

Transoral robotic surgery for obstructive sleep apnea: a single-institution learning-curve and safety analysis

Chung Wui Loi

Universiti Malaya

Hui Yan Ong

Universiti Malaya

Liang Chye Goh

juliusglc@um.edu.my

Universiti Malaya

Siti Nadzrah Yunus

Universiti Malaya

Lim Chee Chean

Universiti Malaya

Maisarah Najihah Ridzuan

Universiti Malaya

Nur Fakhriah Izzati Mohamad Fauzi

Universiti Malaya

Daniella Rossellini Saifuzzman

Universiti Malaya

Mohd Zulkiflee Abu Bakar

Universiti Malaya

Research Article

Keywords: Robotic Surgical Procedures, Sleep Apnea, Obstructive, Learning Curve, Operative Time, Polysomnography, Otorhinolaryngologic Surgical Procedures

Posted Date: September 24th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10986571/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Purpose

In multilevel surgery for obstructive sleep apnea (OSA), total operative time poorly reflects the robotic learning curve because most of each case is non-robotic. We report the learning curve and safety profile of our first year of transoral robotic surgery (TORS), using console time as the primary learning-curve measure, with a concurrent conventional-surgery cohort providing institutional context rather than a controlled comparison.

Methods

We analyzed 26 consecutive TORS cases for OSA (February-December 2025) allocated by drug-induced sleep endoscopy (DISE) collapse pattern, and 26 consecutive conventional multilevel OSA operations (September 2024-October 2025) at the same institution. Comparisons were exploratory, two-sided, and uncorrected.

Results

Console time fell from 51.6 ± 23.6 minutes in the first half to 31.2 ± 13.5 minutes in the second half ($p = 0.018$), while total operative time did not. Every TORS case included tongue-base, lingual tonsil, or epiglottic surgery versus none ($p < 0.001$); groups were similar in age, body mass index, and OSA severity. Hospital stay, pain scores, and complication rates (11.5% in each group, $p = 1.000$) were similar, and none required tracheostomy. Paired AHI change was inconsistent (TORS: 4 of 11 met Sher criteria; conventional: 2 of 4). Paired Epworth scores improved significantly in TORS ($10.0 \rightarrow 5.9$, $p = 0.001$) but not significantly in conventional ($10.8 \rightarrow 8.0$, $p = 0.053$).

Conclusion

TORS was introduced without a detectable safety penalty despite greater anatomic extent, and its learning curve was visible only in console time, which we propose as the learning-curve metric for TORS in multilevel OSA surgery.

Introduction

Obstructive sleep apnea (OSA) is characterized by repetitive collapse of the upper airway during sleep, producing intermittent hypoxia, sleep fragmentation, daytime somnolence, and elevated cardiovascular risk [1]. The sites of collapse most often responsible are the soft palate and the tongue base, and retrolingual obstruction is particularly likely in patients with lingual tonsil hypertrophy.

Continuous positive airway pressure remains first-line therapy, but adherence is its central weakness: by the conventional threshold of more than four hours of use per night, 46–83% of patients are non-adherent [2]. Surgery is an option for selected patients who cannot use the device, yet the operations traditionally available address the palate far better than the tongue base. In the meta-analysis that established the surgical success criteria still in general use (postoperative apnea-hypopnea index [AHI] below 20 events/hour with at least a 50% reduction), Sher and colleagues found that response to palate-level surgery depended heavily on the anatomic site of collapse [3].

Transoral robotic surgery (TORS), first described for base-of-tongue neoplasms [4], brought three-dimensional visualization and instrument articulation to a region that is difficult to reach with line-of-sight transoral techniques. Its application to OSA followed. Vicini and colleagues reported the first preliminary experience of robotic tongue-base resection in OSA-hypopnea syndrome, with mean AHI falling from 38.3 to 20.6 events/hour and few complications [5], and later anchored the field's largest multicenter experience, 243 operations across seven centers, in which mean AHI fell from 43.0 to 17.9 events/hour [6]. Toh and colleagues, in the reference Southeast Asian series, combined TORS tongue-base reduction and partial epiglottidectomy with palatal surgery in 40 Asian patients and observed cure in 35% and surgical success in a further 55% of the 20 patients with complete paired polysomnography [7]. Baptista and colleagues [8] selected 64 patients with retrolingual collapse using drug-induced sleep endoscopy (DISE) graded by the VOTE (Velum, Oropharynx, Tongue base, Epiglottis) classification [9], reporting significant AHI and Epworth improvements and a 35.5% success rate by Sher criteria, and Toppenberg and colleagues described an 86% success rate with preserved swallowing quality of life in 56 patients [10]. Safety data are dominated by the 166-patient series of Glazer and colleagues, in which complications were predicted by American Society of Anesthesiologists (ASA) class and by the number of procedures performed in the same sitting rather than by age, body mass index, or OSA severity [11]. Functional and swallowing recovery after TORS has separately been shown to be good in the oropharyngeal cancer population [12].

TORS for OSA is almost never a stand-alone operation; the robotic resection is typically one component of a multilevel procedure that also includes tonsillectomy, uvulopalatopharyngoplasty, or nasal surgery under the same anesthetic. Learning curves for new programs are nonetheless usually described with total operative time, a measure dominated by the conventional portions of the case. Console time, the interval during which the surgeon is docked and operating at the robotic console, isolates the only part of the operation in which robotic proficiency can be expected to grow, and few TORS-for-OSA series report it separately.

TORS was introduced at Universiti Malaya Medical Centre (UMMC), Kuala Lumpur, in February 2025, extending the institution's transoral surgical reach to the tongue base and epiglottis for the first time. Here we describe the perioperative outcomes, safety profile, and learning curve of the first program year, using console time as the primary learning-curve measure, and we set these outcomes against a cohort of patients treated with the conventional techniques previously available at this center. Our aim is to document what the introduction of TORS changed at this institution, including which patients could be

offered anatomically directed surgery, and to report preliminary sleep-study outcomes transparently, rather than to claim superiority over conventional surgery, a comparison our non-randomized design cannot support.

Materials and methods

Study design and setting

This is a single-institution case series with a comparison cohort, conducted at UMMC, a tertiary referral center in Kuala Lumpur, Malaysia. The TORS cohort comprises all 26 consecutive TORS procedures for OSA performed between the program's first case on 19 February 2025 and 18 December 2025, with data collected prospectively. The comparison cohort comprises 26 consecutive patients who underwent conventional, non-robotic surgery for OSA at the same institution between September 2024 and October 2025. 16 of these operations (62%) took place during the TORS program period and 10 (38%) preceded the first TORS case. No a priori sample size calculation was performed. After initial collection the dataset was audited and supplemented with additional postoperative follow-up data. Operative dates and demographic entries were verified against admission and discharge records before analysis, and discrepancies were resolved from the clinical record.

