

UC Irvine

UC Irvine Previously Published Works

Title

Intersurgeon Variability in Proptosis Reduction After Orbital Decompression for Thyroid Eye Disease: A Multicenter Analysis.

Permalink

<https://escholarship.org/uc/item/1463c2zz>

Journal

Ophthalmic Plastic and Reconstructive Surgery

ISSN

0740-9303

Authors

Hong, Ashley

Shoji, Marissa K

Villatoro, George A

et al.

Publication Date

2025-10-13

DOI

10.1097/iop.0000000000003090

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Intersurgeon Variability in Proptosis Reduction After Orbital Decompression for Thyroid Eye Disease: A Multicenter Analysis

Ashley Hong, B.S.*†‡, Marissa K. Shoji, M.D.*†, George A. Villatoro, M.D.†, Bharanidharan Radha-Saseendrakumar, M.SC.†, Sally L. Baxter, M.D., M.SC.†, Peter J. Dolman, M.D.§, Michael Kazim, M.D.¶, Andrew R. Harrison, M.D.¶, Mark J. Lucarelli, M.D.#, James A. Garrity, M.D.***, Dinesh Selva, M.D.††, Lilangi S. Ediriwickrema, M.D.‡‡, Catherine Y. Liu, M.D., PH.D.*†, Bobby S. Korn, M.D., PH.D.*†§§, and Don O. Kikkawa, M.D., F.A.C.S.*†§§

*Division of Oculofacial Plastic and Reconstructive Surgery, University of California San Diego, La Jolla, California, U.S.A.; †Department of Ophthalmology, University of California San Diego, La Jolla, California, U.S.A.; ‡Rush Medical College, Chicago, Illinois, U.S.A.; §Department of Ophthalmology and Visual Sciences, University of British Columbia, Vancouver, British Columbia, Canada; ¶Department of Ophthalmology, Columbia University Irving Medical Center, New York, New York, U.S.A.; #Department of Ophthalmology and Visual Neurosciences, University of Minnesota, Minneapolis, Minnesota, U.S.A.; #Department of Ophthalmology and Visual Sciences, University of Wisconsin-Madison, Madison, Wisconsin, U.S.A.; ***Department of Ophthalmology, Mayo Clinic, Rochester, Minnesota, U.S.A.; ††Department of Ophthalmology, Royal Adelaide Hospital, Adelaide, South Australia, Australia; ‡‡Department of Ophthalmology, University of California Irvine, Irvine, California, U.S.A.; and §§Division of Plastic Surgery, University of California San Diego, La Jolla, California, U.S.A.

Purpose: This study assesses intersurgeon variability in proptosis reduction after orbital decompression for thyroid eye disease.

Methods: This multicenter retrospective study included patients with thyroid eye disease who underwent orbital decompression from 1 of 7 surgeons at 7 different institutions between January 2002 and December 2018. Data were included if a single decompression technique was performed on ≥ 10 patients by ≥ 2 surgeons. The primary outcome was postoperative change in proptosis with emphasis on comparison among surgeons utilizing similar surgical techniques. Statistical analysis was performed with χ^2 and ANOVA testing, and a multivariable logistic regression model was generated.

Results: Six hundred thirty-three orbits that underwent orbital decompression were included. Five different decompression techniques were analyzed: medial wall ($n = 29$), fat + lateral wall ($n = 113$), medial wall + floor ($n = 123$), fat + lateral wall + medial wall ($n = 140$), and fat + lateral wall + medial wall + floor ($n = 228$), without significant difference in proptosis reduction among surgeons. Surgeons did not demonstrate significant differences in outcomes at different time points compared with each other or themselves. Outcomes when comparing endoscopic versus open medial wall decompression varied among surgery types. Multivariate modeling revealed a statistically significant association between postoperative change in proptosis with preoperative proptosis ($p < 0.001$).

Conclusions: Postoperative change in proptosis did not differ significantly between surgeons utilizing similar orbital decompression techniques for patients with thyroid eye disease. This study may strengthen the statistical validity of multicenter clinical trials assessing orbital decompression outcomes performed by surgeons employing uniform surgical techniques, thereby advancing our understanding of optimal surgical management strategies for thyroid eye disease.

(*Ophthalmic Plast Reconstr Surg* 2015;XX:00–00)

Accepted for publication September 5, 2025.

No conflicts are relevant to this study; other: M.K.S.'s research is supported by the Heed Foundation; D.O.K. is a consultant for Amgen, Immunovant, Genentech, Acelyrin, Argenx, and Lassen Therapeutics and receives royalties from Elsevier; B.S.K. is a consultant/advisor to Amgen, Acelyrin, Argenx, Immunovant, and Viridian and receives royalties from Elsevier; C.Y.L. is a principal investigator for Lassen Therapeutics-sponsored clinical trial, Genentech-sponsored clinical trial, and Amgen-sponsored clinical trial and receives Wolters Kluwer's Health Royalties; A.R.H. reports compensation from RVL Pharmaceuticals and Amgen; P.J.D. is an advisor for Amgen, Immunovant, Sling Therapeutics, and Kriya; and M.J.L. is a consultant for Amgen and Viridian with unrestricted departmental grant for Research to Prevent Blindness.

Supported by the Bell Charitable Foundation, Rancho Santa Fe, CA; and Research to Prevent Blindness, New York, NY, grant NIH DP5OD29610.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's website (www.op-rs.com).

