Supplemental Online Content

Rodriguez PJ, Goodwin BM, Gratzl S, et al. Semaglutide vs tirzepatide for weight loss in adults with overweight or obesity. *JAMA Intern Med*. Published online July 8, 2024. doi:10.1001/jamainternmed.2024.2525

This supplemental material has been provided by the authors to give readers additional information about their work.

1 eMethods

1.1 Quantitative Measure Pre-processing: Weight

Several steps were taken to pre-process weight data. First, weights observations with null/missing values were removed.

In instances where a value was recorded but the unit was missing, we assumed values between 40,824 - 317,520 were grams, values between 1,440 - 11,200 were ounces, values greater than 317 were pounds, and values less than 125 were kilograms. These ranges correspond to values of 90 -700 lbs. For this study of adults with overweight or obesity, values equivalent to <90 or >700 pounds were assumed to be entered in error.

For remaining observations where a unit could not be assumed based on value, we assumed the unit based on consistency with other patient values. These were predominately values >125 and <300, which could plausibly correspond to either pounds or kilograms in our population. For each patient, we calculated the mean and standard deviation of weights with known or assumed units in the 15 months before to 15 months after the index date. For each observation with a value but no unit, we calculated the distance from the mean assuming the observation was measured in (a) pounds and (b) kilograms. We then assigned the unit as whichever was closest to the mean, so long as it fell within a plausible distance. We assumed a plausible distance of 30%. We then removed any additional values where a unit could not be assigned.

After reviewing high-variance patient trajectories to identify common patterns, we removed outliers in the patient trajectory that likely represented data entry errors. The primary pattern was a weight bounce, where one weight value was highly inconsistent with both the previous and subsequent weight values. These largely appeared to be incorrect entry of pounds as kilograms and vice versa, resulting in approximate halving or doubling of a single observation (e.g., day 60: 402 pounds, day 0: 400 pounds, day 45: 180 pounds, day 70: 390 pounds). If the preceding weight value was within 365 days and the absolute change was >40% *and* the subsequent weight value was within 45 days and the absolute change was >40%, we removed the value. We also removed instances where one value is >40% different from the next two weight values and all are within 365 days. This was intended to capture cases of bounce, where the first observation was the aberrant value. Next, we removed biologically implausible changes in a short period, those where the preceding weight was within 5 days and absolute change was > 10% *and* the next weight value is within 5 days and absolute change was > 10%. We explored the use of interquartile based approaches for outlier detection but found these to have poor specificity.

Finally, after the cleaning steps taken above, we averaged weights if multiple were taken on the same day.

1.2 Propensity Score Methods

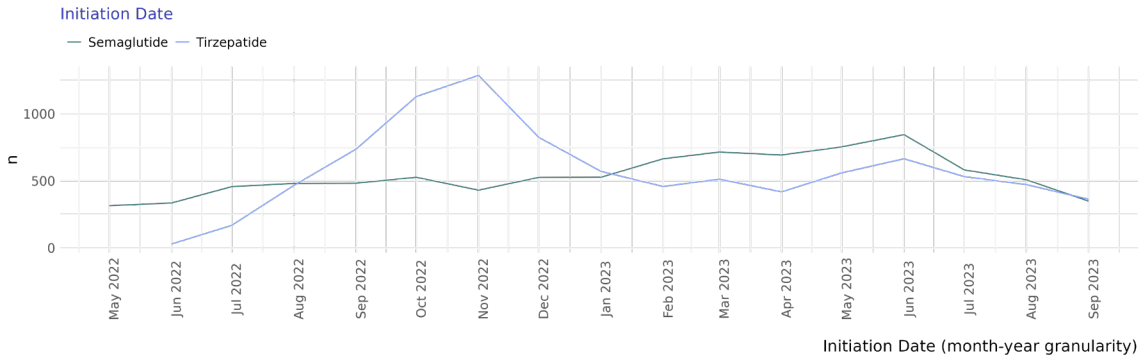
A logistic regression was used to estimate the probability of initiating treatment with tirzepatide, compared to semaglutide, as a function of observed covariates at baseline. Covariates included in the propensity score model were selected based on their potential relationship to treatment choice (tirzepatide vs. semaglutide) and outcomes (magnitude of weight loss). The following baseline covariates were included in the logistic regression: age, race, sex, education (any college on record), Hispanic or Latino ethnicity, patient latest state, months since approval of tirzepatide, baseline weight, duration (years) of overweight/obesity, presence of T2D (e.g., on-label use), duration (years) of T2D (0 for patients without evidence of T2D), number of A1c tests in the previous 2 years, presence of any A1c ≥7 in last 2 years, comorbidity of atrial fibrillation, comorbidity of asthma, comorbidity of CKD, comorbidity of COPD, comorbidity of heart failure, comorbidity of hyperlipidemia, comorbidity of hypertension, comorbidity of ischemic heart disease, comorbidity of major depressive disorder, concurrent or previous use of DPP-4, concurrent or previous use of insulin, concurrent or previous use of metformin, concurrent or previous use of SGLT2i, concurrent or previous use of sulfonylurea, number of interactions in the 6 months before initiation, number of interactions in the 7-12 months before initiation. We did not include previous or concurrent use of AOMs in propensity scores (Orlistat, Phentermine Topiramate) because they were extremely uncommon (<1% of patients). We did not include baseline BMI in propensity scores because it was not always available.

