients who underwent upper airway surgery for OSA, loop gain and arousal threshold were calculated from preoperative polysomnography (PSG). Three models were compared: (1) a multiple regression based on an extensive list of PSG parameters alone; (2) a multivariate regression using PSG parameters plus PSG-derived estimates of loop gain, arousal threshold, and other trait surrogates; (3) a physiological model incorporating selected variables as surrogates of anatomical and nonanatomical traits important for OSA pathogenesis.

Results: Although preoperative loop gain was positively correlated with postoperative apnea-hypopnea index (AHI) (P = .008) and arousal threshold was negatively correlated (P = .011), in both model 1 and 2, the only significant variable was preoperative AHI, which explained 42% of the variance in postoperative AHI. In contrast, the physiological model (model 3), which included AHIRem (anatomy term), fraction of events that were hypopnea (arousal term), the ratio of AHIREM and AHINREM (muscle responsiveness term), loop gain, and central/mixed apnea index (control of breathing terms), was able to explain 61% of the variance in the physiology of postoperative AHI.

Conclusions: Although loop gain and arousal threshold are associated with residual AHI after surgery, only preoperative AHI was predictive using multivariate regression modeling. Instead, incorporating selected surrogates of physiological traits on the basis of OSA pathophysiology created a model that has more association with actual residual AHI.

Commentary: A commentary on this article appears in this issue on page 1023.

Clinical Trial Registration: ClinicalTrials.Gov; Title: The Impact of Sleep Apnea Treatment on Physiology Traits in Chinese Patients With Obstructive Sleep Apnea; Identifier: NCT02696629; URL: https://clinicaltrials.gov/show/NCT02696629

Keywords: obstructive sleep apnea, upper airway surgery, loop gain, critical pressure, polysomnography, phenotyping, lung


INTRODUCTION

Obstructive sleep apnea (OSA) is an extremely common disorder with major cardiometabolic and neurocognitive sequelae. Although continuous positive airway pressure is highly efficacious for OSA, its effectiveness is limited by variable adherence to therapy. As a result, there is considerable interest in alternative therapies for OSA. Although there are existing alternative treatments that might be more acceptable to patients, such as oral appliances or upper airway surgery, enthusiasm for these approaches is limited by variable efficaciousness.

For example, velopharyngeal surgery is highly effective in certain patient subgroups, but yields more variable results in others. Additional testing to characterize individuals might improve the ability to predict response to therapy; however, these approaches have not been adopted into wider clinical use.

We and others have argued that OSA physiology is multifactorial; anatomical as well as nonanatomical factors are important in its pathogenesis. Thus, in addition to a collapsible upper airway, other elements such as upper airway muscle responsiveness, ventilatory control (quantified as loop gain), and sleep state stability as quantified by the respiratory arousal.
Anatomical and nonanatomical traits have been reported.\textsuperscript{17–19} For example, a high loop gain means relatively high ventilatory response to a given blood gas disturbance, which will persistently destabilize breathing. Recently, noninvasive methods to evaluate these surrogates of anatomical and nonanatomical traits have been reported.\textsuperscript{17–19}

A comprehensive physiological model that incorporates all of the aforementioned traits may have utility for determining responsiveness to non-positive airway pressure therapy, allowing a personalized approach.\textsuperscript{16}

Based on this conceptual framework, we hypothesized that we could predict response to non-positive airway pressure therapies using measurement of both anatomical and nonanatomical traits, if combined in a physiological model. In order to test this hypothesis, we studied a relatively homogeneous cohort of patients with OSA who underwent upper airway surgery, and who were evaluated with preoperative and postoperative polysomnography. To determine the value of our approach, we tested 3 models to estimate postoperative apnea-hypopnea index (AHI):

1. A model using traditionally identified demographic and baseline polysomnography (PSG) metrics as independent variables in a multiple linear regression, called “PSG Multiple Regression”
2. A model using these traditional metrics plus PSG-derived surrogates of physiological traits, such as loop gain and arousal threshold, as independent variables in a multiple linear regression, called “PSG Plus Multiple Regression”
3. A model incorporating the same metrics as in model 2, but forced to include scores based on severity categories for anatomical and nonanatomical factors, called “PSG Plus Physiology”