A.H. and M.K.S. are co-first authors.

Presented at the 2023 American Society of Oculofacial Plastic Surgery Fall Meeting on November 2023 in San Francisco, CA.

Address correspondence and reprint requests to Don O. Kikkawa, M.D., F.A.C.S., Shiley Eye Institute, 9350 Campus Point Dr, La Jolla, CA 92037. E-mail: dkikkawa@health.ucsd.edu

DOI: 10.1097/IOP.0000000000003090

Thyroid eye disease (TED), an autoimmune condition that causes enlargement of extraocular muscles and orbital soft tissue, may clinically present with proptosis, restrictive strabismus, lid retraction, and potentially vision-threatening compressive optic neuropathy or exposure keratopathy.^{1–5} Treatment of TED varies but may include glucocorticoids, immunomodulators, insulin growth factor-1 receptor inhibitors, and orbital radiotherapy. Although medical treatment may be effective in some patients, surgical orbital decompression remains a

mainstay of treatment and is typically performed in the setting of disfiguring proptosis or more urgently for compressive optic neuropathy.^{2,4} While different surgical approaches to orbital decompression, including removal of fat and the number of walls removed, may result in varying levels of proptosis reduction, surgeon-dependent variables may also exist.⁴⁻⁸ While literature exists surrounding the efficacy of orbital decompression for TED, no comparative evaluation has been conducted on the variability in outcomes among surgeons employing similar decompression techniques. This is important in standardizing and comparing outcomes in the literature and for future prospective multicenter trials studying outcomes of orbital decompression. Thus, this study aims to explore the intersurgeon variability in proptosis reduction after orbital decompression.

METHODS

This multicenter retrospective study included patients who underwent orbital decompression for TED. The study involved data collected over a 16-year period from January 1, 2002, to December 31, 2018, and analyzed orbital decompression results from 7 surgeons at 7 institutions utilizing similar established surgical techniques.^{4,5,9,10} The exclusion criteria were: 1) patients <18 years of age; 2) patients who received prior radiation or previously underwent orbital decompression; and 3) incomplete clinical data or imaging. This study adhered to the Declaration of Helsinki and was approved by the Institutional Review Board. All research conducted was Health Insurance Portability and Accountability Act-compliant.

Data collected included age, gender, smoking history, preoperative and postoperative exophthalmometry, clinical activity score (CAS), cross-sectional orbital area, and type of orbital decompression (eg, number of walls, fat or bone decompression, and endoscopic or open techniques). The patient was considered a smoker if the patient had smoked within 6 months of the surgery.⁴ The cross-sectional area of the orbit was measured ≈ 1.5 cm anterior to the superior orbital fissure on coronal computerized tomography as previously described.⁴ The surgeons were designated as A to G. All surgical procedures were conducted by fellowship-trained oculoplastic surgeons utilizing uniform surgical techniques, resulting in five primary decompression categories: medial wall alone, fat + lateral wall, medial wall + floor, fat + lateral wall + medial wall, and fat + lateral wall + medial wall + floor. All orbital decompressions followed established approaches described in the previous literature.^{4,5,9,10} To ensure meaningful statistical comparison, data were only considered when ≥ 2 surgeons performed a single technique on ≥ 10 patients.^{4,5,9,10} Surgeon, institutional, and patient data were deidentified and analyzed by masked investigators to mitigate potential biases originating from individual surgeons or institutions.⁴

The primary outcome was the postoperative change in proptosis (in millimeters). Outcomes were compared among surgeons performing the same decompression technique to assess intersurgeon variability. Secondary outcomes included preoperative or postoperative characteristics impacting proptosis, age, CAS, smoking history, radiographically measured orbital area, and follow-up duration.

Subanalyses included single-orbit analysis, comparison of endoscopic and open medial wall decompression, and temporal assessment. As TED may be asymmetric, orbits may undergo different surgical diseases and exhibit individual responses to decompression; thus, all orbits that underwent decompression were included in the overall analysis. To account for the potential inherent link between orbits undergoing binocular decompression, a single-orbit subanalysis was performed; in unilateral decompressions, the operated orbit was included, and in bilateral

decompressions, a single operated orbit (right) was selected for analysis. In addition, subgroup analysis was conducted with 1-way ANOVA to compare outcomes between endoscopic and open surgical approaches to medial wall decompression across surgical technique categories: medial wall only, medial wall + floor, and fat + lateral + medial wall decompression. Last, a subanalysis was performed to assess temporal trends in proptosis outcomes across multiple surgeons performing similar orbital decompression surgeries. Cases were stratified into 3 distinct periods based on the date of operation. Time point 1 encompassed operations performed from 2002 to 2007, time point 2 from 2008 to 2013, and time point 3 from 2014 to 2018. Patients with insufficient data including surgery date were excluded from the temporal analysis. Two additional surgeons were excluded from the temporal analysis due to a lack of time data.

