Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

GLP-1 Use Does Not Explain Bariatric Surgery Volume Decline: Procedure Mix and Patient Case Mix in a Statewide Quality Improvement Collaborative

Akram Irshaid
Michigan State University
Matthew Bucala
Michigan Medicine
Amanda Stricklen
Michigan Medicine
Oliver A. Varban
Henry Ford Health
Jonathan F. Finks
Michigan Medicine
Nabeel Obeid
Michigan Medicine
Michael Kia
McLaren Regional Medical Center
Rachel Ross
Michigan Medicine
Erin Charbeneau
Michigan Medicine
Hope T. Jackson
Michigan Medicine
Amir Ghaferi
University of Chicago
Arthur M. Carlin
acarlin1@hfhs.org

Henry Ford Health

Research Article

Keywords:

Posted Date: November 26th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11446075/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

Introduction:

Bariatric surgery volumes have declined across the United States amid rapid adoption of GLP-1 receptor agonist (GLP-1 RA) medications for obesity management. Whether patient-level or program-level factors predict which sites lose volume remains poorly characterized. This study examined predictors of site-level volume decline in a statewide quality improvement collaborative.

Methods:

Retrospective cohort study of 37 bariatric programs in the Michigan Bariatric Surgery Collaborative from January 2022 through December 2024, including 22,156 adults undergoing primary bariatric surgery at sites with mean annual volume ≥ 50 cases. Volume loss was defined as > 10% decline in annual primary procedures (2022 to 2024). Logistic regression identified independent predictors of site-level volume loss, adjusting for site structure, complication rate, GLP-1 RA use stratified by diabetes status, patient demographics, and comorbidity case mix.

Results:

Twenty-eight programs (76%) experienced volume loss. Higher RYGB procedure share was independently associated with greater odds of volume loss (Model 5 OR per 1%, 1.11; 95% CI, 1.02–1.35; P=0.024), strengthening with case-mix adjustment. Therapeutic GLP-1 use among diabetic patients was associated with reduced odds of volume loss (Model 3 OR per 1%, 0.61; 95% CI, 0.34–0.95; P=0.041); weight-loss GLP-1 use among non-diabetic patients showed a similar nonsignificant trend. MBSAQIP accreditation, mean annual volume, complication rate, teaching status, and robotic approach rate were not independently associated with volume loss.

Conclusions:

Approximately three-quarters of bariatric programs in Michigan experienced significant volume loss between 2022 and 2024, coinciding with rapid GLP-1 RA adoption. However, site-level GLP-1 RA use was not associated with greater volume loss; programs that incorporated GLP-1 prescribing into their practice were less likely to experience volume loss than those that did not. This suggests that multidisciplinary programs offering both medical and surgical obesity management may be partially buffered against volume decline, likely through internal referral pathways that retain patients who do not achieve adequate response to pharmacotherapy alone.

INTRODUCTION

Bariatric surgery volumes have declined substantially across the United States in recent years [1, 2]. The decline has occurred during the same period as a dramatic increase in glucagon-like peptide-1 receptor agonist (GLP-1 RA) prescribing for both diabetes management and weight loss [3–5], a notable pattern

given that adult obesity prevalence in the United States exceeds 42%, suggesting that pharmacotherapy is diverting patients who would otherwise be surgical candidates [6].

Several explanations for declining surgical volumes have been proposed: GLP-1 RA medications diverting potential surgical candidates, reduced referrals from primary care, persistent post-pandemic effects on elective surgery utilization, payer coverage changes, and shifting patient preferences. Distinguishing among these explanations matters because each implies different policy responses. If the decline reflects appropriate medical management of mild-to-moderate obesity, expanded GLP-1 access may be welcome. If instead the decline reflects barriers to definitive surgical treatment for severe obesity, the implications are concerning given that bariatric surgery provides superior and more durable metabolic disease control compared with pharmacotherapy alone, including in head-to-head randomized comparisons [7–11].

The Michigan Bariatric Surgery Collaborative (MBSC) is a statewide quality improvement collaborative that prospectively collects standardized data on bariatric surgery patients across Michigan [12]. The collaborative provides an opportunity to examine site-level predictors of volume change in a population not subject to selection by single-payer or single-system characteristics. We used the MBSC registry to test whether site-level GLP-1 RA exposure, procedure mix, structural characteristics, and patient case mix predict which sites experience volume loss. We hypothesized that if patient-level GLP-1 RA use was driving program-level decline, sites with higher GLP-1 RA exposure would show greater volume loss. We further hypothesized that sites differed in resilience by structural and operational characteristics independent of patient mix.

METHODS

Study Design and Cohort

This retrospective cohort study analyzed prospectively collected data from the Michigan Bariatric Surgery Collaborative (MBSC), a statewide quality improvement collaborative comprising 41 bariatric surgery centers across Michigan [12]. The MBSC collects standardized preoperative, perioperative, and postoperative data on all patients undergoing bariatric surgery at participating centers.

