ORIGINAL ARTICLE

Loop Gain Predicts the Response to Upper Airway Surgery in Patients With Obstructive Sleep Apnea

Simon A. Joosten, MBBS, BMedSc, PhD[1,2]; Paul Leong, MBBS, MPHTM[1]; Shane A. Landry, PhD[3]; Scott A. Sands, PhD[4,5]; Philip I. Terrill, PhD[6]; Dwayne Mann, PhD[6]; Anthony Turton, BSc[1]; Jhanavi Rangaswamy, MBBS[1]; Christopher Andara, BSc(Hons)[3]; Glen Burgess, MBBS[7,8]; Darren Mansfield, MBBS, PhD[1]; Garun S. Hamilton, MBBS, PhD[1,2]; Bradley A. Edwards, PhD[3,9]

[1]Monash Lung and Sleep, Monash Medical Centre, Clayton, Victoria, Australia; [2]School of Clinical Sciences, Monash University, Melbourne, Victoria, Australia; [3]Sleep and Circadian Medicine Laboratory, Department of Physiology Monash University, Melbourne, Victoria, Australia; [4]Division of Sleep and Circadian Disorders, Departments of Medicine and Neurology, Brigham & Women's Hospital & Harvard Medical School, Boston, MA; [5]Department of Allergy, Immunology and Respiratory Medicine and Central Clinical School, The Alfred and Monash University, Melbourne, Victoria, Australia; [6]School of Information Technology and Electrical Engineering, The University of Queensland, Brisbane, Australia; [7]The Ear, Nose and Throat/Head and Neck Surgery Unit, Monash Health, Melbourne, Victoria, Australia; [8]Department of Surgery, School of Clinical Science at Monash Health, Monash University; [9]School of Psychological Sciences and Monash Institute of Cognitive and Clinical Neurosciences, Monash University, Melbourne, Victoria, Australia

Study Objectives: Upper airway surgery is often recommended to treat patients with obstructive sleep apnea (OSA) who cannot tolerate continuous positive airways pressure. However, the response to surgery is variable, potentially because it does not improve the nonanatomical factors (ie, loop gain [LG] and arousal threshold) causing OSA. Measuring these traits clinically might predict responses to surgery. Our primary objective was to test the value of LG and arousal threshold to predict surgical success defined as 50% reduction in apnea-hypopnea index (AHI) and AHI <10 events/hour post surgery.

Methods: We retrospectively analyzed data from patients who underwent upper airway surgery for OSA (n = 46). Clinical estimates of LG and arousal threshold were calculated from routine polysomnographic recordings presurgery and postsurgery (median of 124 [91–170] days follow-up).

Results: Surgery reduced both the AHI (39.1 ± 4.2 vs. 26.5 ± 3.6 events/hour; p < .005) and estimated arousal threshold (−14.8 [−22.9 to −10.2] vs. −9.4 [−14.5 to −6.0] cmH_2O) but did not alter LG (0.45 ± 0.08 vs. 0.45 ± 0.12; p = .278). Responders to surgery had a lower baseline LG (0.38 ± 0.02 vs. 0.48 ± 0.01, p < .05) and were younger (31.0 [27.3–42.5] vs. 43.0 [33.0–55.3] years, p < .05) than nonresponders. Lower LG remained a significant predictor of surgical success after controlling for covariates (logistic regression p = .018; receiver operating characteristic area under curve = 0.80).

Conclusions: Our study provides proof-of-principle that upper airway surgery most effectively resolves OSA in patients with lower LG. Predicting the failure of surgical treatment, consequent to less stable ventilatory control (elevated LG), can be achieved in the clinic and may facilitate avoidance of surgical failures.

Keywords: obstructive sleep apnea, surgery, upper airway physiology, ventilatory control.

Statement of Significance

Upper airway surgery is variably successful in treating obstructive sleep apnea (OSA) and predicting response is an ongoing challenge. Surgical failure may occur, at least in part, because predictive tools used to select patients do not take account of the multifactorial nature of OSA pathogenesis in that a number of important nonanatomical factors also contribute to airway collapse. These include ventilatory control instability (ie, elevated loop gain [LG]) and arousal threshold. This study demonstrates that patients with a low LG at baseline were more likely to respond to upper airway surgery. Given that the physiological measurements in the current study are derived from routine polysomnographic signals, this information brings us closer to applying individualized treatments for OSA based on a patient's underlying pathophysiology.

INTRODUCTION

Upper airway surgery presents an alternative for patients with obstructive sleep apnea (OSA) who cannot tolerate continuous positive airways pressure (CPAP) or other first-line treatments. However, a significant number of patients who undergo surgery experience residual OSA[1,2] and predicting which patient is likely to respond to surgery is difficult. Several investigators have previously attempted to use anthropomorphic,[3] polysomnographic (PSG),[4,5] and anatomical[6–9] factors to predict success, but each of these do not have sufficient accuracy to be applied in clinical practice.[10] As such, there is significant motivation to better understand the mechanistic factors that determine surgery success and, as a result, to develop methods which reliably predict the patients who will respond favorably to upper airway surgery.

A key explanation for upper airway surgery failure is the recognition that OSA is a multifactorial disorder, not simply determined by unfavorable upper airway anatomy. While some instances of failure to respond to surgery will be due to poor anatomy alone, recent evidence has demonstrated that several nonanatomical pathophysiological factors including ventilatory control instability (elevated loop gain [LG]) and low arousal threshold also contribute to OSA.[11,12] In fact, Eckert et al. have demonstrated that up to 19% of all patients with OSA have the same degree of airway collapsibility (as a measure of the anatomical compromise) as control participants without OSA. In these patients, it is the presence of an elevated LG and/or low arousal threshold that is the primary effect modifier predisposing them to OSA.[12] Previous work has demonstrated that other CPAP alternative treatments such as mandibular advancement devices (MADs) and lateral positioning during sleep improve upper airway collapsibility without improving ventilatory control instability or arousal threshold,[13,14] such that these residual causes of OSA can persist despite intervention. While previous evidence suggests that upper airway surgery can improve the anatomy/collapsibility,[15–17] it also remains unclear whether upper airway surgery alters the nonanatomical factors that contribute to OSA.

In this study, we propose that if upper airway surgery only improves the upper airway collapsibility, then any baseline abnormalities in nonanatomical contributors to OSA will impact the effectiveness of the treatment. Indeed, under such

circumstances, a high baseline LG or low arousal threshold is likely to predict treatment failure.[13] Therefore, using our recently developed techniques for assessing two of the key non-anatomical traits from standard overnight PSG,[18,19] we tested the primary hypothesis that LG and arousal threshold can be used to predict those patients who will gain the greatest benefit from surgery. We also sought evidence to confirm that upper airway surgery does not alter these traits.