After estimation of propensity scores, 1:1 nearest neighbor matching was used to pair each patient on tirzepatide to a patient on semaglutide who had a similar likelihood of initiating tirzepatide. Unmatched patients were discarded. As such, the matched sample resembles the tirzepatide population, rather than the semaglutide or overall population, and effect estimates will represent the average treatment effect among the treated (ATT) with tirzepatide, rather than the average treatment effect (ATE) among all new users.

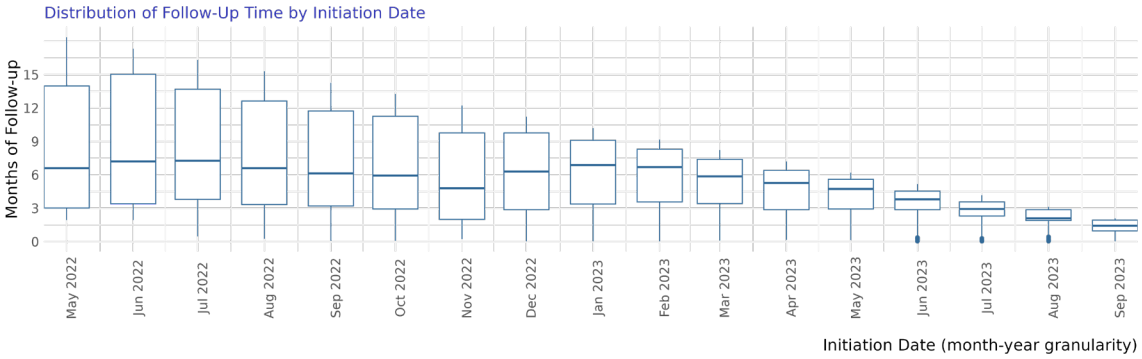
For T2D-stratified subgroup analyses, we first subset to the relevant population (those with T2D and those without T2D, separately) then re-performed propensity score estimation and matching. For these analyses, binary indicators of T2D status were removed from models, given they had no variance in T2D-stratified populations.

1.3 Censoring

Initiation and follow-up times are given in **eFigure 1** and **eFigure 2**. Follow-up times were somewhat shorter for patients initiating medications in Fall 2022, when shortages were reported. Follow-up times increased somewhat for patients initiating in December 2022, around the time some shortages of semaglutide were reduced.



eFigure 1: Distribution of Initiation Time by group



eFigure 2: Distribution of follow-up time by initiation date

1.4 Availability of Follow-up Weight

Of 41,222 patients overall (before matching), 35,097 (85%) had at least one follow-up weight during observation (ITT), and 31,419 (76%) had at least one follow-up weight while on treatment. There was an average of 5.8 follow-up weights per person-year observed (ITT) and 6.12 weights per person-year on treatment. After matching, 15,522 (84%) patients had at least one follow-up weight during observation, and 13,789 (75%) had at least one follow-up weight while on treatment. There was an average of 5.4 follow-up weights per person-year observed (ITT) and 5.7 weights per person-year on treatment.

Analyses of percentage change in body weight at fixed time points only considered follow-up weights taken near the time points of interest, t, where $t \in 3,6,12$ months. As such, patient follow-up weights at t are unavailable for patients censored before t, and for those not censored who lack a follow-up value in the acceptable range (45 days) of t. In **eTable 1**, we present the number and proportion of patients at each t who (1) were theoretically eligible for follow-up at t (initiated at least t months before study end), and (2) were actually observed until t (last encounter $\geq t$), and (3) were still on treatment by t, and (4) had a weight value available in the 45 days before or after t.

eTable 1: Weight Availability at t: On-Treatment Analyses

Step	t = 3 months	t = 6 months	t = 12 months
(1) Initiated t months before study end	39728 (96.4%)	31708 (76.9%)	15192 (36.9%)
(2) Still observed until t or later	37221 (90.3%)	28389 (68.9%)	11861 (28.8%)
(3) On treatment at t	28175 (68.3%)	15335 (37.2%)	3999 (9.7%)
(4) Had on-treatment weight at t	20863 (50.6%)	10427 (25.3%)	2482 (6%)

eTable 2 provides similar information for ITT analyses: the number and proportion of patients at each t who (1) were theoretically eligible for follow-up at t (because they initiated at least t months before study end), and (2) were actually observed until t (last encounter $\geq t$), and (3) had a weight value available in the 45 days before or after t, regardless of treatment status at t.