We included all adult patients (≥ 18 years) who underwent primary bariatric surgery (Roux-en-Y gastric bypass [RYGB], sleeve gastrectomy, biliopancreatic diversion with duodenal switch, or single-anastomosis duodeno-ileal bypass with sleeve gastrectomy) between January 1, 2022, and December 31, 2024. Revisional procedures were excluded from primary analyses. To ensure stable site-level estimates of volume change, sites were included in the volume-decline analysis if they performed a mean of at least 50 primary procedures annually over the study period (37 of 41 sites; n = 22,156 patient procedures). The MBSC participation minimum is 25 cases annually; we applied the more conservative ≥ 50 threshold because year-over-year change at very low volumes is dominated by individual surgeon retirements, single-quarter scheduling shifts, and case-count fluctuations of only a handful of

procedures, which can produce nominal volume-change percentages that do not reflect underlying program trajectory. A sensitivity analysis using the MBSC participation ≥ 25 threshold confirmed the robustness of all key findings (Supplement Table S8).

Outcomes

The primary outcome was site-level volume loss, defined as a decline in primary bariatric procedure volume of more than 10% from 2022 to 2024. Sites were dichotomized as volume-loss (> 10% decline) versus volume-stable (gain of > 10% or change within ± 10%, combined as "not loss"). The 10% threshold was selected a priori to identify clinically meaningful changes affecting program operations.

Secondary outcomes included 30-day surgical complications (anastomotic leak, hemorrhage, obstruction, surgical site infection, venous thromboembolism) using standardized MBSC definitions.

Predictors

Site-level predictors were aggregated from patient-level data and included MBSAQIP accreditation status [13], teaching hospital designation, mean annual surgical volume (2022–2024), percentage of robotic procedures, and percentage of Roux-en-Y gastric bypass procedures. Patient case-mix variables included mean age, percent female, mean body mass index, complication rate, baseline GLP-1 RA use stratified by diabetes status (therapeutic if patients had diabetes; weight-loss if non-diabetic), and three composite comorbidity indices: metabolic burden (≥ 1 of diabetes mellitus, hypertension, hyperlipidemia, gastroesophageal reflux disease), cardiopulmonary burden (≥ 1 of coronary artery disease, congestive heart failure, cerebrovascular disease, peripheral vascular disease, angina, rhythm disorder, sleep apnea, asthma, COPD, lung disease, or supplemental oxygen requirement), and mental health burden (≥ 1 of depression, anxiety, bipolar disorder, eating disorder, or substance use disorder). The composite indices were used in primary analyses; sensitivity analyses replaced them with the corresponding individual flags. Insurance payer mix (percent Medicare, percent Medicaid) was examined separately as patient-level case-mix and was not associated with site-level volume loss in univariate testing (Medicaid OR per 1%, 1.02; 95% CI, 0.97–1.08; P = 0.452; Medicare OR per 1%, 0.98; 95% CI, 0.90–1.07; P = 0.673), so it was not retained in the multivariable models to preserve degrees of freedom for case-mix variables with stronger univariate signal.

Baseline GLP-1 RA use was ascertained from the MBSC preoperative questionnaire item identifying non-insulin injectable medication use. By 2024, the predominant non-insulin injectable medications in use among bariatric candidates were GLP-1 receptor agonists (semaglutide, liraglutide, tirzepatide, dulaglutide) [3, 4]. Patients with diabetes mellitus on this medication class were classified as therapeutic GLP-1 use; patients without diabetes were classified as weight-loss GLP-1 use. Standard perioperative care for bariatric surgery candidates was guided by the joint AACE/TOS/ASMBS/OMA/ASA clinical practice guidelines [14].

Statistical Analysis

Continuous variables are presented as means (standard deviation) or medians (interquartile range) for skewed distributions. Categorical variables are presented as frequencies and percentages. Site characteristics were compared across volume categories using Wilcoxon rank-sum or Fisher's exact tests as appropriate.

Predictors of site-level volume loss were examined using a sequential multivariable regression strategy. Five nested models progressively adjusted for site structural characteristics (Model 1: MBSAQIP accreditation, teaching hospital, mean annual volume, percent robotic, percent RYGB), site quality (Model 2: + complication rate), GLP-1 RA exposure (Model 3: + therapeutic and weight-loss GLP-1 use), patient demographics (Model 4: + age, sex, BMI), and comorbidity case mix (Model 5: + metabolic, cardiopulmonary, and mental health composite indices). Given the sparse outcome (n = 9 not-loss events) and quasi-complete separation arising in adjusted models [15], all regression models used Firth penalized maximum likelihood estimation [16], which provides bias-reduced point estimates and stable confidence intervals in small samples. Profile likelihood confidence intervals were used. Coefficient estimates with 95% confidence intervals spanning more than 50-fold (ratio of upper to lower bound > 50) were flagged as unstable and not interpreted as reliable measures of effect [17].

Sensitivity analyses examined three specifications: (1) replacing the continuous percent RYGB predictor with a binary RYGB-dominant flag (≥ 50% RYGB) consistent with prior characterizations of procedure-dominant cohorts; (2) replacing comorbidity composite indices with individual comorbidity flags (hyperlipidemia, diabetes, sleep apnea, depression); and (3) testing the parallel resilience model with the outcome flipped to volume gain or maintenance. Model fit was compared using likelihood ratio tests on the Firth penalized log-likelihood.

Site-level analyses are inherently ecological, and we interpret findings consistent with American Statistical Association recommendations [18] emphasizing point estimates and confidence intervals over significance dichotomization. Two-sided P values less than 0.05 were considered statistically significant; adjustments for multiple comparisons were not applied given the hypothesis-generating nature of the analyses. All analyses were performed using R version 4.3.0 with the logistf package for Firth penalized regression.