METHODS

Participants

We undertook a retrospective analysis of consecutive patients who underwent anatomically directed upper airway surgery for treatment of OSA (with PSG performed before and after surgery) over a 28-month period at our university teaching hospital. Patients were identified by cross referencing the sleep study database with the surgery database at Monash Health. All patients who had presurgery and post-surgery diagnostic sleep studies and had surgery performed at our institution were included in the analysis. No patients were excluded if they met these criteria. Upper airway surgery was performed through the Ear, Nose and Throat surgical department at Monash Health by board-certified staff surgeons. The majority of patients who had multilevel surgery had so in one surgical visit to hospital. However, in the instances where patients had multilevel surgery performed over more than one visit to hospital, the PSG performed after the completion of all surgical procedures was used as the follow-up PSG. Importantly, the study was not performed to determine the effectiveness of one form of surgery over another but rather to determine factors that predicted surgery success in general. Ethics approval for this study was obtained from Monash Health Human Research Ethics Committee. The medical record was accessed to obtain demographic and anthropomorphic parameters and intraoperative reports. PSG data were accessed from our clinical database of sleep study information. Epworth Sleepiness Scores (ESS) were obtained on all patients at the time of each PSG. The ESS is an eight-point questionnaire of self-reported sleepiness with scores of 10 or more considered to be indicative of excessive sleepiness. The American Society of Anesthesiologists (ASA) physical status classification system is a categorical description of a patient's overall physical status prior to surgery and was obtained at anesthetic review on the day of surgery The scale consists of six categories (although category 5 is reserved for moribund patients) with category 1 consisting of healthy individuals.[20] Finally, a Mallampati score was developed as an indicator of intubation difficulty obtained on the day of surgery and is a four-point categorical classification system based on upper airway configuration, with score of 3 or 4 indicating significant airway crowding.[21]

Data Analysis

PSGs were performed at Monash Health, an academic sleep centre in Melbourne, Australia. Sleep studies were staged and scored according to American Academy of Sleep Medicine criteria[22] using the alternative criteria for scoring respiratory events (specifically rule 1A: hypopnea defined as 30% reduction in respiratory flow amplitude from baseline for 10 or more seconds and one of either 3% desaturation or arousal from sleep). No sedation was administered during the PSGs. LG[18] and arousal threshold[19] analysis was performed on signals obtained from the scored PSG data using our previously described and validated methods.[18]

Determining LG

Briefly, we used routine PSG signals to calculate dynamic LG. The data were exported from the PSG recording as a European Data Format file and imported into Matlab (R2015a version 8.5.0197613 Mathworks Inc., Natick, Massachusetts) for manipulation. All available 7-minute periods of nonrapid eye movement sleep that contain one or more scored obstructive apneas/hypopneas were identified using a software routine. Nasal pressure was square-root transformed and taken as a surrogate of ventilatory flow and integrated and normalized by the mean to provide a ventilation signal for subsequent analysis (see Figure 1). A categorical breath-by-breath time series of scored electroencephalogram arousals and scored obstructed breaths was created. Using these data, a standard ventilatory control model was fit to determine the best set of system parameters (ie, a gain, time constant, and delay) for each 7-minute epoch. Using the best set of parameters, the model outputs an estimated ventilatory drive signal that best fits the ventilation during unobstructed breaths (ie, when ventilation reflects ventilatory drive). These parameters were then used to calculate the magnitude of LG. For consistency with the dynamics of OSA,[23] our primary measure was LG at the natural cycling frequency (LGn, ie, the frequency of periodic breathing if breathing was unstable). We also considered LG at 1 cycle/min (LG_1, see online Supplementary Material). All LG measurements for each 7-minute epoch were then averaged for the entire night.

Determining the Respiratory Arousal Threshold

Standard anthropometric and clinical PSGs values were input into our published model[19] in order to calculate the arousal threshold according the following equation: Arousal threshold = −65.39 + (0.06 × age) + (3.69 × gender [where male = 1, female = 0]) − (0.03 × body mass index) − (0.11 × AHI) + (0.53 × Nadir SpO_2) + (0.09 × % of the overall respiratory events that are hypopneas).

Statistical and Responder Analyses

Statistical analysis was performed using SPSS (version 20, 2011, New York, USA). Continuous variables were expressed as means and standard deviations where normally distributed and as median and interquartile range where not normally distributed. Comparisons were made using paired *t*-tests for normally distributed data, Wilcoxon signed rank or Mann–Whitney U tests for nonparametric data and Fisher's exact test for categorical data. A *p*-value of <.05 was considered significant.

In order to determine the characteristics of those who gained the greatest benefit from upper airway surgery, patients were categorized as "responders" if they displayed an AHI reduction ≥50% and a treatment AHI <10 events/hour.[13] We chose this criterion for our primary analysis as it has the widest applicability for clinical practice. Sensitivity analyses were additionally performed using two alternative definitions of treatment success: (1) AHI reduction ≥50% and (2) AHI ≤5 events/hour (see online