eTable 2: Weight Availability at t: ITT Analyses

Step	t = 3 months	t = 6 months	t = 12 months
(1) Initiated t months before study end	39728 (96.4%)	31708 (76.9%)	15192 (36.9%)
(2) Still observed until t or later	37221 (90.3%)	28389 (68.9%)	11861 (28.8%)
(3) Had weight at t	26596 (64.5%)	18300 (44.4%)	6774 (16.4%)

For all t, the majority of patients still under observation at t had weight values at t. We describe the process for imputing weight among patients for whom weight was unavailable at t in the following section.

eTable 3: Characteristics of Matched Patients with Available vs. Missing Follow-up Weight. *Categorical variables expressed as number (percentage). Durations (overweight/obesity and T2D) refer to time (years) since first evidence in EHR. Other state includes unknown and states with <3% of the pre-match sample: Alaska, Arizona, Arkansas, Colorado, Florida, Georgia, Idaho, Indiana, Iowa, Kansas, Kentucky, Louisiana, Michigan, Missouri, Montana, Nevada, New Mexico, North Carolina, Oklahoma, Oregon, Pennsylvania, South Carolina, Tennessee, Utah, Virginia, and Unknown. Other race includes American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, Other Race, Unknown, Declined to answer. Absolute SMD are reported. Abbreviations: ADM = anti-diabetic medication, AOM = anti-obesity medication, Black = Black or African American, BMI = body mass index, CKD = chronic kidney disease, COPD = chronic obstructive pulmonary disease, DPP4 = dipeptidyl peptidase 4 inhibitor, MI = myocardial infarction, SD = standard deviation, SGLT2i = sodium/glucose cotransporter-2 inhibitor, SMD = standardized mean difference, T2D = type 2 diabetes mellitus.*

Characteristic	Available ITT Weight (N=15,522)	Missing Follow-up ITT Weight (N=2,864)	Available On-Treatment Weight (N=13,789)	Missing On-treatment Weight (N=4,597)	Overall (N=18,386)
Age	52.0 (13.0)	51.9 (12.7)	52.2 (13.0)	51.3 (12.7)	52.0 (12.9)
Female	11,142 (71.8%)	1,828 (63.8%)	9,952 (72.2%)	3,018 (65.7%)	12,970 (70.5%)
Race					
Asian	280 (1.8%)	74 (2.6%)	250 (1.8%)	104 (2.3%)	354 (1.9%)
Black	1,831 (11.8%)	340 (11.9%)	1,617 (11.7%)	554 (12.1%)	2,171 (11.8%)
White	11,999 (77.3%)	2,183 (76.2%)	10,696 (77.6%)	3,486 (75.8%)	14,182 (77.1%)
Other	1,412 (9.1%)	267 (9.3%)	1,226 (8.9%)	453 (9.9%)	1,679 (9.1%)
Ethnicity					
Hispanic or Latino	2,392 (15.4%)	396 (13.8%)	2,120 (15.4%)	668 (14.5%)	2,788 (15.2%)
Not Hispanic or Latino	12,439 (80.1%)	2,332 (81.4%)	11,068 (80.3%)	3,703 (80.6%)	14,771 (80.3%)
Other	691 (4.5%)	136 (4.7%)	601 (4.4%)	226 (4.9%)	827 (4.5%)
Education: Any college on record	9,225 (59.4%)	1,705 (59.5%)	8,174 (59.3%)	2,756 (60.0%)	10,930 (59.4%)
Income Range					
0 - 25,000	431 (2.8%)	79 (2.8%)	379 (2.7%)	131 (2.8%)	510 (2.8%)
25,001 - 50,000	6,451 (41.6%)	1,163 (40.6%)	5,772 (41.9%)	1,842 (40.1%)	7,614 (41.4%)
50,001 - 80,000	5,866 (37.8%)	1,078 (37.6%)	5,180 (37.6%)	1,764 (38.4%)	6,944 (37.8%)
More than 80K	2,447 (15.8%)	471 (16.4%)	2,183 (15.8%)	735 (16.0%)	2,918 (15.9%)
Unknown	327 (2.1%)	73 (2.5%)	275 (2.0%)	125 (2.7%)	400 (2.2%)
State					
Texas	6,870 (44.3%)	1,096 (38.3%)	6,019 (43.7%)	1,947 (42.4%)	7,966 (43.3%)