**METHODS**

**Study Population**

The study population consisted of 31 Chinese adult patients who underwent upper airway surgery for OSA. Data from this cohort have previously been presented (eg, demographic data); however, the current aims, data, and results are novel. Inclusion criteria for the subjects consisted of the following:

1. OSA had been diagnosed in all subjects based on a PSG study (AHI > 5 events/h) and typical clinical symptoms (such as snoring, witnessed apneas, and daytime sleepiness).
2. No other surgical procedures (such as nasal or glossopharyngeal surgical procedures) were performed prior to or during the current treatments.
3. Positive airway pressure treatment was recommended before the current treatment but was refused or the patients were unable to tolerate positive airway pressure treatment.
4. All of the subjects underwent clinical evaluation of upper airway anatomy by one surgeon. They met the indications for velopharyngeal surgery for treating OSA (Chinese guideline for the diagnosis and surgical treatment of obstructive sleep apnea hypopnea syndrome, 2009). Patients with severe maxillofacial or mandibular deformities were excluded.
5. Patients who could accept nasal continuous positive airway pressure therapy, or who had a serious coexisting lung, neurological, cardiovascular, or psychiatric disorder were excluded.

We excluded subjects based on the preoperative polysomnogram if they had less than 30 minutes of stage R sleep.

**Study Design**

This was a retrospective observational cohort study. Surrogates of physiological traits such as loop gain and arousal threshold were assessed from preoperative PSG (see next paragraphs). The association between the surrogates of physiological traits and residual AHI after surgery was analyzed. The efficacy of the 3 models in estimating postoperative AHI was then compared.

**Polysomnography**

PSG was performed according to a standard protocol: electroencephalograms, electrooculograms, and surface electromyograms were applied to score arousals, leg movements, and sleep stage. Abdominal and chest movements, pulse oxygen saturation, and nasal flow were recorded to detect respiratory events. The tests were analyzed according to the scoring guidelines of American Academy of Sleep Medicine Task Force 2007 criteria (hypopneas were defined as events with nasal pressure signal dropping ≥ 50% and ≥ 3% desaturation from baseline or an arousal).\textsuperscript{20} Sleep studies were performed prior to surgery and postoperatively during follow-up.

**Upper Airway Surgery**

Surgical procedures were performed by one surgeon using previously reported techniques, including revised uvulopalatopharyngoplasty (UPPP) with uvula preservation (H-UPPP)\textsuperscript{21} with and without concomitant transpalatal advancement (TA) pharyngoplasty,\textsuperscript{22} genioglossus advancement, or hyoid suspension.

The definition of responders for surgical treatment, defined as a ≥ 50% reduction in AHI to a final AHI of < 20 events/h, was used. Moreover, a postoperative AHI below 10 events/h was considered to be cured.

**Estimation of Physiological Traits**

**Loop Gain**

The stability of the ventilatory control system (loop gain) was quantified by fitting a simplified mathematical model to the spontaneous ventilatory pattern obtained via PSG.\textsuperscript{17} Ventilatory fluctuations are estimated using the square-root transformed nasal pressure waveform. Loop gain was reported as the ventilatory response to a 1 cycle/min disturbance as used previously. Calculations were performed using MATLAB (MathWorks, Natick, Massachusetts, United States). The median value for loop gain, using all available 7-minute windows of non-stage R sleep, was used to provide a single value for each individual for each sleep study.

**Arousal Threshold**

Respiratory arousal threshold was calculated using a previously validated regression equation,\textsuperscript{18} provided by the authors,
which uses baseline AHI, nadir desaturation, and the fraction of events that are hypopneas in order to estimate the epiglottic (eg, downstream, intrathoracic) pressure that typically precedes arousal from sleep.