RESULTS

Cohort and Volume Trends

Between January 1, 2022, and December 31, 2024, a total of 27,732 bariatric procedures were performed at 41 MBSC-participating centers, of which 25,158 (90.7%) were primary procedures. Among 37 sites meeting the ≥ 50 mean annual volume threshold, 22,156 primary procedures were included in the analytic cohort. Of these, 16,929 (76.4%) were sleeve gastrectomies, 4,839 (21.8%) were Roux-en-Y

gastric bypass, and 388 (1.8%) were other procedures. Median patient age was 43 years; 18,326 (82.7%) were female; and 6,053 (27.3%) had diabetes mellitus. Median annual surgical volume per site was 184 cases (interquartile range 109–249). Total annual primary procedure volume across the collaborative declined from 9,744 procedures in 2022 to 6,275 procedures in 2024, a 35.6% decrease (Fig. 1). Twenty-eight of 37 sites (76%) experienced volume loss, while 9 sites (24%) maintained or grew volume during this period (Table 1).

Table 1
Site Characteristics by Volume Category (n = 37 sites with mean annual volume ≥ 50)

Characteristic	Overall (N = 37)	Volume Loss (N = 28)	Not Loss (N = 9)	P-Value
Median annual volume, cases (IQR)	184 (109–249)	189 (115–246)	162 (97–233)	0.51
MBSAQIP accredited, No. (%)	32 (86%)	25 (89%)	5 (56%)	0.030
Teaching hospital, No. (%)	6 (16%)	4 (14%)	2 (22%)	0.62
Roux-en-Y gastric bypass share, % (median, IQR)	21 (10–27)	23 (15–28)	9 (6–16)	< 0.001
Robotic approach, % (median, IQR)	39 (16–58)	39 (15–57)	44 (22–62)	0.51
Complication rate, % (median, IQR)	2.8 (2.0–3.7)	2.9 (2.1–3.8)	2.5 (1.8–3.4)	0.43
Baseline GLP-1 RA use overall, % (median, IQR)	12.5 (8.4–16.2)	12.0 (8.1–15.4)	14.8 (11.0–18.5)	0.038
Therapeutic GLP-1 use (with diabetes), %	9.6 (6.7–12.4)	9.3 (6.5–11.8)	11.7 (8.8–14.2)	0.044
Weight-loss GLP-1 use (without diabetes), %	2.3 (1.6–3.4)	2.2 (1.6–3.2)	2.9 (1.9–4.5)	0.18

Notes: Continuous variables presented as median (interquartile range). P values from Wilcoxon rank-sum tests for continuous variables and Fisher's exact tests for categorical variables. Volume loss defined as >10% decrease in annual primary procedure volume from 2022 to 2024.

Baseline GLP-1 RA Use

Baseline GLP-1 RA use among bariatric surgical candidates increased substantially over the study period. Among patients with completed preoperative questionnaires, baseline GLP-1 RA use rose from 12.8% in 2022 to 17.5% in 2024, with continuing increases observed in 2025. By 2024, approximately three-quarters of GLP-1 RA users had diabetes mellitus, indicating predominant therapeutic rather than weight-loss use. Site-level GLP-1 RA exposure varied widely; 8 sites had ≥ 5% baseline GLP-1 RA use by the start of 2023, while 5 sites did not reach this threshold until 2024 or later.

Predictors of Volume Loss

Univariate site characteristics by volume category are shown in Table 1. Sites that experienced volume loss had a substantially higher percentage of Roux-en-Y gastric bypass procedures than sites that maintained or gained volume (median 22.8% vs 6.6%), a 16 percentage-point absolute difference; however, this comparison did not reach statistical significance in non-parametric univariate testing (P = 0.124), and the independent procedure-mix effect emerged only after multivariable adjustment for case mix. MBSAQIP accreditation was numerically more common among loss sites than non-loss sites (89% [25/28] vs 67% [6/9]) but did not differ significantly (P = 0.140), and the association attenuated further in multivariable models. Teaching hospital status, mean annual volume, robotic approach rate, and complication rate did not differ significantly by volume category.

In multivariable Firth penalized logistic regression (Table 2), higher percent RYGB was independently associated with volume loss across all five sequential models, and the magnitude of the association strengthened with adjustment for patient case mix. In the structural-only Model 1, each 1 percentage-point increase in site RYGB share was associated with 3% higher odds of volume loss (OR 1.03; 95% CI, 0.98–1.11), reaching statistical significance in Model 4 after adjustment for demographics (OR 1.09; 95% CI, 1.02–1.26; P = 0.020) and Model 5 after additional adjustment for comorbidity case mix (OR 1.11; 95% CI, 1.02–1.35; P = 0.024).