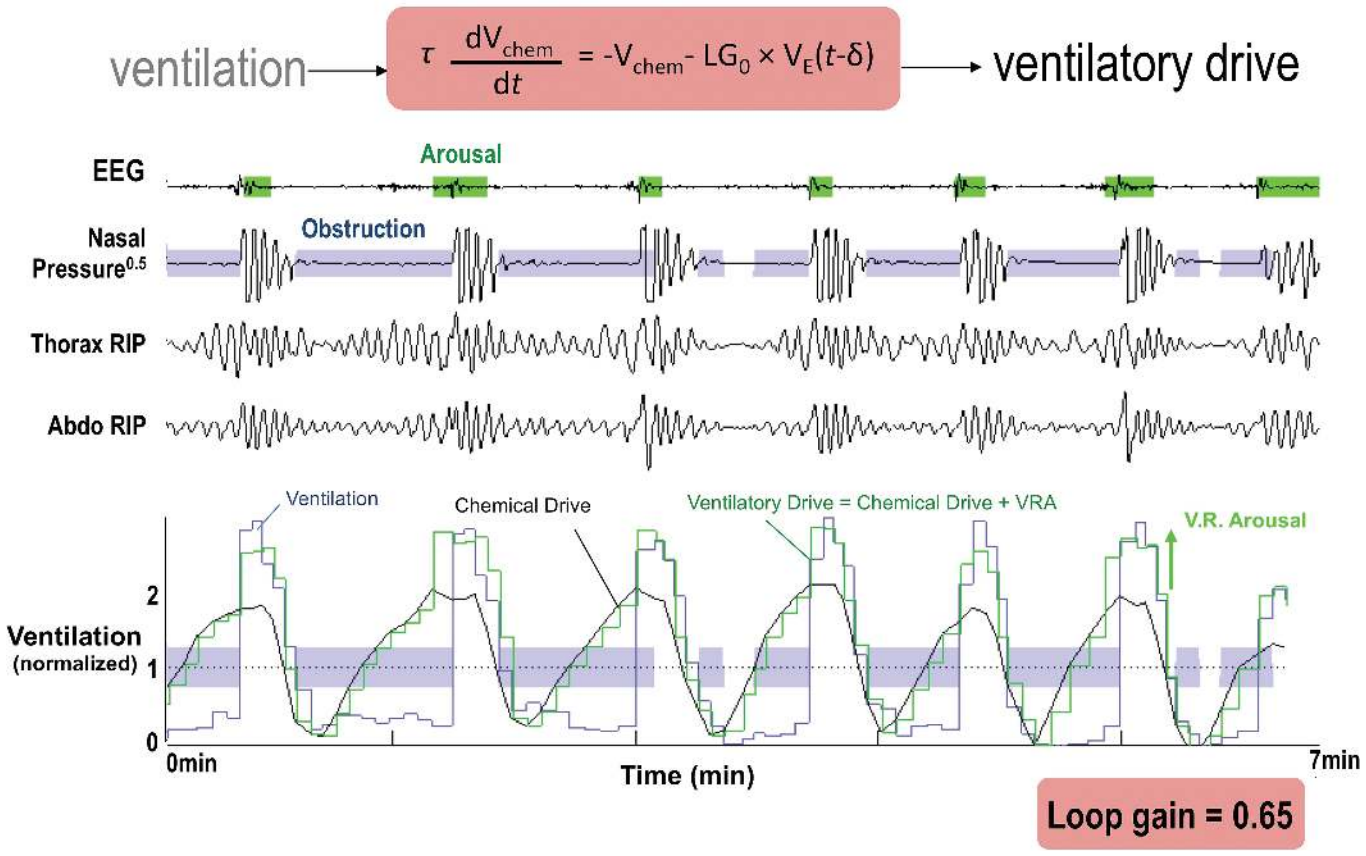

Figure 1—Method for calculating loop gain from standard polysomnographic signals. To determine an individual's loop gain, we used our recently developed technique that fits a mathematical model to the signals recorded in a patient's own PSG data. The key is that the patient's own obstructive events provide perturbations in ventilation, and this allows the model to be fitted. The blue line in the bottom panel is measured ventilation, which is used to model chemical drive (ie, the intended ventilatory response to elevated carbon dioxide and decreased oxygen levels) according to the equation at the top of the diagram and represented graphically by the black line in the bottom panel. Where there is an associated respiratory arousal, total ventilatory drive is equal to chemical drive plus the added drive associated with arousal/waking (ie, ventilatory response to arousal or VRA) and this is represented by the green line in the bottom panel. To characterize the system, the parameters in the equation (τ, δ, LG_0, and VRA) are adjusted until ventilatory drive (ie, green line) best fits the observed ventilation in unobstructed breaths. The parameters can then be used to calculate the magnitude of loop gain at any frequency. EEG, electroencephalogram; τ, time constant (due to the time course of carbon dioxide buffering in tissues and the lungs); δ, delay time (circulatory delay between the chemoreceptors and the lungs); LG_0, the steady-state loop gain; RIP, respiratory inductance plethysmography; *t*, time; V_{chem}, chemical drive due to elevation of carbon dioxide and fall in oxygen; V_E, ventilation; VRA, ventilatory response to arousal.

Supplementary Material). Logistic regression was performed to identify predictors of surgical success. Receiver operating characteristic (ROC) curves were calculated to determine the best threshold values (based on sensitivity and specificity) that differentiated patients' groups based on the significant predictors identified in the logistic regression analyses.

RESULTS

Forty-six patients with OSA who underwent PSGs before and after surgery were identified from our database and included in the analysis. The median delay from initial PSG to surgery was 295 days (interquartile range [IQR] 182–472) and from surgery to the follow-up PSG was 124 days (IQR 91–170). Multilevel surgery was performed in 39/46 patients, tonsil-only surgery in four, and nasal-only surgery in three. For a full description of the surgical procedures performed in each of the 46 patients, see Table E1 in Supplementary material. Table 1 presents patients' baseline characteristics. Patients were generally obese middle-aged males with severe OSA. Comorbidities were generally mild in nature, reflected by low ASA physical status classification.

Effects of Upper Airway Surgery on Sleep and the Nonanatomical Traits

The effects of upper airway surgery on both the characteristics of sleep and the underlying OSA physiology (determined from clinical PSG) are summarized in Table 2. As a group, surgery was associated with a significant reduction in the overall AHI ($p < .005$; Figure 2), driven predominantly by reductions in the obstructive apnea index. All measures of the hypoxemic impact of OSA were also improved following surgery. Self-reported sleepiness (as measured by the ESS) was lowered significantly

Table 1—Baseline Participant Characteristics (n = 46).

Variable	Baseline
Age, years	42.3 ± 2.1
Tried CPAP prior to surgery	10 (21.7%)
Mallampati, *n*, %	1 10 (21.7%) 2 21 (45.7%) 3 10 (21.7%) 4 4 (8.7%)
Gender, male, %	76.1
BMI, kg/m^2, mean (SD)	32.4 (6.9)
ASA, category,n (%)	1 6 (13.0%) 2 29 (63.0%) 3 6 (13.0%) 4 1 (2.2%)
ESS	8 [4 to 13]
Total AHI, events/hour	39.1 ± 4.2
AHI REM, events/hour	44.9 ± 4.7
AHI NREM, events/hour	37.9 ± 4.4

Values are means ± SEM. ASA category n = 42 and Mallampatti n = 45 due to missing data.
AHI, apnea-hypopnea index; ASA, American Society of Anesthesiologists physical status classification; BMI, body mass index; CPAP, continuous positive airways pressure; ESS, Epworth Sleepiness Score; NREM, nonrapid eye movement; REM, rapid eye movement; SD, standard deviation; SEM, standard error of the mean.

by two points (p = .012). In terms of how upper airway surgery altered the physiological characteristics, the key findings were that the arousal threshold was lower after surgery (p < .0005), whereas LG (p = .28) was not altered. Interestingly, the change in the overall AHI was weakly correlated with changes in both LG (r = 0.312, p = .035) and the arousal threshold (r = −0.382, p = .009). When examining whether the baseline physiological measures predicted the change in AHI, we observed that the change in AHI was negatively correlated with baseline arousal threshold (r = −0.495, p = .0005) but not with the baseline LG.