Development of the Models
For all models, postoperative AHI was chosen as the dependent variable. Possible independent variables for inclusion in the models were chosen based on prior literature. In similar fashion, possible PSG-derived surrogates of underlying physiology traits were chosen from the literature, or based on physiological rationale. Candidate variables and the surrogates of physiology traits for further analysis are listed in Table 1. For the PSG Multiple Regression and PSG Plus Multiple Regression models, we performed the same analysis, but used the change in AHI from preoperative to postoperative as our dependent variable.

Models

**PSG MULTIPLE REGRESSION:** Univariate linear regression was used to determine the variance that could be explained by the variables. Variables were considered for the multiple linear regressions if they had an $r^2 \geq .2$ in the univariate linear regression, and were then used in the multiple linear regressions using a backward stepwise procedure. The final model was composed of significant independent variables determined by the regression.

**PSG PLUS MULTIPLE REGRESSION:** In addition to the aforementioned independent variables, PSG-derived surrogates of physiological traits were also included in the candidate parameters. Again, univariate linear regression was used to determine variables to be considered for the multiple linear regression, at a cutoff value of $r^2 > .2$. Similarly, the final model was composed of the significant variables determined by the regression.

---

**PSG PLUS PHYSIOLOGY:** Development of the physiological model included the following steps:

1. Possible surrogates of the physiological traits were considered. All candidate variables thought to represent 1 of the 4 underlying physiological traits were drawn from the literature, or based on opinion of the authors.
2. Correlations between these baseline surrogates and postoperative AHI were examined. Parameters were further considered if they showed significant correlation with postoperative AHI (either Pearson correlation test or Spearman rank correlation test as appropriate, $P < .05$).
3. Correlated surrogates were categorized according to either previously reported physiological cutoffs, or according to cutoffs that best discriminated between cures (AHI < 10 events/h) and failures (AHI > 20 events/h). These cutoffs defined good (0 points), intermediate (1 point), or bad values (2 points) for each surrogate. As stated, physiological traits may not affect the possibility of residual sleep apnea in a linear manner; thus, we categorized each candidate parameter to more clearly explain the interactions of the traits. In short, the cutoff points of each parameter were decided according to either previously reported physiological cutoffs, or grouped into good, intermediate, or poor based on the calculated receiver operating characteristic (ROC) for a residual AHI > 10 and AHI > 20 events/h (see the supplemental material for...
The Effectiveness of the Models in Estimating Residual Sleep Apnea

The correlations between the estimated values derived from each model and the actual postoperative AHI were assessed. ROC for a residual AHI greater than 10 events/h (not cured), greater than 20 events/h and greater than 30 events/h (unsuccessful surgery) were calculated for each of the model for comparison.

Statistical Analysis

Statistical analyses were performed using SPSS 22.0 (SPSS Inc., Chicago, Illinois, United States). Sleep apnea severity before and after surgery was compared using the Wilcoxon signed-rank test. The difference of the variables between responders and nonresponders were compared using Mann-Whitney U test. Univariate linear regression and Spearman correlation coefficient were used to identify significant associations. Statistical significance was set at \( P < .05 \).

RESULTS

All the patients underwent revised H-UPPP,\(^2\) 13 had concomitant TA pharyngoplasty.\(^2\) One patient underwent concomitant genioglossus advancement and one had hyoid suspension. The median (interquartile range) interval between surgery and postoperative PSG was 4 (2, 7) months. The overall severity of OSA improved significantly after surgery (59.1 [34.3, 71.3] versus 18.7 [10.8, 42.3] events/h; \( Z = -4.801, P < .001 \)). There were no significant differences in delta AHI (\( Z = -0.810, P = .417 \)) or postoperative AHI (\( Z = -0.110, P = .913 \)) between the patients who underwent H-UPPP and those had H-UPPP + TA.

A > 50% reduction in AHI to a final AHI of <20 events/h was achieved in 15 patients (48.4%) (defined as responders), 7 were considered cured with a postoperative AHI <10 events/h (22.6%). Sixteen subjects were considered nonresponders and 11 (35.5%) still had severe sleep apnea after surgery. The characteristics of all the participants, responders, and nonresponders are listed in Table 2. The responders had significantly lower loop gain and less severe sleep apnea (lower AHI) than did the nonresponders (\( P < .05 \)).