Stata 14.2 and Excel were used to perform statistical analyses. Bivariate analyses of the measured categorical variables were performed using χ^2 testing, multiple ANOVA testing was used to analyze continuous variables involving ≥ 3 groups, and 2-tailed unpaired *t* test with unequal variances was used for comparisons between 2 groups. A multivariable logistic regression model was generated to predict postoperative change in proptosis with independent preoperative characteristics. For subanalyses assessing temporal trends, 1-way ANOVA testing was applied to evaluate if there was a difference in the change in proptosis across all surgeons performing the same surgical technique for each time point, as well as to assess the change in outcomes over time for the same surgeon doing the same procedure at time points 1, 2, and 3. Bonferroni correction was utilized to account for multiple comparisons.

RESULTS

Six hundred thirty-three orbits of 386 total patients who underwent orbital decompression (247 bilateral and 139 unilateral) by 7 different surgeons were included (Table 1). The average follow-up was 21.9 (range, 3–156) months. Analysis of 5 different decompression techniques included: 1) medial wall (*n* = 29), 2) fat + lateral wall (*n* = 113), 3) medial wall + floor (*n* = 123), 4) fat + lateral wall + medial wall (*n* = 140), and 5) fat + lateral wall + medial wall + floor (*n* = 228).^{4,5,9,10} No statistically significant difference was observed in change in proptosis among different surgeons performing similar surgical techniques, including medial wall surgery (D [*n* = 15]: -1.40 ± 2.06 mm; F [*n* = 14]: -2.29 ± 1.73 mm; *p* = 0.22), fat + lateral wall (A [*n* = 19]: -2.94 ± 2.43 mm; C [*n* = 14]: -3.57 ± 1.73 mm; D [*n* = 80]: -3.14 ± 1.45 mm; *p* = 0.55), medial wall + floor (E [*n* = 46]: -2.61 ± 1.50 mm; F [*n* = 23]: -3.67 ± 0.69 mm; G [*n* = 54]: -3.20 ± 1.79 mm; *p* = 0.06), fat + lateral wall + medial wall (A [*n* = 50]: -3.31 ± 2.43 mm; B [*n* = 51]: -3.76 ± 2.13 mm, C [*n* = 18]: -3.89 ± 1.37 mm; D [*n* = 21]: -3.10 ± 1.69 mm; *p* = 0.45), and fat + lateral wall + medial wall + floor (A [*n* = 94]: -4.55 ± 4.43 mm; B [*n* = 32]: -4.89 ± 2.41 mm; C [*n* = 57]: -4.74 ± 2.00 mm; E [*n* = 45]: -4.98 ± 2.03 mm; *p* = 0.90; Table 2; Fig.). Standard errors ranged from 0.14 to 0.56 mm across all surgeons performing the same technique. The largest postoperative change in proptosis was reported in orbits undergoing 3-wall + fat orbital decompression, with the mean change ranging from -4.55 to -4.98 mm across surgeons, whereas 1-wall decompression via medial decompression alone resulted in the smallest change in proptosis (mean change, -1.40 to -2.29 mm). No significant difference was observed among surgeons performing the same technique when assessing the overall group or single-orbit analysis (see Table 1, Supplemental Digital Content 1, <https://links.lww.com/IOP/A452>).

Further subgroup analysis was performed to assess whether there was a significant difference in the change in proptosis with

TABLE 1. Baseline demographics of patients who underwent orbital decompression among surgeons

	Surgeon A	Surgeon B	Surgeon C	Surgeon D	Surgeon E	Surgeon F	Surgeon G	p value
Orbits (n)	163	83	89	116	91	37	54	NA
Age (years)	47 ± 13.8	47 ± 11.4	50 ± 14.8	52 ± 14.9	52 ± 15.5	57 ± 11.9	56 ± 14.7	<0.0001
Gender (% M)	n = 41 (25%)	n = 18 (20.7%)	n = 26 (29.2%)	n = 19 (16.4%)	n = 34 (37.4%)	n = 14 (37.8%)	n = 12 (20.7%)	0.01
Smoker (% smoker)	n = 50 (30.5%)	n = 59 (71.1%)	n = 21 (23.6%)	n = 52 (44.8%)	n = 42 (46.2%)	n = 6 (16.2%)	n = 24 (41.4%)	<0.0001
Average CAS	2.0 ± 1.5	2.5 ± 1.2	0.8 ± 1.2	1.1 ± 1.2	3.4 ± 2.6	1.7 ± 2.2	3.7 ± 1.6	<0.0001
Average preoperative exophthalmos (mm)	24.2 ± 3.2	25.5 ± 2.5	24.3 ± 2.5	23.5 ± 2.8	24.2 ± 3.8	22.5 ± 5.8	24.3 ± 3.0	0.0001

All data are expressed as mean ± standard deviation.

CAS, clinical activity score; M, male.