Participants and allocation

Patients were eligible for TORS if they were aged 18–65 years and had symptomatic OSA with poor compliance or intolerance to continuous positive airway pressure (use below four hours per night), lingual tonsil hypertrophy of Friedman grade 3 or 4 [13], ASA physical status below 3, and no micrognathia. All patients underwent preoperative DISE, with collapse at each of the four VOTE sites (velum, oropharynx, tongue base, epiglottis) graded by configuration and degree per the VOTE framework [9], and a jaw-thrust maneuver performed to simulate the airway response to mandibular advancement. Those with collapse involving the tongue base or epiglottis were directed to TORS where-else those with collapse confined to the velum or oropharynx, and those treated before the robotic program began, underwent conventional surgery. Allocation was therefore anatomically directed and, for part of the comparison cohort, determined by treatment era. The groups are not a matched comparison, and we analyze them accordingly.

Surgical technique

TORS procedures were performed with the da Vinci Xi Surgical System (Intuitive Surgical, Sunnyvale, CA) and comprised lingual tonsillectomy, base-of-tongue reduction, or partial epiglottectomy, combined in the same sitting with palatal or tonsillar surgery (uvulopalatopharyngoplasty, tonsillectomy, lateral pharyngoplasty, partial uvulectomy) as indicated. Conventional procedures comprised uvulopalatopharyngoplasty, tonsillectomy or adenotonsillectomy, and septoplasty or turbinate surgery,

alone or in combination. By definition, no conventional case addressed the tongue base, lingual tonsil, or epiglottis.

Data collection and outcome measures

Data collection and outcome measures

Data were extracted from electronic medical records, operative and anesthetic notes, histopathology reports, and clinic documentation onto a standardized proforma. Baseline variables were age, sex, self-reported race, body mass index, Epworth Sleepiness Scale (ESS) score, and OSA severity as documented in the clinical diagnosis. Operative variables were total operative time, console time (TORS only, defined as the robot-docked interval), and estimated blood loss, which was recorded qualitatively. Postoperative variables were total and postoperative length of stay, inpatient and discharge pain scores, and all documented complications, including any return to the operating theatre and tracheostomy. A follow-up objective sleep study was available for 11 of 26 TORS patients (42%) and 4 of 26 comparison patients (15%); paired AHI values are reported descriptively for these subsets and were not standardized for modality or interval (see Limitations). A follow-up Epworth score, taken from clinic documentation rather than a fixed-interval research instrument, was documented in 22 of 26 patients per group.

Statistical analysis

Continuous variables are presented as mean $\pm$ standard deviation and compared with the two-sided Mann-Whitney U test; categorical variables are presented as n (%) and compared with Fisher's exact test, with $p < 0.05$ considered significant. Given the sample size and the non-randomized allocation, all between-group comparisons are exploratory and uncorrected for multiplicity. Because every TORS case included tongue-base, lingual tonsil, or epiglottic surgery that no conventional case attempted, operative time is reported descriptively and not tested between groups. The learning curve was assessed within the TORS cohort only: console times of the first and second halves of the chronological sequence were compared with the Mann-Whitney U test, and monotonic trend across the full sequence was assessed with Spearman correlation, applied identically to total operative time as a control comparison. Paired change in AHI and Epworth score was tested with the Wilcoxon signed-rank test within each cohort. Analyses used SPSS Version 30.0.

Ethics

The study was conducted in accordance with the Declaration of Helsinki and approved by the Medical Research Ethics Committee, Universiti Malaya Medical Centre (reference 202639 – 16485). Written informed consent was obtained from all patients. Data were de-identified before analysis.

Results

Cohorts

Twenty-six patients underwent TORS between 19 February and 18 December 2025, and 26 underwent conventional surgery between 25 September 2024 and 1 October 2025. Case volume in the robotic program accelerated as the year progressed: the first 13 TORS cases accrued over 33 weeks and the last 13 over 10 weeks. Baseline characteristics were similar between groups (Table 1). Mean age was 38.4 ± 12.0 versus 36.3 ± 11.9 years ($p = 0.364$) and mean body mass index 32.5 ± 6.7 versus 33.4 ± 6.3 kg/m² ($p = 0.927$); ESS scores and the proportion with severe OSA did not differ significantly.

Table 1
Baseline characteristics

Characteristic	TORS (n = 26)	Conventional (n = 26)	p
Age, years — mean (SD)	38.4 (12.0)	36.3 (11.9)	0.364
Female sex — n (%)	15 (57.7%)	11 (42.3%)	0.406
Race — n (%): Malay	17 (65.4%)	18 (69.2%)	—
Chinese	5 (19.2%)	6 (23.1%)	—
Indian	4 (15.4%)	2 (7.7%)	—
Body mass index, kg/m ² — mean (SD)	32.5 (6.7)	33.4 (6.3)	0.927
Epworth Sleepiness Scale, n = 22/group — mean (SD)*	10.0 (4.4)	10.8 (5.7)	0.564
Severe OSA — n (%)	12 (46.2%)	15 (57.7%)	0.579
<i>SD, standard deviation. OSA severity taken from the documented clinical diagnosis. *ESS documented in 22 of 26 patients per group. p-values: Mann-Whitney U (continuous), Fisher's exact (categorical); exploratory, uncorrected.</i>			

Operative time and the learning curve

Mean total operative time was 117.1 ± 40.9 minutes for TORS and 95.3 ± 40.7 minutes for conventional surgery (documented in 25 of 26 conventional cases); because the procedures differ systematically in anatomic scope, this contrast is descriptive only. Mean console time was 41.4 ± 21.5 minutes. Console time fell from 51.6 ± 23.6 minutes across the first 13 cases to 31.2 ± 13.5 minutes across the last 13 ($p = 0.018$) and declined monotonically over the full sequence (Spearman $\rho = -0.525$, $p = 0.006$). Total operative time showed no corresponding trend ($\rho = 0.076$, $p = 0.713$) (Fig. 1). Blood loss was documented as minimal in all but one TORS case (500 mL, concurrent genioplasty) and two conventional cases (100 mL and 200 mL).