Baseline Loop Gain, Arousal Threshold and Surgical Outcome

Both loop gain and estimated arousal threshold showed significant associations with postoperative AHI. Univariate

---

### Table 2—Characteristics of all the participants, responders, and nonresponders.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>All Subjects (n = 31)</th>
<th>Responders (n = 15)</th>
<th>Nonresponders (n = 16)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td>42 (38, 49)</td>
<td>48 (41, 49)</td>
<td>40 (35, 46)</td>
</tr>
<tr>
<td>Sex, male/female</td>
<td>31/0</td>
<td>15/0</td>
<td>16/0</td>
</tr>
<tr>
<td>Body mass index, kg/m(^2)</td>
<td>27.1 (26.2, 29.4)</td>
<td>27.5 (26.2, 29.0)</td>
<td>26.8 (25.9, 29.8)</td>
</tr>
<tr>
<td>Neck circumference, cm</td>
<td>42.0 (40.0, 43.0)</td>
<td>42.0 (41.0, 43.0)</td>
<td>42.0 (40.0, 42.0)</td>
</tr>
<tr>
<td>AHI, events/h</td>
<td>59.1 (34.3, 71.3)</td>
<td>44.9 (29.1, 69.3)</td>
<td>67.3 (57.7, 73.6)</td>
</tr>
<tr>
<td>AHI(_{\text{REM}}), events/h</td>
<td>66.7 (37.5, 74.5)</td>
<td>47.3 (20.7, 74.1)</td>
<td>70.6 (60.5, 75.4)</td>
</tr>
<tr>
<td>AHI(_{\text{NREM}}), events/h</td>
<td>55.5 (40.3, 62.0)</td>
<td>46.8 (28.2, 59.6)</td>
<td>59.0 (52.1, 63.4)</td>
</tr>
<tr>
<td>Central apnea index, events/h</td>
<td>0.8 (0.3, 1.9)</td>
<td>0.8 (0.3, 1.3)</td>
<td>1.2 (0.3, 4.2)</td>
</tr>
<tr>
<td>Mixed apnea index, events/h</td>
<td>2.7 (0.9, 7.4)</td>
<td>1.9 (0.1, 5.2)</td>
<td>6.0 (1.1, 12.0)</td>
</tr>
<tr>
<td>Fraction of events that were hypopnea, %</td>
<td>4.3 (0.6, 22.5)</td>
<td>9.8 (0.6, 37.7)</td>
<td>1.9 (0.5, 6.3)</td>
</tr>
<tr>
<td>Fraction of events that were obstructive apnea, %</td>
<td>76.9 (57.4, 90.8)</td>
<td>76.9 (51.5, 90.1)</td>
<td>77.5 (59.2, 91.2)</td>
</tr>
<tr>
<td>Sleep time with oxygen saturation ≤ 90%, %</td>
<td>18.4 (3.5, 29.4)</td>
<td>7.0 (2.0, 20.0)</td>
<td>25.8 (4.9, 33.8)</td>
</tr>
<tr>
<td>Nadir oxygen saturation during sleep, %</td>
<td>70.0 (66.0, 83.0)</td>
<td>77.0 (63.0, 85.0)</td>
<td>69.0 (66.3, 78.5)</td>
</tr>
<tr>
<td>Estimated arousal threshold, cm H(_2)O</td>
<td>-26.9 (−32.4, −19.4)</td>
<td>-25.0 (−32.8, −15.5)</td>
<td>-30.4 (−32.2, −23.6)</td>
</tr>
<tr>
<td>Low arousal threshold (estimated arousal threshold greater than −15 cm H(_2)O versus high arousal threshold</td>
<td>4/27</td>
<td>3/15</td>
<td>1/16</td>
</tr>
<tr>
<td>Loop gain</td>
<td>0.71 (0.58, 0.81)</td>
<td>0.58 (0.52, 0.77)</td>
<td>0.78 (0.64, 0.86)</td>
</tr>
<tr>
<td>Ratio of supine versus nonsupine AHI</td>
<td>1.45 (1.05, 3.09)</td>
<td>2.12 (1.02, 4.40)</td>
<td>1.14 (1.05,1.70)</td>
</tr>
</tbody>
</table>