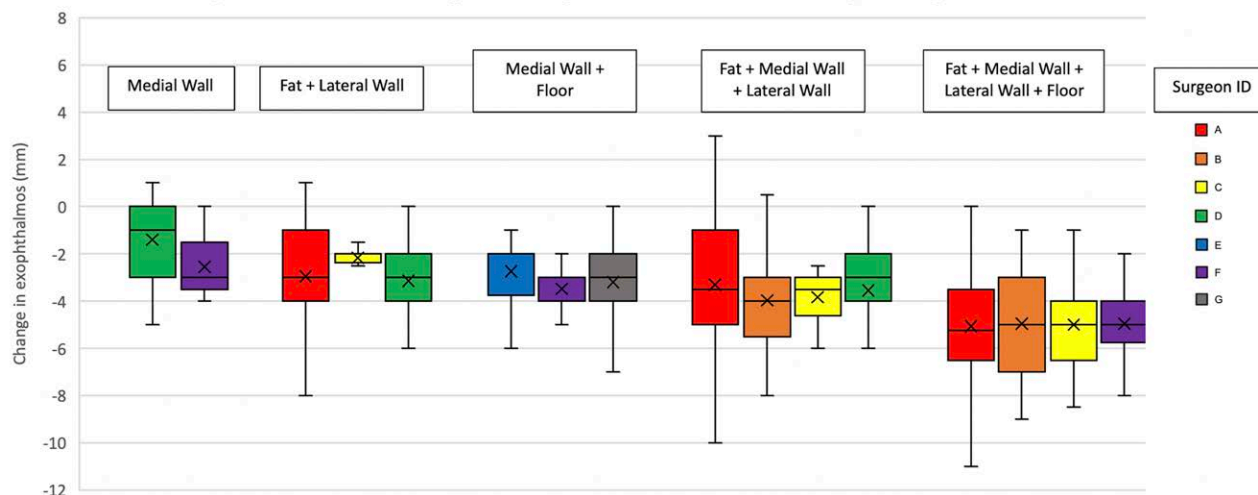
TABLE 2. Comparison of postoperative change in exophthalmos after decompression among surgeons

Decompression technique	n	Surgeon A	Surgeon B	Surgeon C	Surgeon D	Surgeon E	Surgeon F	Surgeon G	p value
Medial wall	29				-1.40 ± 2.06 (n = 15; SE = 0.53)		-2.29 ± 1.73 (n = 14; SE = 0.46)		0.22
Fat + lateral wall	113	-2.94 ± 2.43 (n = 19; SE = 0.56)		-3.57 ± 1.73 (n = 14; SE = 0.46)	-3.14 ± 1.45 (n = 80; SE = 0.16)				0.55
Medial wall + floor	123					-2.61 ± 1.50 (n = 46; SE = 0.22)	-3.67 ± 0.69 (n = 51; SE = 0.14)	-3.20 ± 1.79 (n = 54; SE = 0.24)	0.06
Fat + lateral + medial wall	140	-3.31 ± 2.43 (n = 50; SE = 0.34)	-3.76 ± 2.13 (n = 51; SE = 0.30)	-3.89 ± 1.37 (n = 18; SE = 0.32)	-3.10 ± 1.69 (n = 21; SE = 0.37)				0.45
Fat + lateral wall + medial wall + floor	228	-4.55 ± 4.43 (n = 94; SE = 0.46)	-4.89 ± 2.41 (n = 32; SE = 0.43)	-4.74 ± 2.00 (n = 57; SE = 0.27)		-4.98 ± 2.03 (n = 45; SE = 0.30)			0.90

All data represent postoperative change in exophthalmos (mm) in mean ± standard deviation.

SE, standard error.

Comparison of Post-operative Change in Exophthalmos After Orbital Decompression for Thyroid Eye Disease Among Surgeons



Comparison of postoperative change in proptosis after orbital decompression for thyroid eye disease among surgeons. No significant difference in the change in proptosis was present when comparing surgeons who perform a similar surgical technique. For each group, the box represents the interquartile range (IQR; 25th–75th percentiles), and the horizontal line inside the box marks the median, with “x” denoting the mean. Whiskers extend to the most extreme data points within $1.5 \times \text{IQR}$ from the 25th and 75th percentiles.

endoscopic versus open technique for medial wall decompression when performing the same surgery. When analyzing 1-wall medial wall decompression, no significant difference in outcomes was found between surgeons D and F ($p = 0.22$), both of whom exclusively used endoscopic approaches; no other surgeons performed 1-wall medial wall decompression via either endoscopic or open approach. For patients who underwent medial wall + floor decompression, the average reduction in proptosis was significantly greater in endoscopic cases ($-3.67 \pm 0.96 \text{ mm}$) compared with open cases ($-2.97 \pm 1.64 \text{ mm}$; $p = 0.008$). In contrast, for patients who underwent fat + lateral + medial wall decompression, no statistically significant difference was observed in the change in proptosis between endoscopic cases ($-3.10 \pm 1.69 \text{ mm}$) and open cases ($-3.62 \pm 2.14 \text{ mm}$; $p = 0.168$).

In assessing temporal trends in change in proptosis postoperatively among different surgeons, there were no significant differences observed across all surgeons performing each decompression technique at all 3 different time points. In addition, assessing change in outcomes over time for the same surgeon doing the same procedure at time points 1, 2, and 3 demonstrated no statistically significant differences in outcomes observed with Bonferroni correction (Table 3).

A multivariable logistic regression model was subsequently applied to assess whether any independent preoperative and postoperative characteristics were associated with a change in proptosis. Multivariate modeling showed a statistically significant association between postoperative change in proptosis with preoperative proptosis ($p < 0.001$) but found no significant association between the change in proptosis and age, gender, CAS, smoking history, radiographically measured orbital cross-sectional area, and follow-up duration (Table 4).