Table 2
Multivariable Predictors of Site-Level Volume Loss

Variable	Model 1 (Structure)	Model 2 (+ Quality)	Model 5 (Full Adjustment)
Mean annual volume (per case)	1.00 (0.99–1.01)	1.00 (0.99–1.01)	0.99 (0.98–1.00)
MBSAQIP accreditation	2.17 (0.33–14.28)	2.15 (0.33–14.40)	13.17 (unstable)
Teaching hospital	0.32 (0.05–1.93)	0.35 (0.06–2.06)	1.15 (unstable)
% Robotic (per 1%)	0.99 (0.96–1.02)	0.99 (0.97–1.02)	1.02 (0.98–1.14)
% RYGB (per 1%)	1.03 (0.98–1.11)	1.03 (0.98–1.10)	**1.11 (1.02–1.35)***
Complication rate (per 1%)	—	0.97 (0.54–2.10)	0.48 (unstable)
Therapeutic GLP-1 use (per 1%)	—	—	0.53 (unstable)
Weight-loss GLP-1 use (per 1%)	—	—	0.57 (unstable)
Mean age (per year)	—	—	0.38 (unstable)
% Female	—	—	0.70 (0.19–1.10)
Mean BMI	—	—	0.44 (unstable)
Metabolic comorbidity index, %	—	—	1.21 (unstable)
Cardiopulmonary index, %	—	—	1.10 (unstable)
Mental health index, %	—	—	0.82 (unstable)
Sequential Firth penalized logistic regression models. OR (95% CI). Significant estimates (P < 0.05) shown in bold.			

Notes: All models use Firth penalized maximum likelihood logistic regression with site as the unit of analysis (n=37 sites; 28 loss / 9 not loss). Model 3 (+GLP-1 stratified by diabetes) and Model 4 (+demographics) shown in supplement Table S1. "Unstable" indicates 95% CI spans more than 50-fold, reflecting insufficient information to estimate the predictor reliably given the sample size. *P=0.024 in Model 5; therapeutic GLP-1 use significant in Model 3 (OR 0.61; 95% CI, 0.34–0.95; P=0.041; see supplement Table S1). Model 4 vs Model 3 likelihood ratio test χ^2=8.03, df=3, P=0.045; Model 5 vs Model 4 χ^2=16.81, df=3, P=0.001.

Therapeutic GLP-1 RA use among diabetic patients was associated with reduced odds of volume loss in Model 3 (OR per 1% diabetic GLP-1 RA users, 0.61; 95% CI, 0.34–0.95; P = 0.041), with directionally similar protective estimates in Models 4 and 5 that did not reach statistical significance after adjustment

for full case mix. Weight-loss GLP-1 use among non-diabetic patients showed a similar protective direction (Model 3 OR 0.76; 95% CI, 0.26–1.89) without reaching significance in any model.

MBSAQIP accreditation showed point estimates ranging from OR 2.17 in the structural-only model to higher but unstable estimates in fully adjusted models, with all confidence intervals crossing the null. Mean annual volume, complication rate, teaching hospital status, and robotic approach rate were not independently associated with volume loss in any model. The addition of patient case mix in Models 4 and 5 significantly improved model fit (Model 4 vs Model 3: χ^2=8.03, df = 3, P = 0.045; Model 5 vs Model 4: χ^2=16.81, df = 3, P = 0.001).

Sensitivity Analyses

Three sensitivity analyses confirmed the procedure-mix finding while reframing it methodologically. First, replacing continuous percent RYGB with a binary RYGB-dominant flag (≥ 50% RYGB) yielded directionally consistent estimates but unstable confidence intervals in all models given that only 3 of 37 sites met this threshold; the continuous specification was the appropriate parameterization. Second, replacing comorbidity composite indices with individual flags (hyperlipidemia, diabetes, sleep apnea, depression) yielded directionally similar estimates, although the strain on degrees of freedom widened confidence intervals. Third, the parallel resilience analysis, which flipped the outcome to volume gain or maintenance, yielded mirror-inverted estimates with internal consistency: percent RYGB OR 0.98 per 1% (favoring resilience among sleeve-heavy sites), therapeutic GLP-1 OR 1.44 (favoring resilience among sites with greater diabetic GLP-1 use), and MBSAQIP OR 0.36 (favoring resilience among non-accredited sites). The full sensitivity tables and resilience model results are shown in the supplement.

Univariate Comparison of Resilient and Loss Sites

A univariate descriptive comparison of resilient (Gain or Maintain) versus loss sites (Supplement Table S1) identified two characteristics that significantly differentiated the groups. Therapeutic GLP-1 RA use among diabetic patients was higher at resilient sites (median 6.8% vs 4.8%; P = 0.042), reinforcing the multivariable finding that GLP-1 RA exposure is not associated with volume decline and is directionally protective. Patient depression prevalence was also higher at resilient sites (median 56.0% vs 48.1%; P = 0.045), a pattern consistent with more comprehensive mental health screening at programs that maintained throughput, possibly reflecting integration of multidisciplinary obesity care. Although percent RYGB was not significantly different in non-parametric univariate testing (median 6.6% at resilient sites vs 22.8% at loss sites; P = 0.124), the absolute difference of 16 percentage points is large and the multivariable analysis confirmed an independent procedure-mix effect after adjustment for case mix. MBSAQIP accreditation, teaching hospital status, mean annual volume, robotic approach rate, complication rate, demographics, and other comorbidity indicators did not differ significantly between the two groups.

Patient Outcomes

Despite substantial volume reductions, surgical safety was maintained. Patient-level 30-day complication rates were uniformly low across volume categories (3.0% at loss sites and 2.8% at not-loss sites; P = 0.62); site-level median complication rates are reported in Table 1. No volume category showed statistically significant differences in anastomotic leak, hemorrhage, obstruction, surgical site infection, or venous thromboembolism. Adjusted patient-level analyses confirmed that comorbidities (diabetes, hypertension, sleep apnea), but not site volume category, predicted complications.