Hospital stay, pain, and complications

Total hospital stay (3.4 ± 0.6 vs 3.3 ± 0.7 days, $p = 0.388$; Fig. 2B) and postoperative stay (1.4 ± 0.6 vs 1.3 ± 0.7 days, $p = 0.451$) were similar between groups, as were inpatient pain scores (1.3 ± 1.5 vs 1.8 ± 0.8 , $p = 0.051$) and pain scores at discharge (1.0 ± 0.3 vs 1.2 ± 0.7 , $p = 0.457$) (Table 2).

Table 2
Operative and perioperative outcomes

Outcome	TORS (n = 26)	Conventional (n = 26)	p
Operative time, min – mean (SD) ^{†‡}	117.1 (40.9)	95.3 (40.7)	—
Console time, min – mean (SD)	41.4 (21.5)	N/A	—
Console time, first 13 cases – mean (SD)	51.6 (23.6)	N/A	
Console time, last 13 cases – mean (SD)	31.2 (13.5)	N/A	0.018
Total hospital stay, days – mean (SD)	3.4 (0.6)	3.3 (0.7)	0.388
Post-operative stay, days – mean (SD)	1.4 (0.6)	1.3 (0.7)	0.451
Post-operative pain score – mean (SD)	1.3 (1.5)	1.8 (0.8)	0.051
Pain score at discharge – mean (SD)	1.0 (0.3)	1.2 (0.7)	0.457
Blood loss "minimal" (qualitative) – n (%)§	25 (96.2%)	24 (92.3%)	—
[†] Reported descriptively; the groups were not matched for procedural scope (every TORS case included tongue-base, lingual tonsil, or epiglottic surgery; Table 3) and operative time was not tested between groups. The console-time half-series comparison is within the TORS cohort (Mann-Whitney U). [‡] n = 25 for the conventional group; one case lacked a documented operative time. [§] One TORS case (concurrent genioplasty) recorded 500 mL; two conventional cases recorded a specific volume rather than "minimal" (100 mL and 200 mL) and are excluded from this count.			

Table 3
Complications and anatomic extent

Outcome	TORS (n = 26)	Conventional (n = 26)	p
Any documented complication – n (%)	3 (11.5%)	3 (11.5%)	1.000
Return to operating theatre – n (%)	1 (3.8%)	2 (7.7%)	—
Tracheostomy – n (%)	0 (0.0%)	0 (0.0%)	—
Tongue-base / lingual tonsil / epiglottic component addressed – n (%)	26 (100.0%)	0 (0.0%)	< 0.001
<i>Tongue-base/lingual tonsil/epiglottic component: lingual tonsillectomy, base-of-tongue reduction, or partial epiglottectomy performed as part of the procedure.</i>			

A complication was documented in 3 of 26 TORS cases (11.5%) and 3 of 26 conventional cases (11.5%; p = 1.000), and no patient in either group required tracheostomy (Table 3). In the TORS group, one patient developed a wound infection after palatine and lingual tonsillectomy requiring ward readmission for

intravenous antibiotics; one patient was readmitted with pain (score 8) subsequently attributed to an early postoperative infection; and one patient returned to the operating theatre on postoperative day 3 for secondary hemorrhage. At examination under anesthesia in that case, the bleeding arose from the palatine tonsillar fossae, and the robotic resection sites at the tongue base and epiglottis showed no active bleeding. In the conventional group, one patient underwent hemostasis for a tonsillar bleed seven hours after surgery, one required examination under anesthesia for epistaxis after turbinoplasty, and one had a suspected intraoperative cardiac event, likely unrelated to the surgical procedure, requiring postoperative monitoring in the intensive care unit.

Every TORS case included a tongue-base, lingual tonsil, or epiglottic component, against none of the conventional cases (100% vs 0%, $p < 0.001$; Table 3). This difference describes the anatomic reach of the two approaches rather than an independent treatment effect, and it is examined further in the Discussion.

Exploratory sleep-study outcomes

Among 11 TORS patients with a paired follow-up AHI, the mean fell from 46.4 ± 42.9 to 27.6 ± 26.3 events/hour, a change that did not reach significance (Wilcoxon $p = 0.206$). Individual trajectories varied widely (Fig. 3A): four patients met Sher success criteria, four had a higher AHI after surgery than before, and the remaining three had AHI reduction (21.9–29.2%) that fell short of the 50% threshold. Among 4 conventional patients with a paired follow-up AHI, the mean fell from 37.0 ± 26.5 to 16.9 ± 18.4 events/hour ($p = 0.25$); two met Sher success criteria, one improved substantially without crossing below the AHI 20 threshold (74.4 to 27.1), and one was essentially unchanged (37.0 to 37.7). These AHI data are hypothesis-generating.

Paired Epworth scores showed a more complete, favorable pattern (Fig. 3B, C). Among 22 TORS patients with a documented follow-up score, the mean fell from 10.0 ± 4.4 to 5.9 ± 2.3 (Wilcoxon $p = 0.001$), and 20 of 22 (91%) normalized to below the ESS 10 threshold for excessive daytime sleepiness. Among 22 conventional patients, the mean fell from 10.8 ± 5.7 to 8.0 ± 2.4 ($p = 0.053$), with 18 of 22 (82%) normalizing.

Discussion

Two findings summarize this first program year. TORS was adopted without a detectable safety penalty whereby hospital stay, pain scores, and complication rates were not only statistically indistinguishable but numerically identical between groups (11.5% each), even though every robotic case involved more extensive anatomic work than any conventional case. And a learning curve was present and significant, but only console time revealed it.

How a learning curve is measured determines whether it can be seen. In oncologic TORS, where the robotic resection essentially is the operation, institutional series show whole-case times falling steeply, by 47% across 168 cases in one four-year experience [14], and CUSUM analyses locate proficiency

inflections at roughly 20 to 40 cases using resection-phase rather than whole-case times [15]. A recent scoping review of the TORS learning curve placed inflection points between 25 and 50 cases but found the underlying time definitions inconsistent from study to study [16]. Multilevel OSA surgery is the setting in which that inconsistency matters most. In our series the robotic component accounted for barely a third of the mean case length, and total operative time, bounded by tonsillectomy and palatal work that no amount of console experience shortens, stayed flat while console time fell by 40%. A CUSUM analysis of console time against its own series mean identified case 3 as the proficiency inflection point, with mean console time falling from 88.7 minutes across the first three cases to 35.3 minutes across the remaining 23 (Fig. 1B); with only three cases preceding it, this inflection estimate is imprecise, but consistent with a steep initial gain followed by a longer plateau of incremental refinement. The same difference has been documented outside the head and neck surgery. In a consecutive robotic colorectal series, console time declined significantly across the learning curve while total operative time did not [17]. Even the largest multicenter TORS-for-OSA experience reported operative time without isolating the console interval [6], and we are not aware of a prior TORS-for-OSA series reporting a console-time learning curve. A new program that audits itself on total operative time alone is therefore likely to conclude, that no learning is occurring.