DISCUSSION

The approach to orbital decompression for treatment of TED may be tailored to each patient's degree of proptosis and desired reduction.¹¹ Numerous studies have found that surgical

decompression may significantly reduce proptosis in patients with TED and that increasing the number of walls in the decompression may lead to a significantly higher degree of proptosis reduction.^{8,12,13} However, these studies are often retrospective and involve a single institution, posing challenges for interpretation and cross-study comparisons. Such limitations may hinder the reliable application and comparison of published findings, particularly in evaluating result consistency and providing informed patient counseling. Additionally, while multiple existing studies report the efficacy of various decompression techniques, there is a lack of research regarding whether outcomes in proptosis reduction are consistent among surgeons using the same surgical technique. Thus, this study aimed to evaluate intersurgeon variability in proptosis reduction following decompression surgery to assess whether decompression by one surgeon is comparable to decompression by another.

Our study found that the variability in outcomes for proptosis reduction following orbital decompression for TED did not exhibit statistically significant differences among surgeons performing the same surgical technique. While there are minimal data in the existing oculofacial plastic surgery literature focusing on comparison of outcomes among surgeons performing similar surgical techniques, this finding is consistent with the notion that standardized surgical techniques yield comparable results when executed by trained and experienced surgeons. Interestingly, our findings contrast with the European Group on Graves' Orbitopathy (EUGOGO) study, which reported variability in proptosis reduction and complication rates across centers with surgeons (notably including both otolaryngologists and maxillofacial surgeons in addition to specialized orbital-trained ophthalmologists) employing different decompression techniques.¹⁴ Our multi-center analysis found no significant differences in proptosis outcomes among surgeons utilizing the same standardized surgical approach, suggesting that technique standardization and specialized oculofacial and orbital training may mitigate intersurgeon variability. Mean proptosis outcomes were

TABLE 3. Temporal trends in change in postoperative exophthalmos among surgeons

	Surgeon A	Surgeon C	Surgeon D	Surgeon E	Surgeon F	<i>p</i> value*
Medial wall decompression						
2002–2007	NA	NA	NA	NA	NA	NA
2008–2013	NA	NA	NA	NA	–1.66 ± 1.86	NA
2014–2018	NA	NA	–1.40 ± 2.06	NA	NA	NA
<i>p</i> value†	NA	NA	NA	NA	NA	NA
Fat + lateral wall decompression						
2002–2007	NA	NA	NA	NA	NA	NA
2008–2013	–2.75 ± 1.44	–2.21 ± 0.39	–1.88 ± 2.70	NA	NA	0.69
2014–2018	–3.04 ± 2.83	–2.10 ± 0.55	2.99 ± 1.70	NA	NA	0.61
<i>p</i> value†	0.77	0.70	0.20	NA	NA	NA
Medial wall + floor decompression						
2002–2007	NA	NA	NA	–3.21 ± 1.63	NA	NA
2008–2013	NA	NA	NA	–2.52 ± 1.36	NA	NA
2014–2018	NA	NA	NA	–2.40 ± 0.89	–3.00 ± 1.00	0.44
<i>p</i> value†	NA	NA	NA	0.30	NA	NA
Fat + lateral wall + medial wall decompression						
2002–2007	–2.75 ± 1.06	NA	NA	NA	NA	NA
2008–2013	–4.05 ± 1.94	–4.63 ± 0.95	NA	NA	NA	0.47
2014–2018	–3.14 ± 2.58	–3.62 ± 1.06	–3.10 ± 1.69	NA	NA	0.73
<i>p</i> value†	0.55	0.12	NA	NA	NA	NA
Fat + lateral wall + medial wall + floor decompression						
2002–2007	–6.20 ± 4.55	NA	NA	–4.25 ± 1.39	NA	0.79
2008–2013	–5.16 ± 2.16	–4.83 ± 1.95	NA	–5.85 ± 2.28	NA	0.31
2014–2018	–4.58 ± 3.05	–5.11 ± 1.71	NA	–4.19 ± 1.38	NA	0.37
<i>p</i> value†	0.40	0.60	NA	0.02	NA	NA

All data represent postoperative change in exophthalmos (mm) expressed in mean ± standard deviation.

*Statistical comparison of outcomes among surgeons at a single time point.

†Statistical comparison of outcomes for a single surgeon at multiple time points.

TABLE 4. Multivariate model for factors associated with change in exophthalmos

	Effect (95% CI)	<i>p</i> value
Preoperative exophthalmos	1.49 (1.42–1.56)	<0.001
Clinical activity score	0.96 (0.88–1.06)	0.42
Age (year)	1.004 (0.99–1.02)	0.46
Orbital area	1.03 (0.94–1.14)	0.49
Follow-up duration	0.99 (0.99–10.1)	0.51
Gender	1.07 (0.73–1.56)	0.74
Smoking history (% yes)	1.042 (0.74–1.47)	0.82

consistent across surgeons performing the same technique with the small standard errors for each surgeon-technique group (0.14–0.50 mm), suggesting that the minimal intersurgeon differences observed reflect genuinely consistent outcomes for a given technique.