Values presented as median (interquartile range). * = \( P < .05 \) between responders and nonresponders. AHI = apnea-hypopnea index, AHI\(_{\text{REM}}\) = apnea-hypopnea index during REM sleep, AHI\(_{\text{NREM}}\) = apnea-hypopnea index during non-REM sleep.
regression showed that high loop gain is positively correlated with the postoperative AHI ($r = .465$, $P = .008$) whereas low arousal threshold is negatively correlated with the postoperative AHI ($r = -.453$, $P = .011$).

**Models for Estimating Postoperative AHI**

**PSG Multiple Regression and PSG Plus Multiple Regression**

Multiple regression analysis using a backward method suggested that preoperative AHI was the only significant independent variable associated with postoperative AHI in both PSG Multiple Regression model and PSG Plus Multiple Regression model ($r^2 = .427$; adjusted $r^2 = .406$, $P < .001$; $\beta = 0.555$, SE of $\beta = 0.124$; standard $\beta = 0.653$, $P < .001$). Figure 1A shows the scatter plot of estimated AHI using regression model and the actual postoperative AHI. ROC curves of using physiological model scores to predict postoperative AHI ≥ 10 events/h, AHI ≥ 20 events/h, and AHI ≥ 30 events/h are shown in Figure 2. Similarly, when considering delta AHI as the dependent variable, a marker of preoperative severity—the 3% oxygen desaturation index—was the only significant independent variable for both the PSG Multiple Regression and PSG Plus Multiple Regression models ($r^2 = .340$; adjusted $r^2 = .316$, $P < .001$; $\beta = 0.436$, standard error of $\beta = 0.117$; standard $\beta = 0.583$, $P = .001$).
PSG Plus Physiology Model

The final model with the highest coefficient of determination included AHI_{REM} (anatomy), fraction of events that were hypopnea (arousal term), central or mixed apnea index, loop gain (both control of breathing), and ratio of AHI_{REM} versus AHI_{NREM} (upper airway dilator muscle responsiveness term).

Note that two separate parameters considered surrogates of control of ventilation were retained in the physiology model. The details of the surrogate trait cutoffs are listed in Table 3.

The final equation for the physiology model is a sum of the points from all the surrogate measures. For example, consider a patient with an AHI_{REM} of 35 events/h, fraction of events that were hypopnea of 40%, AHIREM ≥ 130% of AHI_{NREM}, loop gain of 0.7, both central and mixed apnea index less than 5 events/h. The patient’s score will be 1 + 0 + 0 + 1 + 0 = 2. The physiology model could explain 61% of the variance in postoperative AHI (r² = .607; adjusted r² = .593, P < .001; β = 5.530, standard error of β = 0.826; standard β = 0.779, P < .001). The estimated postoperative AHI equation was: AHI = −5.898 + 5.530 × the physiological score. Figure 1B shows the scatterplot of estimated AHI using the physiology model and the actual postoperative AHI. ROC curves that demonstrate the association between the model and actual postoperative AHI ≥ 10 events/h, ≥ 20 events/h or ≥ 30 events/h are shown in Figure 2 (solid lines).

Table 3—Parameters in the physiology model and their cutoffs.