The comparison cohort is best understood as a measure of institutional reach rather than a rival treatment arm. Ten of its 26 patients were treated before February 2025, when no technique for reaching the tongue base transorally existed at this center; in that era a patient with DISE-confirmed retrolingual collapse could be offered palate-level surgery or no surgery at all. After the program began, allocation followed the DISE collapse pattern, a practice with support beyond institutional preference: DISE changes the surgical plan relative to awake examination in roughly half of patients, most often at hypopharyngeal and tongue-base structures, although whether DISE-directed selection improves eventual outcomes remains unproven [18]. The 100% versus 0% difference in tongue-base surgery (Table 3) is therefore a description of what changed at this institution, not a finding about comparative efficacy. What a first-year robotic program contributed here was not faster or demonstrably better surgery, but surgery in a place it could not previously go.

Our safety profile compares favorably with published experience. The two dedicated meta-analyses report pooled complication rates of 21.3% and 22.3% after TORS for OSA, with bleeding near 4% and transient dysphagia the most common event [19, 20]; our 11.5% falls below that range. Baptista and colleagues reported 15.8%, with every bleeding complication in their first 15 cases, a clustering they attributed to the institutional learning curve [8], and Glazer and colleagues found complications tracked ASA class and the number of procedures performed in one sitting [11]; our cases were multilevel by design. The anatomy of our single hemorrhage also matches the literature. In a 509-patient administrative cohort, post-TORS hemorrhage occurred in 8% of patients at a median of nine days, and the tonsillar subsite was the independent risk factor for bleeding [21]; adult tonsillectomy itself carries a substantially higher hemorrhage rate than the pediatric operation in prospective multicenter data [22]. Our patient bled on day 3 from the palatine tonsillar fossae, with the robotic resection sites clean at re-exploration. In multilevel TORS, a meaningful share of the hazard belongs to the conventional

components of the operation, and attributing it to the robot would mislead both audit and consent. Continued prospective monitoring as volume accrues remains more appropriate than treating a first-year rate as steady state.

The sleep-study data warrant a candid reading against both the literature and the measurement itself. Pooled success after TORS for OSA ranges from 48.2% to 69% across meta-analyses, depending on the success definition and study mix [20, 23, 24], with pooled cure near 24% [23]; our exploratory 4 of 11 sits below that band, while the 90% combined response of an established Singaporean program [7] illustrates what mature case selection can deliver. The comparison cohort's paired AHI data, available for only 4 of 26 patients, are too sparse to interpret as a group result; two met Sher criteria, one improved substantially without crossing below the 20-event threshold, and one showed essentially no change. Before reading any of these trajectories literally, the instability of the AHI deserves weight. In a multi-night cohort of more than ten thousand adults, a single night misclassified the severity category of roughly one in five patients with mild or moderate disease [25]; in repeated home testing, the AHI varied by a mean of 57% between nights, and 12–14% of untreated patients met common surgical success criteria on a second night with no intervention at all [26]. Our paired comparisons rest on one study before and one after surgery, unmatched for modality and interval, so part of the apparent deterioration, and equally part of the apparent improvement, is likely measurement noise. Both AHI follow-up subsets may also carry an informative-dropout bias in an unexpected direction: patients who felt substantially better were, in our experience, less likely to pursue a repeat objective study than those with persistent symptoms, which would enrich the subset for less complete responders and bias the observed change toward the null rather than overstate benefit. These results establish neither efficacy nor its absence; they establish the need for the standardized, complete-cohort follow-up now under way.

Symptomatic outcomes were more complete and favorable than the AHI data (Fig. 3B, C): the paired Epworth subset was nearly complete, with significant improvement after TORS and a similar, non-significant trend after conventional surgery. That subjective improvement was common even among patients whose paired AHI did not improve is not necessarily contradictory: some patients in both cohorts underwent surgery as the first stage of a planned two-stage procedure, with skeletal or nasal surgery deferred to a second sitting (one of the two TORS non-normalizers, for example, has an explicit plan for further nasal surgery not yet performed). Postoperative CPAP use, documented in several patients, is a related confound: a normalized AHI or Epworth score may reflect surgery having enabled CPAP adherence, not a surgery-alone effect. Neither factor could be identified across the cohort, so both remain qualitative caveats.

TORS is also no longer the only surgical route to the tongue base. Hypoglossal nerve stimulation reduced median AHI by 68% in its pivotal trial [27] and has outperformed TORS on AHI reduction and success in a direct institutional comparison (86.8% versus 54.2%) [28] and in network meta-analysis [29]. For patients eligible for both hypoglossal nerve stimulation and TORS, the current evidence favors stimulation. Eligibility and cost decide how often that choice actually arises. Selection criteria exclude a body mass index of 35 kg/m² or more and complete concentric palatal collapse on DISE [30, 31]; 10 of

our 26 TORS patients (38%) exceeded that threshold alone. Cost compounds this barrier: in a recent single-center analysis, total costs were 25,583 US dollars for stimulation against 5,833 for TORS, the difference driven almost entirely by the implanted hardware [31]. In a middle-income healthcare system, anatomically directed surgery therefore remains the realistic alternative to positive airway pressure for a substantial share of patients, and that is the treatment-algorithm context in which this series belongs.

This series also adds to regional literature. Asian patients develop OSA with a more restricted craniofacial skeleton than Caucasian patients and show more severe disease at the same degree of obesity [32], an anatomic phenotype that shapes both the disease and the surgical access of the populations this center serves. Reported TORS-for-OSA outcomes from Southeast Asia rest largely on the Singaporean experience [7]; a Malaysian cohort drawn from Malay, Chinese, and Indian populations at a mean body mass index above 32 kg/m² supplies reference perioperative values for the regional programs most likely to face similar anatomy and similar resource constraints.