Predictors of Response to Therapy

Twelve (26%) out of the 46 patients were classified as "responders" to therapy (by the AHI <50% and <10/hour criterion). The baseline clinical and physiological characteristics (ie, traits) of "responders" are summarized in Table 3 and Figure 3. In terms of their baseline (ie, presurgery) clinical and PSG characteristics, responders were younger compared to nonresponders, but BMI (body mass index) and AHI were not different between groups. When examining the underlying physiological traits, responders had a lower baseline LG (assessed either using LGn or LG_1), whereas there was no difference in arousal threshold between groups at baseline. Importantly, the type of surgery received did not differ between groups (p = .4) and the Mallampatti score and ESS did not differ significantly between groups at baseline (p = .5, Mann–Whitney U test, see Table 3). The ESS was significantly lower following surgery in the nonresponder group (7.5 [4–13] presurgery, 6 [2.5–10.75] post surgery, p = .036, Wilcoxon signed rank test, see Table 3) while the ESS was not significantly lower in the responder group. Similar to the main findings, surgery significantly reduced the AHI in both responders and nonresponders (although to a lesser extent in nonresponders by definition) and it significantly lowered the arousal threshold despite not altering any of the other physiological variables. LG_n was also significantly lower at baseline in the responder group using both of the alternative definitions of treatment success (see Table E5 and E6 in Supplementary material).

Logistic regression analysis was performed, utilizing a definition of success of AHI reduced by >50% to an AHI <10 events/hour, to determine predictors of surgical success. The factors included in the model were baseline age, BMI, AHI, LG, and gender. The model was statistically significant and predicted 44.2% of the variance in surgical success (X^2 (5) = 16.52, p = .002, Nagelkerke R^2 = 0.442). LG remained a strong predictor of a successful response to surgery even after controlling for age, BMI, and baseline AHI (p = .018; see Table E2 in Supplementary material). Using ROC analysis, the area under curve was 0.80 (95% confidence interval 0.66–0.94) for LG_n (see Figure E1 in Supplementary material). Table 4 presents sensitivity, specificity, and predictive values for various clinically relevant cut points with respect to ruling in or ruling out surgical success. A baseline LGn of ≤0.50 was 100% sensitive for subsequent surgical success (ie, all patients with a presurgical LGn >0.50 [n = 12] were surgical failures), and an LGn ≥0.39 was 94% specific for failure to respond to surgery (ie, the majority of patients with an LG <0.39 were surgical successes). Similar results were observed when LG_1 was used (see Table E4 in Supplementary material).

DISCUSSION

The results of upper airway surgery directed toward anatomical abnormalities are difficult to predict, potentially because current predictive tools do not account for the contribution of nonanatomical pathophysiological factors to OSA. Unlike CPAP, the degree to which a patient's anatomical compromise is reversed with surgery may not be powerful enough to render the nonanatomical factors inconsequential. The major novel finding of the current study is that responders to surgery had a lower baseline LG and were younger than nonresponders; and furthermore, only LG remained a strong and significant predictor of surgical success when these parameters were combined in a logistic regression (also controlling for age, BMI, and baseline AHI). These findings suggest that a less unstable ventilatory control system is a recognizable factor predicting successful upper airway surgery. Perhaps, most importantly, our findings also demonstrated that a high baseline LG (>0.5) was 100% predictive of surgical failure—thus highlighting a group who potentially could be spared from undergoing futile surgery.

Effect of Surgery on Nonanatomical Contributors to OSA

A major finding of our study is that upper airway surgery for OSA does not alter LG. Such a finding is similar to what has previously been observed in OSA patients treated with a MAD[13] or lying in the lateral position.[14] The observed lack of change in

Table 2—Physiological Characteristics in All Patients (n = 46) Before and After Surgery.

Parameter	Before surgery	After surgery	p-Value*
Sleep characteristics			
Time in bed (TIB), minutes	474.5 ± 13.4	437.7 ± 9.3	**.018**
Total sleep time (TST), minutes	373.0 ± 11.3	346.9 ± 14.0	.160
% Sleep efficiency (TST/TIB)	79.5 ± 1.7	80.9 ± 2.7	.524
N1 duration, minutes	35.5 [17.8 to 61.3]	35.3 [24.4 to 53.0]	.973
N2 duration, minutes	212.8 [161.3 to 254.3]	169.8 [139.4 to 233.6]	**.037**
N3 duration, minutes	61.0 [28.6 to 85.6]	69.8 [36.8 to 91.1]	.108
NREM duration, minutes	311.8 [275.4 to 360.1]	316.5 [261.9 to 354.0]	.238
REM duration, minutes	59.5 [39.8 to 74.4]	55.3 [42.3 to 71.3]	.797
Mean SpO_2, %	93.3 ± 0.6	94.5 ± 0.3	**.006**
Lowest SpO_2, %	80.1 ± 1.6	85.9 ± 0.9	**.0005**
ODI 3%, events/hour	34.0 ± 5.0	21.5 ± 3.8	**.007**
ODI 4%, events/hour	26.0 ± 4.7	14.7 ± 3.2	**.007**
Obstructive apnea index, events/hour	18.2 ± 3.4	10.7 ± 2.6	**.004**
Hypopnea index, events/hour	22.7 ± 2.8	18.5 ± 2.7	.180
Respiratory arousal index, events/hour	17.6 ± 3.3	16.5 ± 2.6	.391
Total AHI, events/hour	39.1 ± 4.2	26.5 ± 3.6	**.003**
AHI REM, events/hour	44.9 ± 4.7	28.4 ± 4.2	**.001**
AHI NREM, events/hour	37.9 ± 4.4	26.2 ± 3.7	**.010**
Epworth sleepiness score	8 [4 to 13]	6 [2 to 10.75]	**.012**
Physiological characteristics			
LGn	0.45 ± 0.01	0.45 ± 0.02	.278
Tau, seconds	175.1 [101.1 to 180.0]	175.2 [132.2 to 179.8]	.362
Chemoreflex delay, seconds	11.8 [10.3 to 14.2]	12.5 [10.7 to 15.1]	.299
Arousal threshold, cmH_2O	−14.8 [−22.9 to −10.2]	−9.4 [−14.5 to −6.0]	**<.0001**
Tn, seconds	43.4 [37.7 to 53.2]	45.3 [39.9 to 52.4]	.232

*Paired samples t-test for parametric data, Wilcoxon sign-rank for nonparametric data. Values are means ± SEM. or medians [interquartile range].
Arousal threshold, expressed as a proportion of mean ventilation; LGn, loop gain at natural frequency; NREM, nonrapid eye movement; ODI 3%, oxygen desaturation index 3%; REM, rapid eye movement; SEM, standard error of the mean; Tau, time constant; Tn, natural cycling period defined as the period of sinusoidal disturbance that results in an "in phase" feedback response.

LG in the current study remained even within responder/nonresponder subgroups. The fact that LG does not fall following surgery in the responder group, despite significant reductions in AHI, suggests that we obtained an estimate of the individual's "intrinsic" LG.[24] That is, LG was not elevated in these patients as a consequence of OSA but rather that it was an "intrinsic" factor in the development of the disease in the first place. If our measured LG was an "acquired" feature subsequent to the development of OSA, we would have expected it to fall as the AHI was lowered (particularly in responders)—a phenomenon that has been observed in other studies assessing the effect of LG following 1 month CPAP therapy.[25,26] While it remains a challenge to resolve the disparate findings between studies, we hypothesize that such differences may be due to the methods used to measure LG (and its components, ie, controller and plant gain).