<table>
<thead>
<tr>
<th></th>
<th>0 Points</th>
<th>1 Point</th>
<th>2 Points</th>
<th>Standard β Values</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Anatomy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AHI_{REM}, events/h</td>
<td>&lt; 20</td>
<td>20–50</td>
<td>≥ 50</td>
<td>.438</td>
</tr>
<tr>
<td><strong>Arousals</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction of events that were hypopnea, %</td>
<td>≥ 20</td>
<td>10–20</td>
<td>&lt; 10</td>
<td>.595</td>
</tr>
<tr>
<td><strong>Control of Breathing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loop gain</td>
<td>&lt; 0.6</td>
<td>0.6–1.0</td>
<td>&gt; 1.0</td>
<td>.536</td>
</tr>
<tr>
<td>Central/mixed apnea index, events/h</td>
<td>&lt; 5</td>
<td>≥ 5</td>
<td></td>
<td>.655</td>
</tr>
<tr>
<td><strong>Muscle Responsiveness</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AHI_{REM} versus AHI_{NREM}</td>
<td>AHI_{REM} ≥ 130% AHI_{NREM}</td>
<td>130% AHI_{NREM} &lt; AHI_{REM} &lt; 130% AHI_{NREM}</td>
<td>AHI_{NREM} ≥ 130% AHI_{REM}</td>
<td>.582</td>
</tr>
</tbody>
</table>

0 points indicates a relatively good trait. 1 point indicates an intermediate trait and ≥ 2 points indicate a bad trait for passive anatomy, arousal threshold or upper airway muscle or control of breathing. AHI = apnea-hypopnea index, AHI_{NREM} = apnea-hypopnea index during non-REM sleep, AHI_{REM} = apnea-hypopnea index during REM sleep.

DISCUSSION

Our findings are important for a number of reasons. We show that consideration of both anatomical and non-anatomical traits can improve the ability to anticipate the response to non-positive airway pressure therapy (in this case upper airway surgery). Moreover, we show that an understanding of physiology, rather than unguided reliance on statistical methods, should be used to incorporate this additional information into a model. Although our results may be specific to the population and intervention we studied, they provide proof of concept to this approach that might be used in other populations and with other interventions.

The Value of Estimates of Nonanatomical Traits for Upper Airway Surgery

An elevated loop gain at baseline is associated with poor surgical outcome in the current study. Patients with unstable ventilatory control may have primarily nonanatomical factors underlying OSA and thus may not be ideal candidates for upper airway surgery, which might only improve anatomy. Similarly, in patients in whom surgery does not achieve complete airway patency, unstable respiratory control (high loop gain) may continue to drive OSA. This logic can be used for the other traits as well; for example, relatively less responsive upper airway muscles (or negative effort dependence) may add to the risk of residual sleep apnea after surgery designed to improve anatomy.

One could argue that deleterious physiological traits (eg, high loop gain/arousal threshold) or high AHI_{REM} are markers of severity of OSA. As such, these features may be associated with surgical failure mainly because of the known suboptimal results of surgery in patients with severe OSA. Nonetheless, sparse data are available to understand why patients with severe OSA are not as responsive to upper airway surgery than to less severe disease. Some have argued that those with a high AHI are less likely to be cured purely for statistical reasons. Conversely, our data might suggest that the mechanisms underlying severe OSA may be a factor underlying poor surgical outcomes. However, in this study, as in clinical practice, a relatively high preoperative AHI did not uniformly result in a high postoperative AHI (Figure 1A).

Modeling Response to Surgery

Although loop gain, as well as a number of trait surrogates (eg, AHI_{REM} or arousal threshold), might be predictors of surgical
The importance of nonanatomical traits is dependent on anatomy.

For the same given response to surgery (improvement in anatomy), the role of nonanatomical traits will be very different. For subject A, surgery is likely to improve obstructive sleep apnea (OSA) no matter the nonanatomical traits (ie, the patient has little airway collapsibility and surgery may cure OSA). For subject B, surgery has improved the anatomy but not enough to completely eliminate obstruction. Whether the subject has OSA after surgery will very much depend on the nonanatomical traits. Subject C has an improvement in anatomy, but the anatomy is still very poor. In this case, again, the nonanatomical traits will still contribute to the residual apnea-hypopnea index but not matter very much in deciding the success of surgery. Another difficulty with predication is the variable effect of surgery on the individual’s anatomy. Pcrit = critical pressure.