DISCUSSION

In this statewide cohort study, approximately three-quarters of bariatric programs experienced significant volume loss between 2022 and 2024. Programs with higher rates of GLP-1 RA use among their patient population were less likely to experience volume loss, with therapeutic use among diabetic patients showing a statistically significant protective association and weight-loss use among non-diabetic patients trending in the same direction. Higher RYGB procedure share was associated with greater odds of volume loss after adjustment, though this likely reflects differences in program orientation rather than a direct effect of procedure selection. Taken together, these findings suggest that bariatric programs embracing GLP-1 prescribing as part of a multidisciplinary obesity treatment pathway may be partially buffered against the volume declines affecting programs that offer surgery alone.

The procedure-mix finding has both clinical and operational implications. Programs that maintained or grew volume during 2022–2024 performed predominantly sleeve gastrectomy. Several mechanisms could account for this pattern. Sleeve gastrectomy may be perceived by referring providers and patients as a less invasive option more compatible with concurrent GLP-1 therapy, with shorter hospitalization and recovery, and randomized comparisons of sleeve gastrectomy versus RYGB show comparable weight loss at 5 years for many patients [19, 20]. Sleeve-dominant programs may have built referral networks and patient education materials around the simpler procedure that maintained throughput in a contracting referral environment. Alternatively, sites performing more RYGB may serve more medically complex patient populations with higher GERD or diabetic comorbidity for whom RYGB is preferred [21, 22]' and whose providers are more likely to defer surgical evaluation in favor of medical management once GLP-1 therapy becomes available. The cross-sectional ecological nature of this analysis cannot distinguish among these mechanisms.

The protective directional association between baseline GLP-1 RA use and volume preservation is consistent across the primary stepwise analysis, the resilience sensitivity analysis, and the univariate descriptive comparison. Sites that captured patients on GLP-1 therapy were not the sites losing volume; in fact, resilient sites had significantly higher therapeutic GLP-1 use among diabetic patients (median 6.8% vs 4.8%; P = 0.042) than loss sites. This finding is incompatible with a simple substitution model in which GLP-1 therapy diverts surgical candidates one-for-one. A more plausible interpretation is that GLP-1 prescribing prevalence among surgical candidates is a marker of integrated obesity care: programs that effectively communicate with primary care, manage medication adjuncts perioperatively, and operate within multidisciplinary obesity care frameworks may be the same programs maintaining

surgical throughput. The univariate finding that resilient sites also had higher patient depression prevalence (median 56.0% vs 48.1%; P = 0.045) is consistent with this interpretation, as comprehensive mental health screening and management is a hallmark of integrated obesity care programs [23]. Patient-level GLP-1 use, in this framing, is a surrogate for organizational engagement with comprehensive obesity care, not a competing therapy [14].

Population-level rather than patient-level GLP-1 effects may still drive aggregate volume decline. The MBSC registry only captures patients who reach surgical evaluation, and therefore cannot observe the larger pool of potential candidates whose primary care providers prescribe GLP-1 RA medications and never refer them onward [2, 5]. Pharmacy claims data and primary care registries will be needed to test whether community-level GLP-1 prescribing volume correlates with bariatric surgery referral declines, particularly in settings where insurance coverage of GLP-1 therapy is generous and bariatric coverage is restrictive.

MBSAQIP accreditation showed a directionally consistent association with volume loss across models that did not reach statistical significance after adjustment. The structural-only Model 1 estimate of OR 2.17 attenuated and stabilized substantially compared to a prior estimate from the same registry using a less restrictive cohort definition. The current finding does not support an MBSAQIP-specific paradox; rather, the apparent association reflects the correlation between accreditation, urban or suburban location, and procedure mix, with the procedure-mix variable absorbing most of the structural signal once introduced as a covariate.

Surgical safety was maintained across all volume categories and adoption cohorts. The uniform low complication rates (~ 3%) provide reassurance that contemporary bariatric surgery quality has not degraded despite substantial volume reductions, consistent with prior evidence of a safety plateau in mature programs [24]. However, sustained volume reductions over multi-year horizons raise concerns about training pipelines, surgeon experience maintenance, and program viability that this cross-sectional snapshot cannot address.

This study has several limitations. First, the ecological nature of site-level analyses limits causal inference; site-level associations cannot be interpreted as patient-level effects. Second, the small number of not-loss sites (n = 9) constrains statistical power and produces wide confidence intervals on several adjusted estimates. We addressed this with Firth penalized estimation and explicit flagging of unstable confidence intervals, but the fundamental constraint remains. Third, our analysis includes only 2022–2024 in the primary cohort because the MBSC preoperative GLP-1 questionnaire item achieved meaningful coverage only in 2022 and beyond; longer panels would require alternative data sources for early-period exposure ascertainment. Fourth, we cannot observe the upstream referral pipeline; community-level prescribing data and primary care referral pattern data are needed to test population-level mechanisms. Fifth, our analytic window ends in 2024 to ensure complete abstraction of follow-up data; observed continued declines through 2025 (presented descriptively in the supplement) suggest the trends documented here have continued. Sixth, although the procedure-mix association was robust to a

sensitivity analysis using the MBSC participation ≥ 25 threshold (Table S8), the binary RYGB-dominant specification yielded unstable estimates given that only 3 of 37 sites met the ≥ 50% RYGB threshold; the continuous parameterization is the appropriate inference, and the finding warrants replication in larger cohorts.