Given that LG was not altered by surgery, the finding that the respiratory arousal threshold decreased post surgery may at first seem unexpected. This finding implies that patients are prone to arousing from sleep at a lower intensity respiratory stimulus following surgery. In contrast to our LG findings, this may suggest that an elevated arousal threshold is an acquired feature of OSA. This is further supported by our finding that the arousal threshold remained higher in nonresponders post surgery and previous studies showing that (1) the severity of OSA has been strongly linked to the arousal threshold[19,27–29] and (2) the arousal threshold increases following discontinuation of CPAP[30] and decreases immediately following institution of CPAP therapy,[31] a trend which continues up until at least 3 months.[25,31]

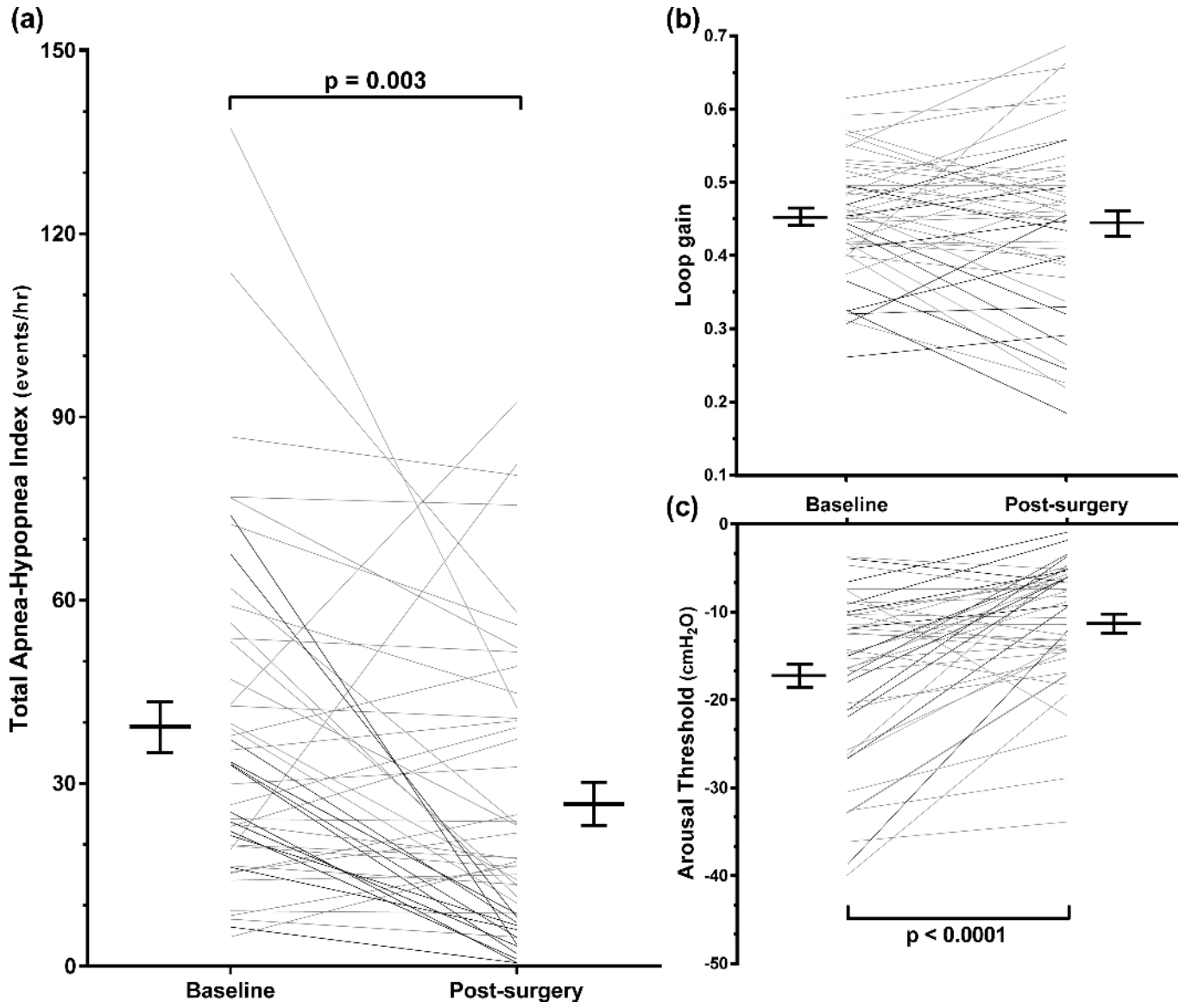

Figure 2—The effects of upper airway surgery on sleep and physiological variables. Upper airway surgery significantly (A) reduced the overall apnea–hypopnea index (AHI). When examining the effects on the underlying physiology, surgery did not alter our estimate of loop gain (B) but was associated with a lower estimated arousal threshold (C). Black lines represent those patients considered responders to therapy whereas gray lines represent nonresponders (see text for definitions of responders/nonresponders). Loop gain was estimated using the ventilatory response to apneas/hypopneas (nasal pressure) and arousal threshold was estimated using a linear regression equation (see text for details). Loop gain here is LGn.

Why Might Surgery Fail in a Given Patient?

Patients with OSA can be divided into three distinct groups based on the presence of anatomical and/or nonanatomical physiological factors that predispose to airway collapse, as demonstrated by Eckert et al.[12] The first group of patients have a very severe anatomical deficit with a highly collapsible airway and a smaller contribution from nonanatomical factors. The second group of patients has an intermediate anatomical deficit—with 64% having contribution from nonanatomical factors. The third group of patients has mild vulnerability to airway collapse but major contributions from nonanatomical factors. With this paradigm in mind, we feel there are three potential reasons for surgical failure in OSA:

a. The collapsibility/anatomical deficit is so severe that current surgical techniques (other than tracheotomy, which bypasses the site of obstruction) simply cannot overcome the abnormality.
b. The collapsibility/anatomical deficit is potentially reversible with surgery, however, the technique or procedure applied does not adequately (or specifically) target the site of upper airway collapse.
c. In a group of patients, despite potential surgical improvements in upper airway anatomy/collapsibility trait, there are remaining abnormalities in other nonanatomical traits that are not modifiable by surgery (eg, high LG, low arousal threshold).

Table 3—Clinical and Physiological Characteristics in Surgical Responders and Nonresponders Before and After Upper Airway Surgery.