Figure 3

Figure 4—Actual postoperative apnea-hypopnea index (AHI) versus model-derived AHI by PSG Multiple Regression model or PSG Plus Physiology model in selective patients.

Reference line: x = y. Four patients 1, 2, 3, and 4 with a similar preoperative AHI (50.3, 58.5, 65.2, and 51.3, respectively) had very different postoperative AHI. The estimated postoperative AHI of the same 4 patients are marked in the PSG Plus Physiology model figure as well (asterisks, with the patient’s total physiological score in parenthesis; same patient’s markers are linked with a dash line). Patients 1, 3, and 4 were all stratified to the same anatomy based on AHIres (all bad anatomy). However, their scores from the nonanatomical surrogates were quite different (eg, Patient 1 had a total of 7 points from nonanatomy traits [high arousal threshold, bad upper airway muscle response and control of breathing]; Patient 3 only had 2 points from nonanatomy traits [intermediate sleep stability, intermediate upper airway response, good control of breathing]; and Patient 4 only had 1 point based on the nonanatomical trait surrogates [intermediate control of breathing, good upper airway response and low arousal threshold]). Although patient 2 had better (intermediate) anatomy, the nonanatomical surrogates of physiological traits score was 7 in total, thus the actual residual AHI was higher than estimated by multiple regression model. AHI = apnea-hypopnea index.
that two terms for control of breathing added value to our model suggests that this is the case. The addition of loop gain estimated by the method of Terrill and colleagues\(^7\) did improve the ability of the model to estimate actual surgical outcome, although a simplified version without estimated loop gain (using only the central or mixed apnea index as the sole control of breathing term) still had greater coefficient of determination value than reliance on statistical methods alone. Similarly, a larger fraction of events that are hypopneas was taken to indicate a lower arousal threshold, but it may also reflect a reduced collapsibility of upper airway,\(^{26}\) which may explain why low arousal threshold was correlated with better surgical outcome.

**Limitations of This Study**

Despite the novelty and rigorous approach in our study, we acknowledge a number of limitations. First, the efficiency of structural change by upper airway surgery is likely to be variable and we do not include an estimate of the extent of surgical resection. The effectiveness of an intervention to manipulate anatomy is also an important contributor to the success of the treatment. We did try to minimize this variability by focusing on one surgeon performing a consistent procedure. Nevertheless, the decrease of upper airway collapsibility by a given surgical procedure may be more related to the baseline anatomical features of the patient\(^{1,27}\) (ie, obstruction sites and the structural components of the collapse). This is one reason why the baseline anatomical features are important predictors in previous studies.\(^{1,2,27,28}\) Obtaining this information before and after surgery either by invasive (endoscopy) or noninvasive (inspiratory flow pattern)\(^{25}\) methods and adding to the model may further improve its efficacy. Second, it is likely that the values of the nonanatomical traits contain inherent and acquired (due to OSA) components. For example, loop gain may increase with OSA and may be reduced as the severity of sleep apnea is improved. Currently, no method can distinguish one’s inherent versus acquired loop gain. Third, other changes may occur after surgery, such as weight loss that will affect the model’s strength. Indeed, some subjects did have weight change because all of the patients are encouraged to lose some weight as part of their clinical care. Furthermore, all the surrogates of the traits were derived from clinical PSG, which will not be as precise as direct measurements of the traits. In fact, some of the surrogates may carry information from multiple underlying traits. Ideally, these would be measured directly, although this may not be feasible given the few laboratories currently performing such measurements. Our goal was to use clinically accessible data to allow clinicians to use these concepts to understand/estimate potential surgical outcomes. Moreover, we argue that even precisely measured physiological traits have complex interactions (eg, high loop gain can be masked by low arousal threshold), emphasizing the complex nature of the topic rather than inadequacy of our clinical assessments.