However, it is important to acknowledge that unmeasured confounding variables such as surgeon experience, intraoperative decision-making, and subtle variations in technique may still influence surgical outcomes. Interestingly, there have been data in other surgical subspecialties suggesting that surgeon volume, gender, subspecialization, and patient complexity may contribute to outcomes, typically assessing mortality, readmission, or major medical complications.^{15–19} Given the nature of orbital decompressions as an outpatient procedure, many of these metrics do not apply, and as such, this article focuses on the efficacy of treatment as measured by postoperative change in proptosis compared among different surgeons using similar surgical techniques. In evaluating our data, all surgeons were fellowship-trained oculofacial plastic surgeons, treating patients with approximately equivalent patient complexity (diagnosis of TED requiring surgery), which left surgical volume as a potential factor to assess. However, there was no statistically

significant difference among surgeon outcomes even when comparing surgeons with relatively higher or lower surgical volume, as indicated by the total number of cases a surgeon performed within a given decompression technique during the study period, when controlling for a single surgical technique. Moreover, assessing temporal trends in outcomes serves as a surrogate for measuring the effect of volume, as later years are a surrogate measure for an individual surgeon having more experience and more cumulative volume relative to earlier years. There were no significant differences observed in outcomes over time for the same surgeon doing the same procedure at time points. This suggests that high volume or experience may not necessarily affect patient outcomes of change in proptosis after decompression although, of note, this study only included data if the surgeon had performed at least 10 cases of a single surgical technique. The results of our study strengthen the statistical validity of future multicenter clinical trials assessing outcomes in orbital decompression and facilitate comparison of the outcomes reported by single institutions when assessing outcomes from surgeons employing uniform surgical techniques, thereby advancing understanding of optimal surgical management strategies for TED.

While this article primarily focuses on comparing outcomes among surgeons performing the same decompression technique, there are several other notable findings. The greatest postoperative change in proptosis was reported in patients who underwent 3-wall decompression via fat + lateral wall + medial wall + floor decompression, with mean changes ranging from –4.55 to –4.98 mm across surgeons, and the smallest change in proptosis was noted for 1-wall medial wall decompression (mean change, –1.4 to –2.29 mm). This is consistent with the current literature demonstrating an association between the number of walls involved in decompression and the change in proptosis postoperatively.^{5,8,11–13} In addition, a multivariate model assessing factors influencing proptosis reduction

identified preoperative proptosis as a significant predictor ($p < 0.001$), consistent with prior reports that patients with greater baseline proptosis often experience greater postoperative improvement,^{20,21} potentially reflecting that patients with more severe proptosis may be more likely to undergo more extensive decompressions. Separate multivariate analysis from a similar cohort demonstrated that larger preoperative exophthalmometry values and a greater number of walls decompressed were both independently associated with greater proptosis reduction (unpublished data). Although baseline characteristics such as age, smoking status, CAS, and preoperative proptosis varied significantly among surgeons, CAS, age, orbital area, follow-up duration, gender, and smoking history were found not to be significantly associated with postoperative change in proptosis. Importantly, there was no significant difference in outcomes among surgeons performing the same surgical technique despite these baseline differences in characteristics, which supports the generalizability of our results when standardized surgical techniques are used. Thus, our findings may also increase the ability to compare and apply results from the literature even for studies or groups assessing orbital decompression outcomes for patients with TED with different baseline demographics.

Interestingly, our multivariate analysis found that smoking history was not significantly associated with change in proptosis, despite prior studies suggesting that smoking may negatively impact surgical outcomes after orbital decompression.^{12,20} Smokers have previously been reported to have more severe TED and are more likely to undergo surgical decompression.^{22–26} Furthermore, smokers may be more likely to have worse outcomes after other treatment modalities including radiation, corticosteroids, or teprotumumab.^{23–25,27} Our findings suggest that surgical decompression may remain an effective option for patients with TED who smoke despite counseling, reinforcing the benefit of decompression surgery even in high-risk groups that may have a poor response to other treatment options. In addition, while a larger orbital area/volume may potentially suggest more space for decompression and, therefore, a greater reduction in proptosis, orbital cross-sectional area also was not found to significantly correlate with postoperative change in proptosis. This supports previous findings by Oeverhaus et al.²⁸ that demonstrated that CT morphological analysis of the orbital area does not reliably predict proptosis reduction; as such, preoperative imaging may be helpful, but clinical measurement of proptosis preoperatively may offer the greatest potential predictor of response to surgery. Last, we explored whether an endoscopic or open surgical approach for medial wall decompression resulted in differences in postoperative proptosis reduction among surgeons. There is currently debate regarding whether endoscopic medial wall decompression versus open decompression leads to better outcomes.^{29–31} Although there are few studies comparing open and endoscopic decompression outcomes for TED, one systematic review suggests that the endoscopic medial approach in 2- and 3-wall procedures resulted in significantly lower proptosis reduction compared with transconjunctival or transcaruncular approaches.³¹ In contrast, in our study, patients who underwent endoscopic medial wall + floor decompression had significantly greater proptosis reduction compared with open cases, while fat + lateral + medial wall decompression outcomes revealed no significant difference when comparing endoscopic and open surgical approaches. Endoscopic approaches may provide increased exposure to the medial orbital apex and posterior floor, potentially allowing for more bone removal and, thus, greater proptosis reduction. However, in cases where fat removal and lateral wall decompression are also performed, the overall decompression may be sufficiently robust such that the relative contribution of the medial wall technique becomes less clinically significant.