Limitations

This is a learning-curve and safety analysis with exploratory efficacy outcomes, not a powered comparative-effectiveness study, and several design features constrain what it can show. Allocation was not randomized: patients were directed to TORS or conventional surgery by DISE-identified collapse pattern and, for ten comparison patients, by treatment era before the robotic program existed; the conventional cohort is therefore a convenience cohort, not a true control arm, and between-group tests are exploratory and uncorrected for multiplicity. All surgery was performed by a single surgeon at a single institution, so these results describe one program's early trajectory rather than a generalizable learning curve. Follow-up objective sleep-study data were available for only 11 of 26 TORS patients (42%) and 4 of 26 conventional patients (15%) and were not standardized for modality (in-laboratory polysomnography versus home sleep apnea testing) or interval; as discussed above, these subsets may also be enriched for less complete responders through informative dropout. Paired Epworth data were more complete (22 of 26 patients per group) but were extracted from clinic documentation rather than a fixed-interval research instrument. OSA severity itself was classified from the documented diagnosis rather than re-derived from raw polysomnography. Some patients in both cohorts were awaiting a planned second-stage procedure or using adjunctive CPAP postoperatively; only one could be confidently identified from available records, so this remains a qualitative caveat. Twenty-six patients per arm at a single center also limit statistical power, and operative time could not be compared inferentially at all given the scope mismatch between arms. Finally, the entire series describes a program's first year, during which selection and technique were themselves evolving; steady-state results may differ.

Conclusion

TORS for OSA was introduced at this institution without a detectable safety penalty relative to conventional surgery, despite consistently greater anatomic extent. The learning curve was visible only in console time, which fell by 40% within 26 cases while total operative time stayed flat; we propose

console time as the learning-curve metric of choice for TORS embedded in multilevel surgery. The program's principal early contribution has been anatomic, extending surgery to the tongue base and epiglottis for patients selected by DISE. Whether that reach translates into durable sleep outcomes is a question for the standardized follow-up this first-year experience now makes possible.

Declarations

Funding None.

Conflicts of interest None.

Ethics approval Medical Research Ethics Committee, Universiti Malaya Medical Centre (reference 202639-16485); conducted in accordance with the Declaration of Helsinki.

Consent Written informed consent was obtained from all patients.

Data availability The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

Authors' contributions Conceptualization: LCG, CWL, MZAB. Methodology: LCG, SNY, MZAB. Formal analysis: CWL, LCG. Investigation: CWL, HYO, LCG. Resources: LCG, LCC, MZAB. Data curation: CWL, HYO, MNR, NFI, DRS. Validation: LCG, CWL. Visualization: CWL, LCG. Writing—original draft: CWL, LCG. Writing—review and editing: CWL, HYO, LCG, SNY, LCC, MNR, NFI, DRS, MZAB. Supervision: LCC, MZAB, LCG. Project administration: CWL. CWL, Chung Wui Loi; HYO, Hui Yan Ong; LCG, Liang Chye Goh; SNY, Siti Nadzrah Yunus; LCC, Lim Chee Chean; MNR, Maisarah Najihah Ridzuan; NFI, Nur Fakhriah Izzati Mohamad Fauzi; DRS, Daniella Rossellini Saifuzzman; MZAB, Mohd Zulkiflee Abu Bakar. All authors read and approved the final manuscript.

Use of artificial intelligence A large language model (Claude, Anthropic) was used under the corresponding author's direction to assist with paraphrasing, grammar checking and spelling checks to improve on document readability. The authors had also employed AI to improve upon table and figure representations for better clarity. The corresponding author reviewed and edited all content and takes full responsibility for the accuracy and integrity of the work.

References

1. Jordan AS, McSharry DG, Malhotra A (2014) Adult obstructive sleep apnoea. *Lancet* 383:736–747. [https://doi.org/10.1016/S0140-6736\(13\)60734-5](https://doi.org/10.1016/S0140-6736(13)60734-5)
2. Weaver TE, Grunstein RR (2008) Adherence to continuous positive airway pressure therapy: the challenge to effective treatment. *Proc Am Thorac Soc* 5:173–178. <https://doi.org/10.1513/pats.200708-119MG>

3. Sher AE, Schechtman KB, Piccirillo JF (1996) The efficacy of surgical modifications of the upper airway in adults with obstructive sleep apnea syndrome. *Sleep* 19:156–177.
<https://doi.org/10.1093/sleep/19.2.156>
4. O'Malley BW Jr, Weinstein GS, Snyder W, Hockstein NG (2006) Transoral robotic surgery (TORS) for base of tongue neoplasms. *Laryngoscope* 116:1465–1472.
<https://doi.org/10.1097/01.mlg.0000227184.90514.1a>
5. Vicini C, Dallan I, Canzi P, Frassinetti S, La Pietra MG, Montevecchi F (2010) Transoral robotic tongue base resection in obstructive sleep apnoea-hypopnoea syndrome: a preliminary report. *ORL J Otorhinolaryngol Relat Spec* 72:22–27. <https://doi.org/10.1159/000284352>
6. Vicini C, Montevecchi F, Campanini A et al (2014) Clinical outcomes and complications associated with TORS for OSAHS: a benchmark for evaluating an emerging surgical technology in a targeted application for benign disease. *ORL J Otorhinolaryngol Relat Spec* 76:63–69.
<https://doi.org/10.1159/000360768>
7. Toh ST, Han HJ, Tay HN, Kiong KL (2014) Transoral robotic surgery for obstructive sleep apnea in Asian patients: a Singapore sleep centre experience. *JAMA Otolaryngol Head Neck Surg* 140:624–629. <https://doi.org/10.1001/jamaoto.2014.926>
8. Baptista PM, Diaz Zufiaurre N, Garaycochea O et al (2022) TORS as part of multilevel surgery in OSA: the importance of careful patient selection and outcomes. *J Clin Med* 11:990.
<https://doi.org/10.3390/jcm11040990>
9. Kezirian EJ, Hohenhorst W, de Vries N (2011) Drug-induced sleep endoscopy: the VOTE classification. *Eur Arch Otorhinolaryngol* 268:1233–1236. <https://doi.org/10.1007/s00405-011-1633-8>
10. Toppenberg AGL, Lodder WL, Plaat RE, Schwandt LQ (2025) Patient centered outcomes of transoral robotic surgery for obstructive sleep apnea: a retrospective analysis. *Eur Arch Otorhinolaryngol* 282:1459–1468. <https://doi.org/10.1007/s00405-024-09102-0>
11. Glazer TA, Hoff PT, Spector ME (2014) Transoral robotic surgery for obstructive sleep apnea: perioperative management and postoperative complications. *JAMA Otolaryngol Head Neck Surg* 140:1207–1212. <https://doi.org/10.1001/jamaoto.2014.2299>
12. Achim V, Bolognone RK, Palmer AD et al (2018) Long-term functional and quality-of-life outcomes after transoral robotic surgery in patients with oropharyngeal cancer. *JAMA Otolaryngol Head Neck Surg* 144:18–27. <https://doi.org/10.1001/jamaoto.2017.1790>
13. Friedman M, Yalamanchali S, Gorelick G, Joseph NJ, Hwang MS (2015) A standardized lingual tonsil grading system: interexaminer agreement. *Otolaryngol Head Neck Surg* 152:667–672.
<https://doi.org/10.1177/0194599815568970>
14. White HN, Frederick J, Zimmerman T, Carroll WR, Magnuson JS (2013) Learning curve for transoral robotic surgery: a 4-year analysis. *JAMA Otolaryngol Head Neck Surg* 139:564–567.
<https://doi.org/10.1001/jamaoto.2013.3007>