Parameter	Surgical responders (n = 12)		Surgical nonresponders (n = 34)	
	Baseline	Postsurgery	Baseline	Postsurgery
Clinical				
Age, years	31.0 [27.3 to 42.5]	33.5 [29.0 to 47.8]*	43.0 [33.0 to 55.3]$^{+}$	44.5 [34.8 to 55.3]*^
Gender, Male:Female	7:5		28:6	
BMI, kg/m^2	29.9 [26.3 to 37.7]	29.9 [25.3 to 35.1]	30.4 [28.4 to 38.0]	30.5 [27.5 to 39.8]
AHI, events/hour	29.0 [21.3 to 36.0]	4.0 [1.3 to 7.0]*	36.5 [18.3 to 56.5]	24.5 [15.8 to 49.8]*^
ESS	8 [1 to 12]	6 [2 to 14]	7.5 [4 to 13]	6 [2.5 to 10.75]*
Mallampatti score	2 [1 to 2.75]		2 [2 to 3]	
Physiological				
LGn	0.38 ± 0.02	0.37 ± 0.03	0.48 ± 0.01^{+}	0.47 ± 0.02$^{\wedge}$
LG_1	0.49 ± 0.04	0.43 ± 0.02	0.66 ± 0.03^{+}	0.64 ± 0.03$^{\wedge}$
Arousal threshold, cmH_2O	−16.1 [−21.7 to −9.3]	−5.6 [−8.6 to −3.5]*	−14.5 [−25.0 to −10.3]	−13.3 [−17.0 to −7.5]*^

Values are means ± SEM. or medians [interquartile range]. Significant comparisons at $p < .05$ for *baseline versus post surgery within responders and nonresponders; $^{+}$at baseline in responder versus nonresponder groups; ^post surgery in responder versus nonresponders. Independent samples t-test for parametric data, Mann–Whitney U for nonparametric unpaired data, Wilcoxon signed rank for nonparametric paired data. Data pertaining to Tau (time constant), Tn (natural cycling period), and chemoreflex delay are included in Supplementary material, Table E7.
AHI, apnea–hypopnea index; Arousal threshold, proportion of mean ventilation; BMI, body mass index; ESS, Epworth Sleepiness Score; LG_1, loop gain, ventilatory response to 1 cycle/minute disturbance; LGn, loop gain at natural frequency; SEM, standard error of the mean.

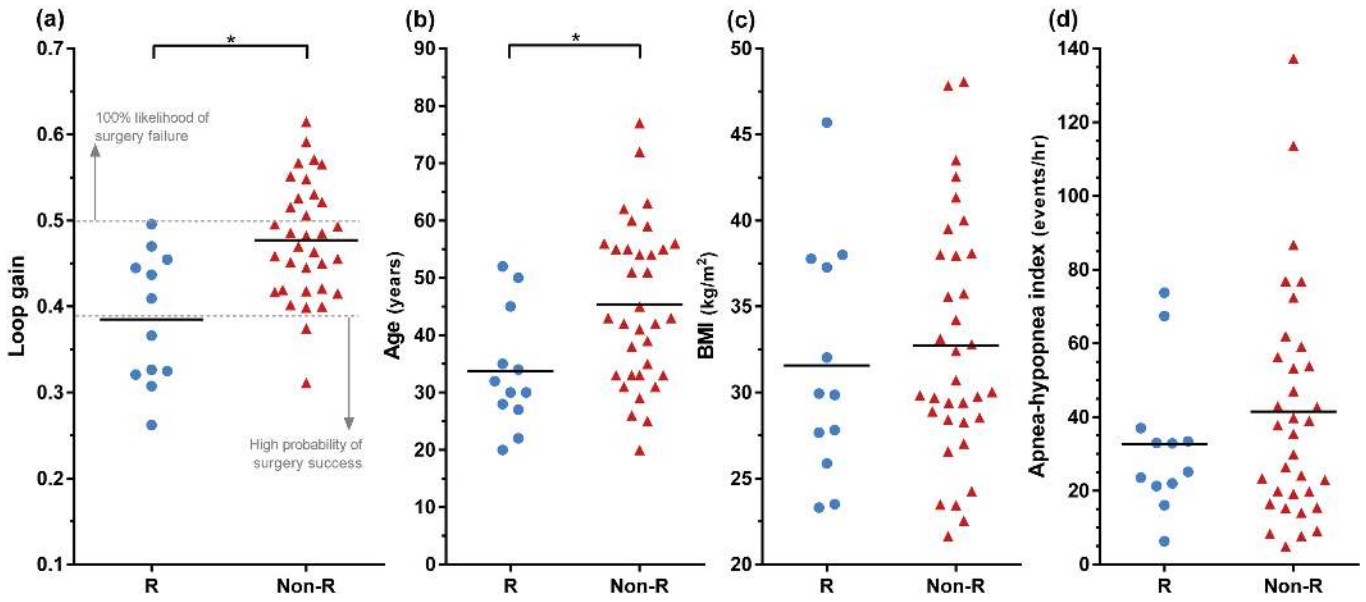

Figure 3—Baseline physiological and clinical characteristics of responders to surgery. Responders to surgery had (A) a lower loop gain and (B) were younger. There was no difference in the presurgical (C) BMI or (D) apnea–hypopnea index. R, responders (R); Non-R, nonresponders; LGn, loop gain; BMI, body mass index.

Selecting Patients for Upper Airway Surgery

It is evident from the current literature that upper airway surgery only resolves OSA in a proportion of patients and that the treatment response is difficult to predict. While previous studies suggest responders to surgery typically have a lower BMI and AHI as well as certain anatomical features,[3–9] current predictive tools are not accurate enough for clinical use. For instance, surgical failure can result despite an apparently favorable preoperative anatomical milieu.[10] The crucial missing information is an understanding both of the dominant pathophysiological mechanisms responsible for airway obstruction in a given patient (particularly nonanatomical

Table 4—Loop Gain: Sensitivity and Specificity for Surgical Success.

Loop gain type	Cutoff	Sensitivity	Specificity	PPV	NPV	TP (#)	TN (#)	FP (#)	FN (#)
LGn ≤	0.50	1.00	0.35	0.35	1.00	12	12	22	0
	0.48	0.92	0.57	0.38	0.94	11	16	18	1
	0.39	0.50	0.94	0.75	0.84	6	32	2	6

FN, false negative; FP, false positive; LGn, loop gain at natural frequency; NPV, negative predictive value; PPV, positive predictive value; TN, true negative; TP, true positive.

contributors) and the effect of surgery on these processes. Our study demonstrates that although in general nonresponders have a higher age and AHI compared to responders, having a lower age or AHI does not predict surgical success in our cohort. LGn, however, is predictive of either surgical success or failure—depending on where one sets the cutoff point. Using an LG threshold of ≤0.50 had a sensitivity for surgical success of 100%. This means that all surgical successes had an LG ≤ 0.50 and thus an LGn above this level was 100% predictive of surgical failure. Importantly, regardless of the definition of success utilized, LG remains a predictor of surgical success indicating the robust nature of our findings (see Supplementary material, Tables E2–E7). If these findings are confirmed prospectively, then knowing this information preoperatively would be of enormous value to the clinician, as it would allow futile surgery to be avoided.