There are several limitations to our study. First, the multicenter retrospective nature of this study introduces inherent limitations, including that data collection methods such as measurements of proptosis may not be uniform across sites. At the time of data collection, exophthalmometry measurements were performed by the same person, typically the surgeon, using the same exophthalmometer at each visit if possible. In addition, this study attempted to limit bias by assessing the change in proptosis rather than the absolute value of proptosis, given that the measurement technique was most likely standard within a site. Second, all surgeons are associated with major academic institutions with assumed similar resources, which may limit the generalizability of the findings. Third, while this study includes 7 oculofacial plastic surgery fellowship-trained surgeons from a variety of training backgrounds, it is possible that this may not be fully representative of training backgrounds across the United States and, thus, may underestimate variability in outcomes. All included surgeons were also male, which is not representative of all oculofacial plastic surgeons. Fourth, although the study included comparison among surgical techniques performed by at least 2 surgeons on at least 10 occasions each to ensure adequate volume for comparison, minor unstandardized variations may still influence results. Fifth, the effect of the TED phase on decompression outcomes may not have been fully captured with the available data. While CAS was not significantly associated with postoperative change in proptosis on multivariate analysis, this score may not reliably define the true state of disease activity. Sixth, while temporal trends were assessed by stratifying data into 3 time points (2002–2007, 2008–2013, and 2014–2018), this method may not fully reflect individual surgeon career progression. Although stratification by career growth may offer insight into the impact of surgical experience on results, surgeon identity was blinded during analysis to minimize bias; as such, temporal trends were used as a surrogate measure but may not truly reflect the impact of surgeon career growth. Nonetheless, surgeons likely experience a continual process of improvement throughout their entire career. Seventh, post hoc power analysis demonstrated the study was highly powered overall ($\approx 100\%$) to detect differences across all surgeon-technique groups, but power for within-technique comparisons was limited by smaller subgroup sizes; therefore, large intersurgeon differences are unlikely though small differences cannot be excluded. However, most within-technique effect sizes were small (0.16, 0.12, 0.10, and 0.09 for fat + lateral wall, medial wall + floor, fat + lateral + medial wall, and fat + lateral + medial wall + floor, respectively), indicating that between-surgeon differences in mean proptosis reduction were minimal relative to variability in individual patient outcomes. Only the medial wall effect size was moderate (0.48), likely in the setting of the smallest sample size ($n = 29$), which may have inflated variability estimates. Furthermore, most observed differences fell within the known 1- to 2-mm variability of exophthalmometry,³² suggesting that any undetected differences are unlikely to be clinically meaningful. Eighth, postoperative complications (including postoperative diplopia) were not analyzed in this study. Postoperative diplopia outcomes have been previously reported,⁴ and other complication data were not available in the dataset for analysis but may be an avenue for future research. Ninth, surgeon decision-making for the decompression technique may have varied among surgeons, reflecting individual preferences, experience, and institutional standards. Although greater preoperative proptosis typically prompts removal of more walls, some surgeons may use different treatment algorithms or approaches. While we compared outcomes only among

surgeons using the same technique, details regarding individual decision-making algorithms were unavailable, given the retrospective nature of this study and the blinded surgeon identity. Tenth, patient disease characteristics, including orbital fat-predominant, muscle-predominant, or mixed subtypes, were not uniformly available across all participating sites in this dataset. As such, the potential influence of orbital phenotype on proptosis reduction after decompression and volume outside of the orbit could not be assessed in the current study but may represent an area of future research. Last, this study is retrospective in nature with postoperative change in proptosis measured at an average follow-up duration of 21.9 months with a wide range in follow-up duration (3–156 months). It is possible that variability in long-term outcomes beyond 22 months may exist; thus, future studies with longer follow-up may be indicated.

In conclusion, this study demonstrates that there is no significant difference in outcomes for proptosis reduction following orbital decompression for TED among surgeons when performing the same surgical technique. These findings are useful when considering future prospective multicenter trials studying outcomes of orbital decompression. Preoperative proptosis and the number of walls included in the surgical decompression are associated with postoperative change in proptosis, and the timing of surgery within a surgeon's career as approximated by temporal stratification do not appear to significantly affect outcomes. While limitations exist, the results of this study may strengthen the statistical validity of multicenter clinical trials assessing outcomes in orbital decompression and comparison of outcomes among multiple studies in the literature when performed by trained oculofacial plastic and orbital surgeons employing standardized techniques, which may facilitate a better understanding of optimal surgical management strategies for TED.