Future research should examine longer-term volume trajectories beyond 2024 once full abstraction is available, conduct comparative effectiveness studies of GLP-1 monotherapy versus surgery versus combined sequenced approaches, characterize referral pattern changes through linked primary care data, and assess health equity implications of differential surgical access across patient populations [25, 26].

CONCLUSION

In this statewide cohort of 37 bariatric surgery programs from 2022 through 2024, programs performing a higher share of Roux-en-Y gastric bypass experienced greater volume decline. Patient-level GLP-1 RA use was not associated with site-level volume loss and was directionally protective when therapeutic. These findings reframe the bariatric volume decline as a procedure-mix phenomenon affecting RYGB-heavier programs rather than evidence of patient-level substitution by GLP-1 RA medications. Sustained low complication rates indicate surgical quality has been preserved despite volume reductions. Population-level mechanisms operating upstream of the surgical referral pipeline likely account for aggregate volume declines and warrant investigation in linked claims and primary care data.

DECLARATIONS

Author Contribution

Akram Irshaid: Conceptualization, methodology, investigation, writing (original draft), writing (review and editing), supervision, project administration.Matthew Bucala: Methodology, formal analysis, data curation, software, visualization, writing (original draft), writing (review and editing).Amanda Stricklen: Project administration, investigation, data curation, writing (review and editing).Oliver Varban: Writing (review and editing), supervision.Jonathan Finks: Writing (review and editing), supervision.Nabeel Obeid: Writing (review and editing), supervision.Mina Kia: Investigation, writing (review and editing).Rachel Ross: Project administration, data curation, writing (review and editing).Erin Charbeneau: Data curation, investigation, writing (review and editing).Hope Jackson: Writing (review and editing), supervision.Amir Ghaferi: Writing (review and editing), supervision.Arthur Carlin: Conceptualization, methodology, investigation, writing (original draft), writing (review and editing), supervision, project administration.

REFERENCES

1. English WJ, Spaniolas K, Brethauer SA, et al. American Society for Metabolic and Bariatric Surgery 2022 estimate of metabolic and bariatric procedures performed in the United States. Surg Obes

Relat Dis. 2024;20(5):425–31.
2. Westbrook JD, Dimick JB, Sheetz KH. Trends in inpatient bariatric surgery procedures in the United States, 2012–2022. JAMA. 2024;331(13):1148–50.
3. Wilding JPH, Batterham RL, Calanna S, et al. Once-weekly semaglutide in adults with overweight or obesity. N Engl J Med. 2021;384(11):989–1002.
4. Jastreboff AM, Aronne LJ, Ahmad NN, et al. Tirzepatide once weekly for the treatment of obesity. N Engl J Med. 2022;387(3):205–16.
5. Watanabe JH, Kwon J, Nguyen B, Yount RA, Banta JE. Trends in glucagon-like peptide-1 receptor agonist use, 2014 to 2022. J Am Pharmacists Association. 2024;64(1):133–e1382.
6. Hales CM, Carroll MD, Fryar CD, Ogden CL. Prevalence of obesity and severe obesity among adults: United States, 2017–2018. Natl Cent Health Stat Data Brief. 2020;(360):1–8.
7. Sjöström L, Peltonen M, Jacobson P, et al. Association of bariatric surgery with long-term remission of type 2 diabetes and with microvascular and macrovascular complications. JAMA. 2014;311(22):2297–304.
8. Arterburn DE, Olsen MK, Smith VA, et al. Association between bariatric surgery and long-term survival. JAMA. 2015;313(1):62–70.
9. Schauer PR, Bhatt DL, Kirwan JP, et al. Bariatric surgery versus intensive medical therapy for diabetes — 5-year outcomes. N Engl J Med. 2017;376(7):641–51.
10. Mingrone G, Panunzi S, De Gaetano A, et al. Metabolic surgery versus conventional medical therapy in patients with type 2 diabetes: 10-year follow-up of an open-label, single-centre, randomised controlled trial. Lancet. 2021;397(10271):293–304.
11. Courcoulas AP, Patti ME, Hu B, et al. Long-term outcomes of medical management vs bariatric surgery in type 2 diabetes. JAMA. 2024;331(8):654–64.
12. Birkmeyer NJO, Dimick JB, Share D, et al. Hospital complication rates with bariatric surgery in Michigan. JAMA. 2010;304(4):435–42.
13. American College of Surgeons. Metabolic and Bariatric Surgery Accreditation and Quality Improvement Program: Standards Manual. American College of Surgeons; 2022.
14. Mechanick JI, Apovian C, Brethauer S, et al. Clinical practice guidelines for the perioperative nutrition, metabolic, and nonsurgical support of patients undergoing bariatric procedures – 2019 update: cosponsored by American Association of Clinical Endocrinologists/American College of Endocrinology, The Obesity Society, American Society for Metabolic & Bariatric Surgery, Obesity Medicine Association, and American Society of Anesthesiologists. Surg Obes Relat Dis. 2020;16(2):175–247.
15. Heinze G, Schemper M. A solution to the problem of separation in logistic regression. Stat Med. 2002;21(16):2409–19.
16. Firth D. Bias reduction of maximum likelihood estimates. Biometrika. 1993;80(1):27–38.

17. Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and Cox regression. Am J Epidemiol. 2007;165(6):710–8.
18. Wasserstein RL, Lazar NA. The ASA statement on p-values: context, process, and purpose. Am Stat. 2016;70(2):129–33.
19. Peterli R, Wölnerhanssen BK, Peters T, et al. Effect of laparoscopic sleeve gastrectomy vs laparoscopic Roux-en-Y gastric bypass on weight loss in patients with morbid obesity: the SM-BOSS randomized clinical trial. JAMA. 2018;319(3):255–65.
20. Salminen P, Helmiö M, Ovaska J, et al. Effect of laparoscopic sleeve gastrectomy vs laparoscopic Roux-en-Y gastric bypass on weight loss at 5 years among patients with morbid obesity: the SLEEVEPASS randomized clinical trial. JAMA. 2018;319(3):241–54.
21. Ali M, El Chaar M, Ghiassi S, Rogers AM. American Society for Metabolic and Bariatric Surgery updated position statement on sleeve gastrectomy as a bariatric procedure. Surg Obes Relat Dis. 2017;13(10):1652–7.
22. Aminian A, Brethauer SA, Andalib A, et al. Individualized metabolic surgery score: procedure selection based on diabetes severity. Ann Surg. 2017;266(4):650–7.
23. Sogg S, Lauretti J, West-Smith L. Recommendations for the presurgical psychosocial evaluation of bariatric surgery patients. Surg Obes Relat Dis. 2016;12(4):731–49.
24. Birkmeyer JD, Finks JF, O'Reilly A, et al. Surgical skill and complication rates after bariatric surgery. N Engl J Med. 2013;369(15):1434–42.
25. Kim DD, Arterburn DE, Sullivan SD, Basu A. Economic value of greater access to bariatric procedures for patients with severe obesity and diabetes. Med Care. 2018;56(7):583–8.
26. Aminian A, Zajichek A, Arterburn DE, et al. Association of metabolic surgery with major adverse cardiovascular outcomes in patients with type 2 diabetes and obesity. JAMA. 2019;322(13):1271–82.

Figures

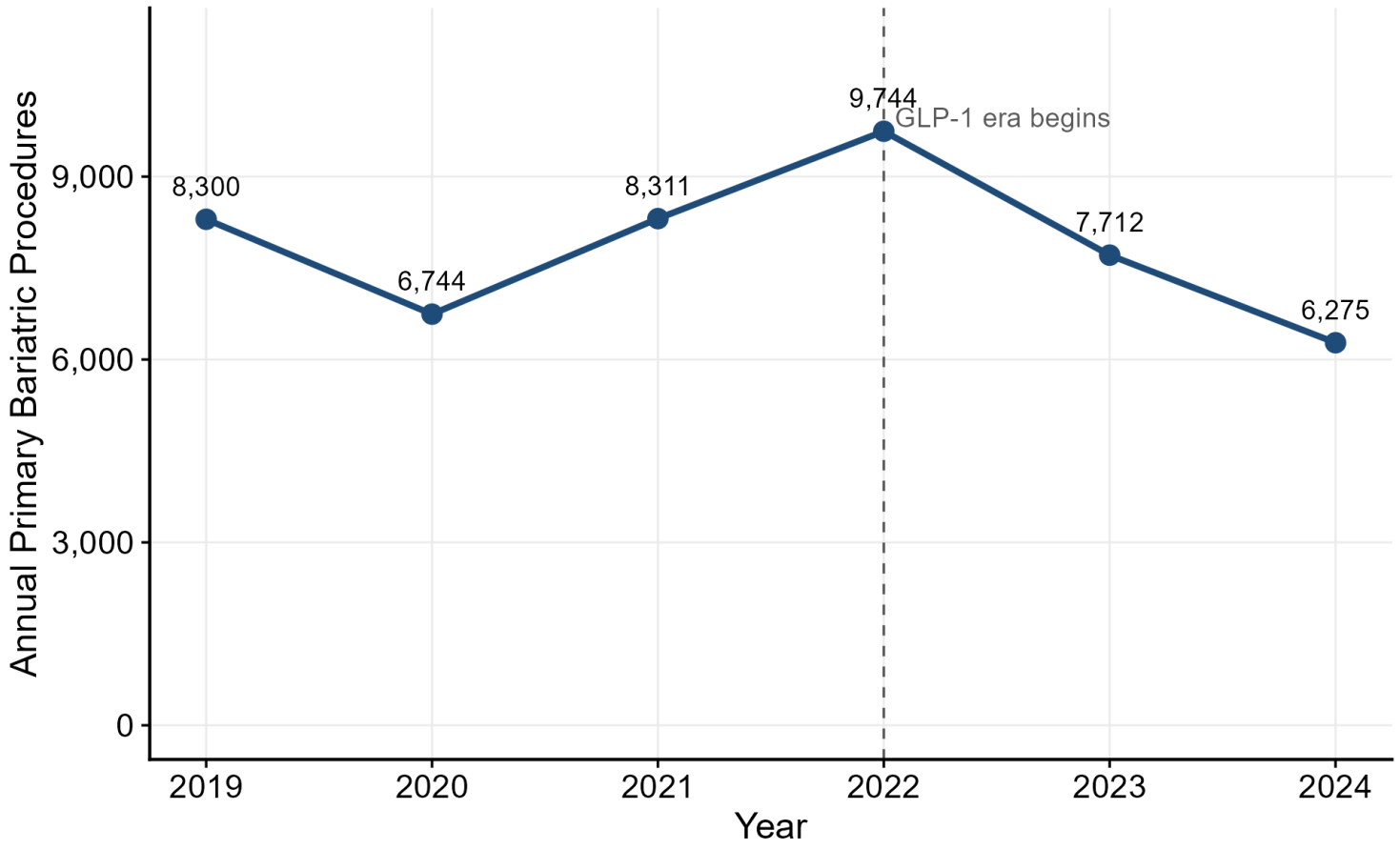

Figure 1

Annual primary bariatric procedure volume across the Michigan Bariatric Surgery Collaborative, 2019–2024.