When examining Figure 3, it is clear that the outcome of surgery is difficult to predict for patients with an intermediate LG (0.39 < LGn < 0.50). In a similar examination of patients' response to treatment with MAD,[13] we found that the addition of an anatomical measurement to the predictive model helped predict surgical response for patients with an LG in this "zone of uncertainty". In fact, the addition of a measurement of a patient's anatomical compromise (ie, collapsibility) in our previous study improved the predictive accuracy to almost 100%. Crucially, the only anatomical measurement included in our current study was the Mallampati Score, which is inferior to several other routinely utilized anatomical models/scores.[32] This opens up the possibility that baseline LG could be coupled with some simple measure of the patients' anatomical deficiency (ie, obtained using a Friedman score, drug-induced sleep endoscopy findings, or newer techniques measuring upper airway collapsibility during wakefulness[33] or sleep,[34] which could further improve predictive power and reduce unnecessary surgery. The anatomical measure applied to improve the predictive power of the model would ideally be predictive, reproducible, cost-effective, and easy to apply.

The current work highlights the importance of understanding a patient's baseline physiology before embarking on CPAP-alternative therapies such as upper airway surgery. It suggests that knowing the contribution of ventilatory control instability to a given patient's OSA pathogenesis will enable a more accurate prediction of the likelihood of surgical success. Critically, for those patients failing surgery because of a high LG, the possibility exists that their treatment could be salvaged by additional combinations of therapy targeting elevated LG.

Limitations

There are several limitations to our work that require consideration. Firstly, the retrospective nature and relatively small number of patients included in the study mean that the impact of our findings will only become clear if the results are replicated in a prospective study design. However, the statistical strength of the associations makes us confident of our findings. Secondly, we have not been able to include a robust and consistent measure of anatomy in our predictive model, which is important for those responders/nonresponders with overlapping LGn measurements (ie, those with LGn between 0.39 and 0.50, 26/46 patients). This clearly hampers the effectiveness of the model, although it makes the finding of the predictive nature of LG for the success and failure of surgery all the more impressive. It is likely that the addition of a robust measure of either upper airway anatomy (eg, Friedman classification or drug-induced sleep endoscopy to determine site of collapse) or collapsibility (ie, critical closing pressure or Pcrit) to the model will improve our ability to better predict those likely to have their OSA successfully resolved following surgery. Thirdly, we have included patients who have undergone a variety of surgical procedures. Clearly, some specific clinical situations and procedures are more likely to result in surgical success (eg, tonsillectomy for grossly enlarged tonsils), whereas other procedures and situations are less likely to result in success (eg, nasal surgery as a standalone procedure).[35] Despite this, we feel that the current analysis represents a real-world examination of surgical practice and the inclusion of these procedures/clinical situations adds strength to the analysis. Fourthly, caution is warranted when interpreting the results of the changes in the arousal threshold with surgery as the clinical prediction algorithm used for estimating the arousal threshold used in the current study utilizes three measures of OSA severity. Given that OSA severity will differ systematically between success versus failure groups by definition, the greater reductions in the estimated arousal threshold in responders compared to nonresponders is to be expected. It is worth noting that previous studies utilizing gold-standard measurements of the respiratory arousal threshold have shown similar changes: the arousal threshold increases following discontinuation of CPAP[30] and decreases immediately following institution of CPAP therapy.[31] Thus, while future studies will need to confirm the current findings with gold-standard measurements of the arousal threshold (ie, methods utilizing esophageal pressure changes), this evidence supports our observation that the reduction in the arousal threshold with upper airway surgery is likely a real phenomena. Finally, our definition of "responder" was solely based on

a reduction in AHI, but clearly this is not the only factor that could be used to judge surgical success. It may be just as important to abolish snoring as well as improve the adverse health consequences associated with OSA (eg, reduced daytime sleepiness). It is worth noting that the surgical nonresponders had a significant reduction in ESS post surgery of two points, which is a clinically significant difference for the ESS,[36] this finding was observed in nonresponders regardless of the definition of success applied (see Supplementary material, Tables E5–E7).

CONCLUSION

Our study is the first to test whether deficits in nonanatomic traits predict the resolution of OSA with upper airway surgery. We show that baseline LG, measured using a clinically applicable tool, is a strong predictor of surgical success and that a high-baseline LG greatly increases the risk of surgical failure. The retrospective nature of our study demands further investigation with prospective study design; however, our findings (and the noninvasive methods used to obtain them from standard clinical PSG recordings) represent a significant clinical advance in the understanding of which patients are best suited to upper airway surgery. The findings also highlight the importance of understanding a given patient's underlying physiology in order to successfully tailor personalized therapy for patients experiencing OSA.