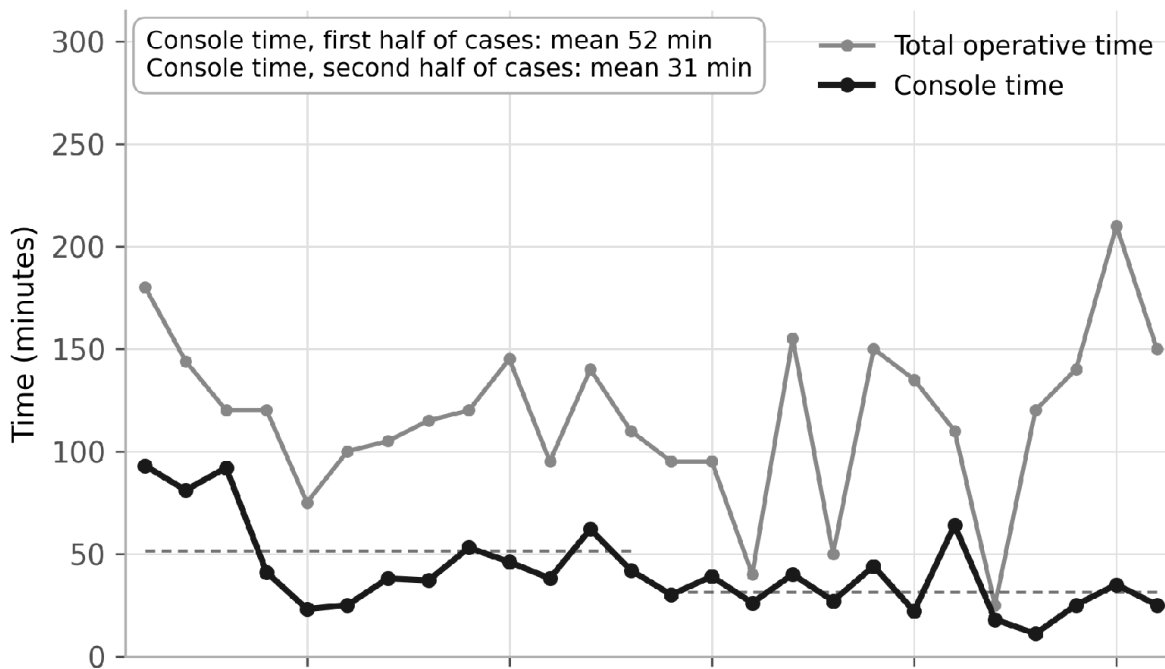
15. Albergotti WG, Gooding WE, Kubik MW et al (2017) Assessment of surgical learning curves in transoral robotic surgery for squamous cell carcinoma of the oropharynx. *JAMA Otolaryngol Head Neck Surg* 143:542–548. <https://doi.org/10.1001/jamaoto.2016.4132>
16. Lechien JR (2025) Learning process of transoral robotic surgery for head and neck cancers: a scoping review. *Oral Oncol* 171:107788. <https://doi.org/10.1016/j.oraloncology.2025.107788>
17. Nasser Y, Stettler I, Shen W et al (2021) Learning curve in robotic colorectal surgery. *J Robot Surg* 15:489–495. <https://doi.org/10.1007/s11701-020-01131-1>
18. Certal VF et al (2016) Awake examination versus DISE for surgical decision making in patients with OSA: a systematic review. *Laryngoscope* 126:768–774. <https://doi.org/10.1002/lary.25722>
19. Meccariello G et al (2017) Transoral robotic surgery for the management of obstructive sleep apnea: a systematic review and meta-analysis. *Eur Arch Otorhinolaryngol* 274:647–653. <https://doi.org/10.1007/s00405-016-4113-3>
20. Justin GA et al (2016) Transoral robotic surgery for obstructive sleep apnea: a systematic review and meta-analysis. *Otolaryngol Head Neck Surg* 154:835–846. <https://doi.org/10.1177/0194599816630962>
21. Zenga J et al (2017) Postoperative hemorrhage and hospital revisit after transoral robotic surgery. *Laryngoscope* 127:2287–2292. <https://doi.org/10.1002/lary.26626>
22. Sarny S, Ossimitz G, Habermann W, Stammberger H (2011) Hemorrhage following tonsil surgery: a multicenter prospective study. *Laryngoscope* 121:2553–2560. <https://doi.org/10.1002/lary.22347>
23. Miller SC et al (2017) Transoral robotic base of tongue reduction for obstructive sleep apnea: a systematic review and meta-analysis. *Laryngoscope* 127:258–265. <https://doi.org/10.1002/lary.26060>
24. Lechien JR et al (2021) Surgical, clinical, and functional outcomes of transoral robotic surgery used in sleep surgery for obstructive sleep apnea syndrome: a systematic review and meta-analysis. *Head Neck* 43:2216–2239. <https://doi.org/10.1002/hed.26702>
25. Punjabi NM, Patil S, Crainiceanu C, Aurora RN (2020) Variability and misclassification of sleep apnea severity based on multi-night testing. *Chest* 158:365–373. <https://doi.org/10.1016/j.chest.2020.01.039>
26. Tschopp S, Wimmer W, Caversaccio M, Borner U, Tschopp K (2021) Night-to-night variability in obstructive sleep apnea using peripheral arterial tonometry: a case for multiple night testing. *J Clin Sleep Med* 17:1751–1758. <https://doi.org/10.5664/jcsm.9300>
27. Strollo PJ Jr et al (2014) Upper-airway stimulation for obstructive sleep apnea. *N Engl J Med* 370:139–149. <https://doi.org/10.1056/NEJMoa1308659>
28. Huntley C et al (2019) Comparing upper airway stimulation to transoral robotic base of tongue resection for treatment of obstructive sleep apnea. *Laryngoscope* 129:1010–1013. <https://doi.org/10.1002/lary.27484>
29. Kim DH et al (2026) Comparison of upper airway stimulation, tongue base surgeries, and continuous positive airway pressure for moderate-to-severe obstructive sleep apnea: a systematic review and

network meta-analysis. *Sleep Breath*. <https://doi.org/10.1007/s11325-026-03772-2>