References

- Shu X, Shao Y, Chen Y, et al. Immune checkpoints: new insights into the pathogenesis of thyroid eye disease. *Front Immunol* 2024;15:1–13.
- Epstein FH, Bahn RS, Heufelder AE. Pathogenesis of Graves' ophthalmopathy. *N Engl J Med* 1993;329:1468–1475.
- Mourits MP, Prummel MF, Wiersinga WM, et al. Clinical activity score as a guide in the management of patients with Graves' ophthalmopathy. *Clin Endocrinol (Oxf)* 1997;47:9–14.
- Nair AA, Ediriwickrema LS, Dolman PJ, et al. Predictive modeling of new-onset postoperative diplopia following orbital decompression for thyroid eye disease. *Ophthalmic Plast Reconstr Surg* 2022;38:551–557.
- Kikkawa DO, Pornpanich K, Cruz RC, et al. Graded orbital decompression based on severity of proptosis. *Ophthalmology* 2002;109:1219–1224.
- Braun T, Bhadkamkar M, Jubbal K, et al. Orbital decompression for thyroid eye disease. *Semin Plast Surg* 2017;31:040–045.
- Rootman DB. Orbital decompression for thyroid eye disease. *Surv Ophthalmol* 2018;63:86–104.
- Ediriwickrema LS, Korn BS, Kikkawa DO. Orbital decompression for thyroid-related orbitopathy during the quiescent phase. *Ophthalmic Plast Reconstr Surg* 2018;34(4S):S90–S97.
- Perry JD, Kadakia A, Foster JA. Transcaruncular orbital decompression for dysthyroid optic neuropathy. *Ophthalmic Plast Reconstr Surg*. 2003;19:353–358.
- Juniat V, Abbeel L, McGilligan JA, et al. Endoscopic orbital decompression by oculoplastic surgeons for proptosis in thyroid eye disease. *Ophthalmic Plast Reconstr Surg*. 2019;35:590–593.
- Tooley AA, Godfrey KJ, Kazim M. Evolution of thyroid eye disease decompression—dysthyroid optic neuropathy. *Eye (Lond)* 2019;33:206–211.
- Cheng AMS, Wei YH, Liao SL. Strategies in surgical decompression for thyroid eye disease. *Oxid Med Cell Longev* 2020;2020:3537675.
- Choe CH, Cho RI, Elner VM. Comparison of lateral and medial orbital decompression for the treatment of compressive optic neuropathy in thyroid eye disease. *Ophthalmic Plast Reconstr Surg*. 2011;27:4–11.
- Mourits MP, Bijl H, Altea MA, et al; European Group on Graves' Orbitopathy (EUGOGO). Outcome of orbital decompression for disfiguring proptosis in patients with Graves' orbitopathy using various surgical procedures. *Br J Ophthalmol* 2009;93:1518–1523.
- Wallis CJ, Ravi B, Coburn N, et al. Comparison of postoperative outcomes among patients treated by male and female surgeons: a population based matched cohort study. *BMJ* 2017;359:j4366.
- Wallis CJD, Jerath A, Aminoltejeri K, et al. Surgeon sex and long-term postoperative outcomes among patients undergoing common surgeries. *JAMA Surg* 2023;158:1185–1194.
- Birbeck KF, Macklin CP, Tiffin NJ, et al. Rates of circumferential resection margin involvement vary between surgeons and predict outcomes in rectal cancer surgery. *Ann Surg* 2002;235:449–457.
- Campbell RJ, El-Deffrawy SR, Gill SS, et al. Surgical outcomes among focused versus diversified cataract surgeons. *Ophthalmology* 2021;128:827–834.
- Kelz RR, Sellers MM, Niknam BA, et al. A national comparison of operative outcomes of new and experienced surgeons. *Ann Surg* 2021;273:280–288.
- Wu CY, Niziol LM, Musch DC, et al.^{5,6,12} *Ophthalmic Plast Reconstr Surg*. 2017;33:189–195.
- Peraza LR, Fearington FW, Hernandez-Herrera GA, et al. Orbital decompression for thyroid eye disease: outcomes by preoperative severity and technique. *Am J Rhinol Allergy* 2025;39:284–292.
- O'Dell J, Mussatto C, Chu R, et al. Effects of smoking on outcomes of thyroid eye disease treated with teprotumumab: a retrospective cohort study. *Kans J Med* 2023;16:62–64.
- Bartalena L, Marcocci C, Tanda ML, et al. Cigarette smoking and treatment outcomes in Graves ophthalmopathy. *Ann Intern Med* 1998;129:632–635.
- Eckstein A, Quadbeck B, Mueller G, et al. Impact of smoking on the response to treatment of thyroid associated ophthalmopathy. *Br J Ophthalmol* 2003;87:773–776.
- Xing L, Ye L, Zhu W, et al. Smoking was associated with poor response to intravenous steroids therapy in Graves' ophthalmopathy. *Br J Ophthalmol* 2015;99:1686–1691.
- Prummel MF, Wiersinga WM. Smoking and risk of Graves' disease. *JAMA* 1993;269:479–482.
- Oke I, Reshef ER, Elze T, et al; IRIS Registry Analytic Center Consortium. Smoking is associated with a higher risk of surgical intervention for thyroid eye disease in the IRIS registry. *Am J Ophthalmol* 2023;249:174–182.
- Oeverhaus M, Copei A, Mattheis S, et al. Influence of orbital morphology on proptosis reduction and ocular motility after decompression surgery in patients with Graves' orbitopathy. *PLoS One* 2019;14:e0218701.
- Gulati S, Ueland HO, Haugen OH, et al. Long-term follow-up of patients with thyroid eye disease treated with endoscopic orbital decompression. *Acta Ophthalmol* 2015;93:178–183.
- Silver RD, Harrison AR, Goding GS. Combined endoscopic medial and external lateral orbital decompression for progressive thyroid eye disease. *Otolaryngol Head Neck Surg* 2006;134:260–266.
- Boboridis KG, Uddin J, Mikropoulos DG, et al. Critical appraisal on orbital decompression for thyroid eye disease: a systematic review and literature search. *Adv Ther* 2015;32:595–611.
- Mourits MP, Lombardo SH, van der Sluijs FA, et al. Reliability of exophthalmos measurement and the exophthalmometry value distribution in a healthy Dutch population and in Graves' patients. An exploratory study. *Orbit* 2004;23:161–168.