REFERENCES

1. Sundaram S, Bridgman SA, Lim J, Lasserson TJ. Surgery for obstructive sleep apnoea. Cochrane Database Syst Rev. 2005;(4): CD001004.
2. Smith DF, Cohen AP, Ishman SL. Surgical management of OSA in adults. Chest. 2015; 147(6): 1681–1690.
3. Hoff PT, Glazer TA, Spector ME. Body mass index predicts success in patients undergoing transoral robotic surgery for obstructive sleep apnea. ORL J Otorhinolaryngol Relat Spec. 2014; 76(5): 266–272.
4. Xiong YP, Yi HL, Yin SK, et al. Predictors of surgical outcomes of uvulopalatopharyngoplasty for obstructive sleep apnea hypopnea syndrome. Otolaryngol Head Neck Surg. 2011; 145(6): 1049–1054.
5. Millman RP, Carlisle CC, Rosenberg C, Kahn D, McRae R, Kramer NR. Simple predictors of uvulopalatopharyngoplasty outcome in the treatment of obstructive sleep apnea. Chest. 2000; 118(4): 1025–1030.
6. Koutsourelakis I, Safiruddin F, Ravesloot M, Zakynthinos S, de Vries N. Surgery for obstructive sleep apnea: sleep endoscopy determinants of outcome. Laryngoscope. 2012; 122(11): 2587–2591.
7. Zhang P, Ye J, Pan C, Sun N, Kang D. The role of obstruction length and height in predicting outcome of velopharyngeal surgery. Otolaryngol Head Neck Surg. 2015; 153(1): 144–149.
8. Sher AE, Schechtman KB, Piccirillo JF. The efficacy of surgical modifications of the upper airway in adults with obstructive sleep apnea syndrome. Sleep. 1996; 19(2): 156–177.
9. Friedman M, Ibrahim H, Bass L. Clinical staging for sleep-disordered breathing. Otolaryngol Head Neck Surg. 2002; 127(1): 13–21.
10. Hessel NS, Vries N. Increase of the apnoea-hypopnoea index after uvulopalatopharyngoplasty: analysis of failure. Clin Otolaryngol Allied Sci. 2004; 29(6): 682–685.
11. Wellman A, Edwards BA, Sands SA, et al. A simplified method for determining phenotypic traits in patients with obstructive sleep apnea. J Appl Physiol (1985). 2013; 114(7): 911–922.
12. Eckert DJ, White DP, Jordan AS, Malhotra A, Wellman A. Defining phenotypic causes of obstructive sleep apnea. Identification of novel therapeutic targets. Am J Respir Crit Care Med. 2013; 188(8): 996–1004.
13. Edwards BA, Andara C, Landry S, Sands SA, Joosten SA, Owens RL, et al. Upper-airway collapsibility and loop gain predict the response to oral appliance therapy in obstructive sleep apnea patients. Am J Respir Crit Care Med. 2016.
14. Joosten SA, Edwards BA, Wellman A, et al. The effect of body position on physiological factors that contribute to obstructive sleep apnea. Sleep. 2015; 38(9): 1469–1478.
15. Tucker Woodson B. Structural effectiveness of pharyngeal sleep apnea surgery. Sleep Med Rev. 2008; 12(6): 463–479.
16. Tucker Woodson B. Acute effects of palatopharyngoplasty on airway collapsibility. Otolaryngol Head Neck Surg. 1999; 121(1): 82–86.
17. Schwartz AR, Schubert N, Rothman W, et al. Effect of uvulopalatopharyngoplasty on upper airway collapsibility in obstructive sleep apnea. Am Rev Respir Dis. 1992; 145(3): 527–532.
18. Terrill PI, Edwards BA, Nemati S, et al. Quantifying the ventilatory control contribution to sleep apnoea using polysomnography. Eur Respir J. 2015; 45(2): 408–418.
19. Edwards BA, Eckert DJ, McSharry DG, et al. Clinical predictors of the respiratory arousal threshold in patients with obstructive sleep apnea. Am J Respir Crit Care Med. 2014; 190(11): 1293–1300.
20. Owens WD, Felts JA, Spitznagel EL Jr. ASA physical status classifications: a study of consistency of ratings. Anesthesiology. 1978; 49(4): 239–243.
21. Mallampati SR, Gatt SP, Gugino LD, et al. A clinical sign to predict difficult tracheal intubation: a prospective study. Can Anaesth Soc J. 1985; 32(4): 429–434.
22. Iber C, American Academy of Sleep Medicine. The AASM manual for the scoring of sleep and associated events: rules, terminology, and technical specifications. Westchester, IL: American Academy of Sleep Medicine; 2007. 59 p.
23. Edwards BA, Sands SA, Eckert DJ, et al. Acetazolamide improves loop gain but not the other physiological traits causing obstructive sleep apnoea. J Physiol. 2012; 590(5): 1199–1211.
24. Orr JE, Edwards BA, Malhotra A. CrossTalk opposing view: loop gain is not a consequence of obstructive sleep apnoea. J Physiol. 2014; 592(14): 2903–2905.
25. Loewen A, Ostrowski M, Laprairie J, et al. Determinants of ventilatory instability in obstructive sleep apnea: inherent or acquired? Sleep. 2009; 32(10): 1355–1365.
26. Salloum A, Rowley JA, Mateika JH, Chowdhuri S, Omran Q, Badr MS. Increased propensity for central apnea in patients with obstructive sleep apnea: effect of nasal continuous positive airway pressure. Am J Respir Crit Care Med. 2010; 181(2): 189–193.
27. Eckert DJ, Younes MK. Arousal from sleep: implications for obstructive sleep apnea pathogenesis and treatment. J Appl Physiol (1985). 2014; 116(3): 302–313.
28. Suzuki M, Ogawa H, Okabe S, et al. Digital recording and analysis of esophageal pressure for patients with obstructive sleep apnea–hypopnea syndrome. Sleep Breath. 2005; 9(2): 64–72.
29. Krieger J, Sforza E, Boudewijns A, Zamagni M, Petiau C. Respiratory effort during obstructive sleep apnea: role of age and sleep state. Chest. 1997; 112(4): 875–884.
30. Berry RB, Kouchi KG, Der DE, Dickel MJ, Light RW. Sleep apnea impairs the arousal response to airway occlusion. Chest. 1996; 109(6): 1490–1496.
31. Haba-Rubio J, Sforza E, Weiss T, Schröder C, Krieger J. Effect of CPAP treatment on inspiratory arousal threshold during NREM sleep in OSAS. Sleep Breath. 2005; 9(1): 12–19.
32. Aktas O, Erdur O, Cirik AA, Kayhan FT. The role of drug-induced sleep endoscopy in surgical planning for obstructive sleep apnea syndrome. Eur Arch Otorhinolaryngol. 2015; 272(8): 2039–2043.
33. Hirata RP, Schorr F, Kayamori F, et al. Upper airway collapsibility assessed by negative expiratory pressure while awake is associated with upper airway anatomy. J Clin Sleep Med. 2016; 12(10): 1339–1346.
34. Azarbarzin A, Sands SA, Taranto-Montemurro L, et al. Estimation of pharyngeal collapsibility during sleep by peak inspiratory airflow. Sleep. 2016.
35. Choi JH, Kim EJ, Kim YS, et al. Effectiveness of nasal surgery alone on sleep quality, architecture, position, and sleep-disordered breathing in obstructive sleep apnea syndrome with nasal obstruction. Am J Rhinol Allergy. 2011; 25(5): 338–341.
36. Chai-Coetzer CL, Antic NA, Hamilton GS, et al. Physician decision making and clinical outcomes with laboratory polysomnography or limited-channel sleep studies for obstructive sleep apnea: a randomized trial. Ann Intern Med. 2017; 166(5): 332–340.

SUPPLEMENTARY MATERIAL

Supplementary material is available at *SLEEP* online.

FUNDING

This work was supported by the National Health and Medical Research Council of Australia project grant (1064163). SAL is supported by the NHMRC Centre of Research Excellence, NeuroSleep. SAS was supported by an NHMRC Early Career Fellowship and R.G. Menzies award (1053201), the American Heart Association (15SDG25890059), and an American Thoracic Society Foundation Unrestricted Grant. BAE was supported by the National Health and Medical Research Council (NHMRC) of Australia's CJ Martin Overseas Biomedical Fellowship (1035115) and is now supported by a Heart Foundation of Australia Future Leader Fellowship (101167).

SUBMISSION & CORRESPONDENCE INFORMATION

Submitted for publication November, 2016
Submitted in final revised form March, 2017
Accepted for publication May, 2017

Address correspondence to: Simon A. Joosten, MBBS, BMedSc, PhD, Monash Lung and Sleep, Monash Medical Centre, 246 Clayton Rd, Clayton, Victoria, 3168 Australia. Telephone: 613 95942900; Fax: 613 95946311. Email: drjoosten@hotmail.com

DISCLOSURE STATEMENT

None declared.