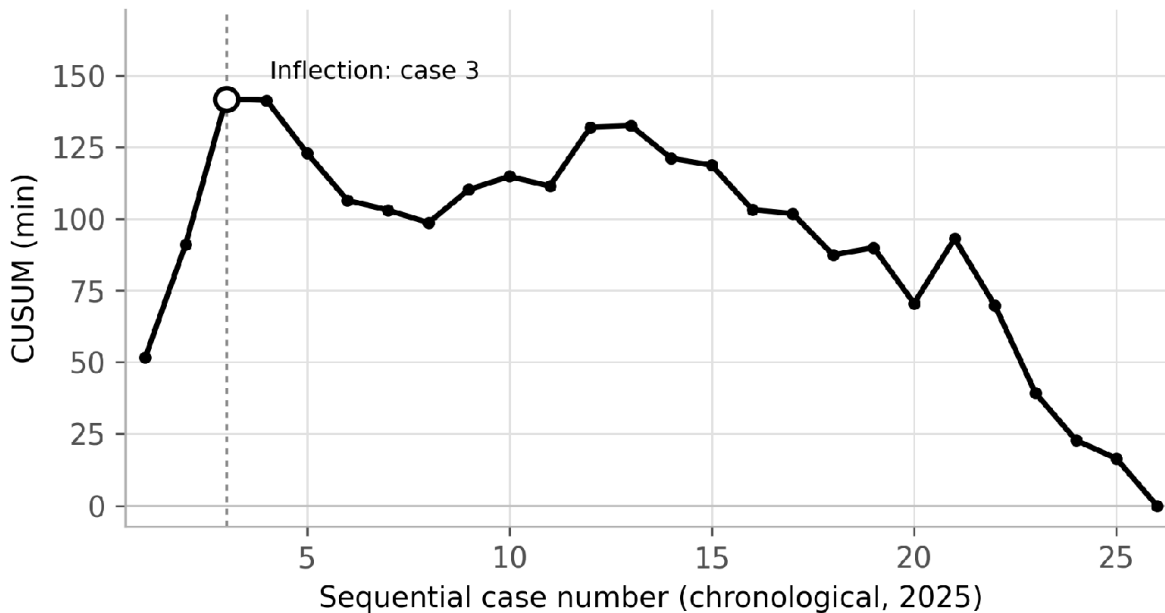
30. Vanderveken OM et al (2013) Evaluation of drug-induced sleep endoscopy as a patient selection tool for implanted upper airway stimulation for obstructive sleep apnea. *J Clin Sleep Med* 9:433–438. <https://doi.org/10.5664/jcsm.2658>
31. Mandloi S, Garg N, Naimi B et al (2025) Transoral robotic surgery versus hypoglossal nerve stimulation for OSA: a cost analysis study. *Laryngoscope* 135:457–462. <https://doi.org/10.1002/lary.31697>
32. Lee RWW, Vasudavan S, Hui DS et al (2010) Differences in craniofacial structures and obesity in Caucasian and Chinese patients with obstructive sleep apnea. *Sleep* 33:1075–1080. <https://doi.org/10.1093/sleep/33.8.1075>

Figures

A. Console time fell across the learning curve; total operative time did not



B. CUSUM of console time identifies the proficiency inflection



A: console time $\rho = -0.525$, $p = 0.006$; operative time $\rho = 0.076$, $p = 0.713$ (Spearman).
 B: CUSUM vs. overall mean (41 min); peak (case 3) = proficiency inflection.
 Mean console time 89 min before it ($n=3$) vs 35 min after ($n=23$); post hoc.

Figure 1

Learning curve for transoral robotic surgery. (A) Console time and total operative time across the chronological case sequence; dashed lines mark the mean console time of each half of the series. Console time declined across the sequence (Spearman $\rho = -0.525$, $p = 0.006$); total operative time, which includes concurrent non-robotic procedures, did not ($\rho = 0.076$, $p = 0.713$). (B) CUSUM of console time against its own series mean; the peak, after case 3, marks the proficiency inflection point (mean console

time 88.7 minutes for cases 1-3 versus 35.3 minutes for cases 4-26). Post hoc and hypothesis-generating given the small pre-inflection sample.