Notes: Solid line shows total annual procedures across all 41 participating sites. Dashed reference line marks the 2022 GLP-1 era inflection point. Total volume declined from 9,744 procedures in 2022 to 6,275 procedures in 2024 (35.6% decrease).

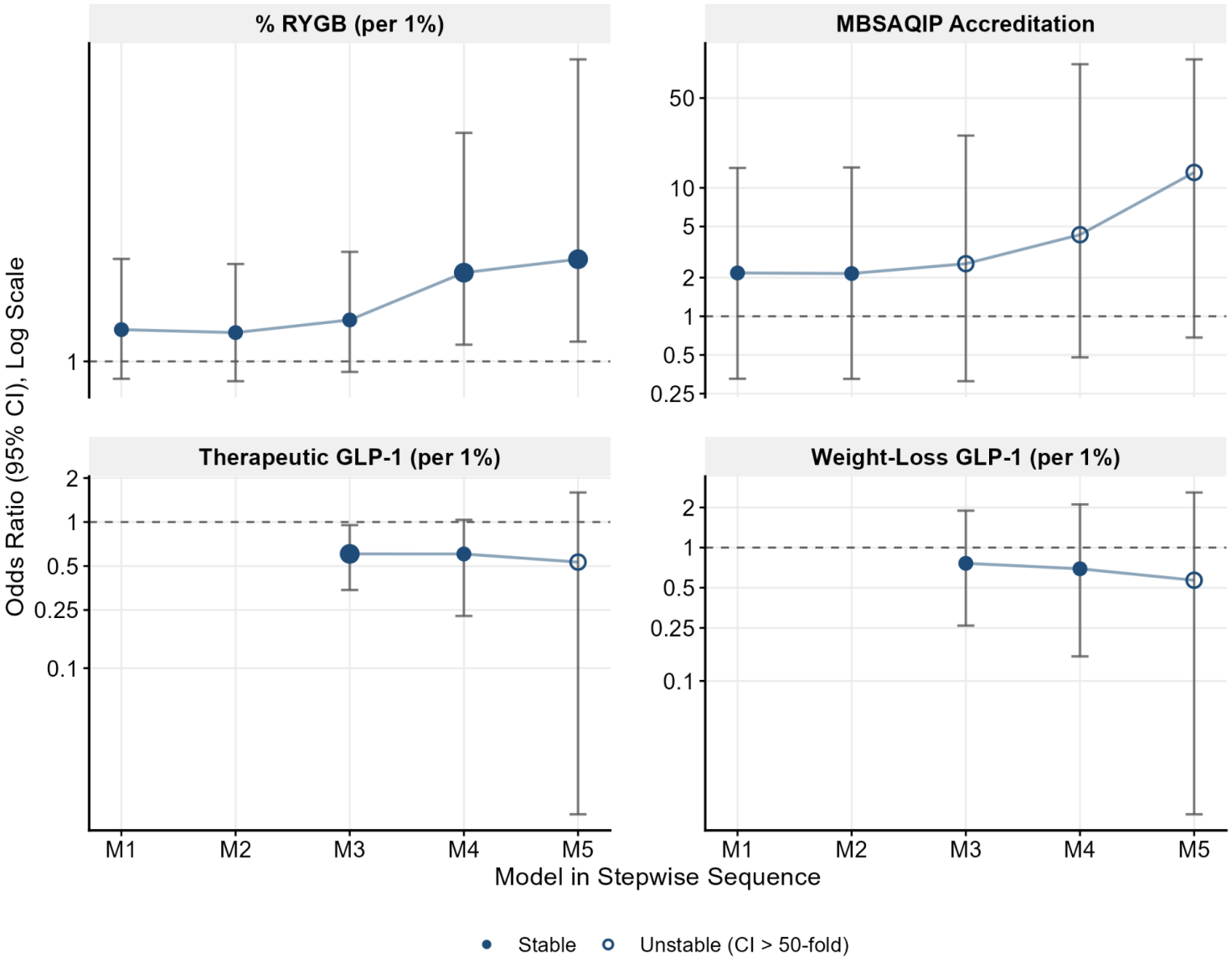

Figure 2

Coefficient stability across the stepwise model sequence (Models 1–5)

Notes: Coefficient stability across the stepwise model sequence (Models 1–5) for four key predictors: percent RYGB, MBSAQIP accreditation, therapeutic GLP-1 use, and weight-loss GLP-1 use. Odds ratios on log scale; 95% confidence intervals shown as error bars; dashed horizontal line at OR=1 (null). Open circles flag estimates with confidence intervals spanning more than 50-fold (unstable). Percent RYGB shows a stable, increasingly significant association with volume loss across the model sequence.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SurgicalVolumeManuscriptSupplementvF.docx