We had a modest sample size of a relatively homogenous group of patients that preclude a model with stratified interactions among physiological traits based on anatomy.\(^{10}\) Moreover, our cohort was selected to undergo upper airway surgery, and thus one could argue that the participants in our study may not be representative of a general OSA population. We acknowledge this issue, but argue that because our goal was to define predictors of surgical success/failure our study population was appropriate. We also recognize that the extent or type of surgery may vary somewhat depending on the anatomical characteristics of the patient, but again this variance reflects the reality of upper airway surgery. Based on our study design, we could not reasonably remove surgical judgment or variation from clinical practice, but again minimized this issue by focusing on one surgeon performing a consistent procedure on clinically homogeneous patients based on the current clinical guidelines.

Most importantly, these models have not been validated in a second sample to assess whether they have predictive value. It is possible that the cutoffs and variables are overfit to this training set, and our findings represent a type 1 error. Although the model requires further confirmation by applying the model to patients undergoing future similar surgeries, we believe that the concepts we propose (incorporation of nonanatomical traits based on underlying physiological concepts) are important and could be developed in other populations of patients and for other interventions (oral appliances, hypoglossal nerve stimulation therapy). When applying these ideas to another population or other interventions, different surrogates, cutoffs, or weights of the parameters may be needed.

**CONCLUSIONS**

Nonanatomical traits such as high loop gain and arousal threshold are associated with residual AHI after upper airway surgery. However, an understanding of OSA physiology, rather than reliance on unguided statistical methods, might be necessary to incorporate the anatomical and nonanatomical surrogates of physiological traits in estimating potential surgical outcomes.

**ABBREVIATIONS**

AHI, apnea-hypopnea index
H-UPPP, uvulopalatopharyngoplasty with uvula preservation
OSA, obstructive sleep apnea
PSG, polysomnography
ROC, receiver operating characteristic
TA, transpalatal advancement
UPPP, uvulopalatopharyngoplasty

**REFERENCES**

1037

Y Li, J Ye, D Han, et al. Physiology Based Modeling and Surgical Treatment for OSA

ACKNOWLEDGMENTS

The authors thank Dr. Bradley A. Edwards for providing us the formula to calculate estimated arousal threshold.

SUBMISSION & CORRESPONDENCE INFORMATION

Submitted for publication March 9, 2017
Submitted in final revised form May 30, 2017
Accepted for publication June 21, 2017

Address correspondence to: Jingyey Ye, MD; No. 168, Litang Road, Beijing, 102218, China; Tel: 086 13701396970; Email: yejingyey@yahoo.net or Demin Han, MD, PhD, No. 1, Dongjiaoinxiang, Beijing, 100730, China; Tel: 086 58269331; Fax: 086 58269858; Email: enthandm@126.com

DISCLOSURE STATEMENT

Work for this study was performed in Beijing Tongren Hospital, Capital Medical University, Key Laboratory of Otolaryngology Head and Neck Surgery (Ministry of Education of China) Beijing, China and data analysis was done in University of California at San Diego, CA, United States. The study was supported by the National Natural Science Foundation of China (8120735) and the National Science and Technology Support Plan (Ministry of Science and Technology of the People’s Republic of China No.2013BAI03B05). Dr. Sands was supported by the American Heart Association (15SDG25890059), the National Health and Medical Research Council of Australia (NHMRC, Early Career Fellowship 1053201), the Menzies Foundation, the American Thoracic Society Foundation and was Co-Investigator on grants from the NHMRC (1064163) and NIH (R01 HL128658). Dr. Malhotra is PI on NIH R01 HL085188, K24 HL121794, and co-investigator on R21 HL121794, R01 HL 119201, R01 HL081823. Yuanru Li; Jingyey Ye; Demin Han; Xin Cao; Xu Ding; Yuhuan Zhang; Scott Sands; Jeremy Orr; Rachel Jen; Wen Xu; Atul Malhotra have no conflicts of interest to disclose. As an Officer of the American Thoracic Society, Dr. Malhotra has relinquished all outside personal income since 2012. Dr. Owens has received travel and honoraria from ResMed and Itamar Medical. ResMed, Inc. has received travel and honoraria from ResMed and Itamar Medical. ResMed, Inc. has provided a philanthropic donation to the UC San Diego in support of a sleep center.