Perioperative outcomes by surgical approach

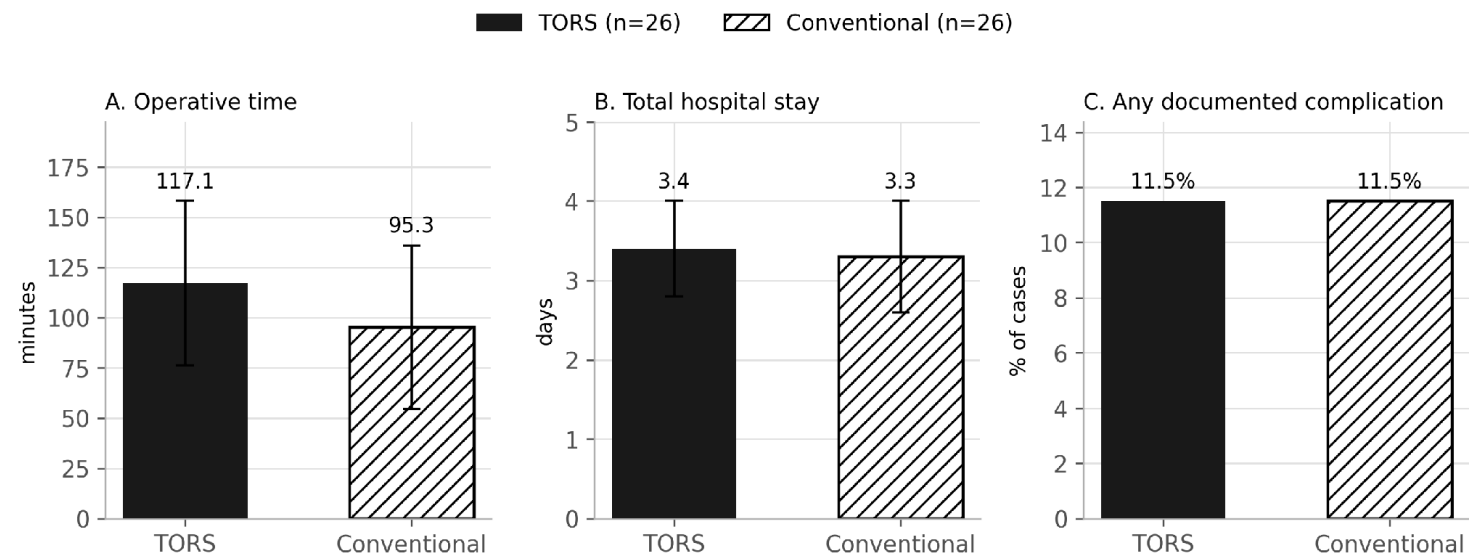
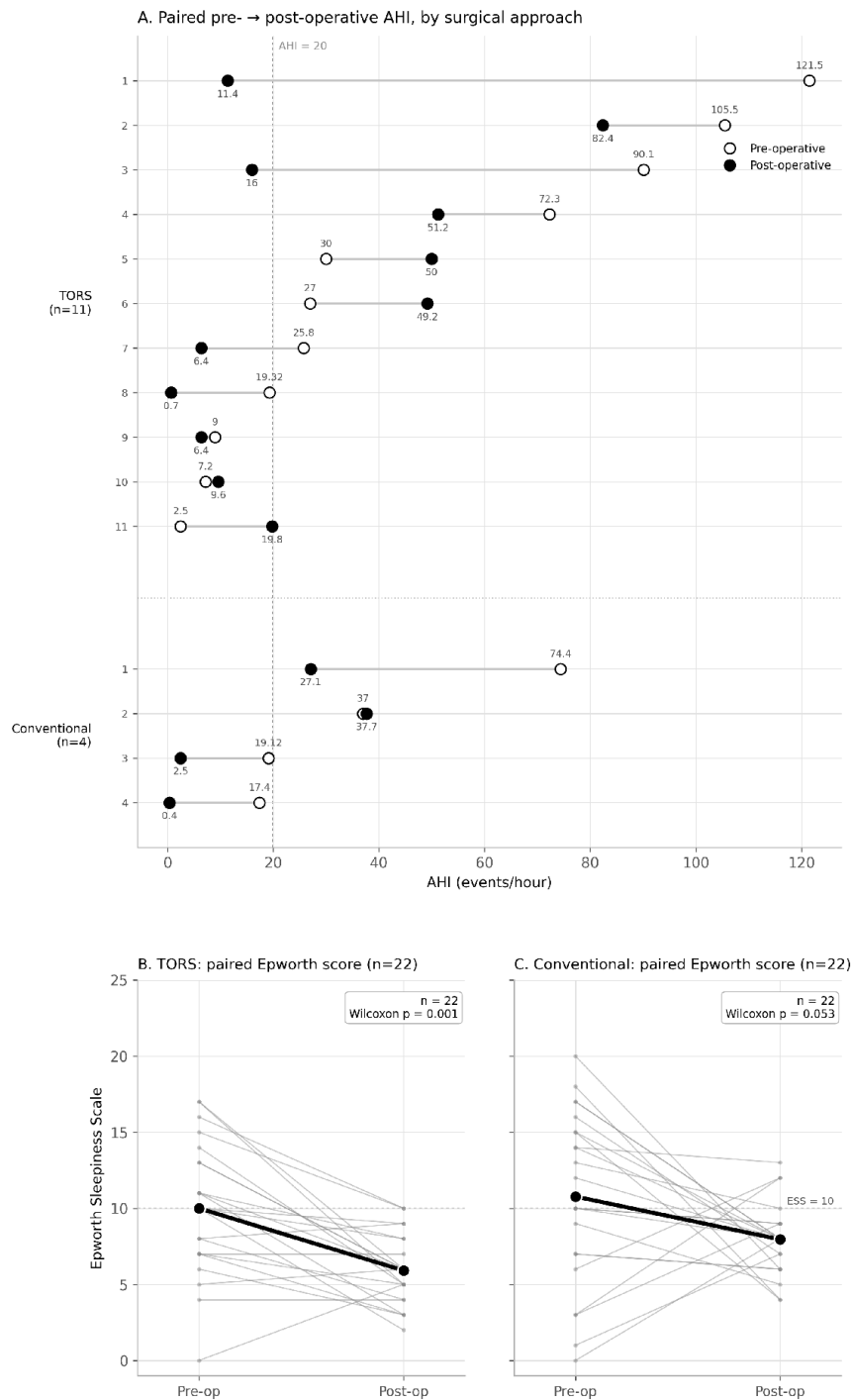


Figure 2

Perioperative outcomes by surgical approach. Panel A: operative time (descriptive only; groups differ in procedural scope, Table 3). Panel B: total hospital stay. Panel C: proportion of cases with any documented complication. Error bars are ± 1 SD (panels A-B).



Case order in (A) reflects de-identified sequence numbers, not chart/registration numbers, ranked by pre-operative AHI within each group.
 Pre/post AHI studies were not consistently matched for testing modality (in-laboratory PSG vs home sleep apnea test) or follow-up interval.
 Post-operative ESS was extracted from clinic follow-up notes rather than an instrument re-administered at a fixed interval; both are exploratory.

Figure 3

Paired sleep-study outcomes by surgical approach. (A) Paired pre- and postoperative AHI for the 11 TORS and 4 conventional patients with a completed follow-up sleep study; case labels are de-identified sequence numbers ranked by pre-operative AHI within each group, and the dashed reference line marks AHI 20 events/hour. (B, C) Paired pre- and postoperative Epworth Sleepiness Scale score for the 22 TORS and 22 conventional patients with a documented follow-up score; thin lines are individual patients and

the bold line is the group mean, with the dashed reference line marking the ESS 10 threshold for excessive daytime sleepiness. AHI studies were not consistently matched for modality or follow-up interval, postoperative Epworth scores were extracted from clinic notes rather than a fixed-interval research instrument, and some patients in both panels were awaiting a planned second-stage procedure or using adjunctive CPAP at the time of assessment; all panels are exploratory.