Wisconsin	912 (5.9%)	261 (9.1%)	825 (6.0%)	348 (7.6%)	1,173 (6.4%)
Illinois	882 (5.7%)	287 (10.0%)	781 (5.7%)	388 (8.4%)	1,169 (6.4%)
Ohio	1,605 (10.3%)	218 (7.6%)	1,442 (10.5%)	381 (8.3%)	1,823 (9.9%)
Washington	772 (5.0%)	143 (5.0%)	704 (5.1%)	211 (4.6%)	915 (5.0%)
California	1,315 (8.5%)	189 (6.6%)	1,187 (8.6%)	317 (6.9%)	1,504 (8.2%)
Other or Unknown	3,166 (20.4%)	670 (23.4%)	2,831 (20.5%)	1,005 (21.9%)	3,836 (20.9%)
Weight (in kg)	110 (25.8)	110 (25.7)	110 (25.7)	111 (25.8)	110 (25.8)
BMI*	39.2 (8.12)	38.5 (7.87)	39.2 (8.10)	38.7 (8.01)	39.1 (8.09)
Unknown	2,021 (13.0%)	392 (13.7%)	1,719 (12.5%)	694 (15.1%)	2,413 (13.1%)
Years since First Overweight/Obesity	4.47 (3.09)	4.59 (3.25)	4.51 (3.10)	4.42 (3.15)	4.49 (3.11)
T2D	7,893 (50.9%)	1,670 (58.3%)	7,102 (51.5%)	2,461 (53.5%)	9,563 (52.0%)
Years since First T2D (among all patients)	3.49 (3.31)	3.46 (3.19)	3.51 (3.34)	3.42 (3.13)	3.48 (3.29)
Number of HbA1c Tests in Previous 2 Years	2.19 (1.89)	2.25 (1.84)	2.21 (1.90)	2.18 (1.83)	2.20 (1.88)
Months since May 2022	7.95 (4.02)	12.1 (3.61)	8.10 (4.07)	10.0 (4.38)	8.59 (4.23)
Bariatric Surgery History	662 (4.3%)	90 (3.1%)	606 (4.4%)	146 (3.2%)	752 (4.1%)
Comorbidities					
Atrial Fibrillation	701 (4.5%)	108 (3.8%)	642 (4.7%)	167 (3.6%)	809 (4.4%)
Asthma	2,878 (18.5%)	469 (16.4%)	2,600 (18.9%)	747 (16.2%)	3,347 (18.2%)
CKD	1,458 (9.4%)	227 (7.9%)	1,337 (9.7%)	348 (7.6%)	1,685 (9.2%)
COPD	728 (4.7%)	100 (3.5%)	682 (4.9%)	146 (3.2%)	828 (4.5%)
Glaucoma	228 (1.5%)	40 (1.4%)	212 (1.5%)	56 (1.2%)	268 (1.5%)
Heart Failure	788 (5.1%)	106 (3.7%)	733 (5.3%)	161 (3.5%)	894 (4.9%)
Hyperlipidemia	9,737 (62.7%)	1,801 (62.9%)	8,674 (62.9%)	2,864 (62.3%)	11,538 (62.8%)
Hypertension	9,745 (62.8%)	1,743 (60.9%)	8,721 (63.2%)	2,767 (60.2%)	11,488 (62.5%)
Ischemic Heart Disease	786 (5.1%)	109 (3.8%)	712 (5.2%)	183 (4.0%)	895 (4.9%)
Osteoporosis	488 (3.1%)	62 (2.2%)	453 (3.3%)	97 (2.1%)	550 (3.0%)
Acute MI	266 (1.7%)	40 (1.4%)	238 (1.7%)	68 (1.5%)	306 (1.7%)

Ischemic Stroke	15 (0.1%)	4 (0.1%)	14 (0.1%)	5 (0.1%)	19 (0.1%)
Major Depressive Disorder	3,531 (22.7%)	513 (17.9%)	3,181 (23.1%)	863 (18.8%)	4,044 (22.0%)
ADM Use					
DPP-4	1,126 (7.3%)	199 (6.9%)	1,004 (7.3%)	321 (7.0%)	1,325 (7.2%)
Insulin	810 (5.2%)	163 (5.7%)	734 (5.3%)	239 (5.2%)	973 (5.3%)
Metformin	6,958 (44.8%)	1,438 (50.2%)	6,238 (45.2%)	2,158 (46.9%)	8,396 (45.7%)
SGLT2i	1,956 (12.6%)	382 (13.3%)	1,763 (12.8%)	575 (12.5%)	2,338 (12.7%)
Sulfonylurea	1,835 (11.8%)	340 (11.9%)	1,625 (11.8%)	550 (12.0%)	2,175 (11.8%)
AOM Use					
Orlistat	34 (0.2%)	4 (0.1%)	32 (0.2%)	6 (0.1%)	38 (0.2%)
Phentermine Topiramate	146 (0.9%)	18 (0.6%)	129 (0.9%)	35 (0.8%)	164 (0.9%)

1.5 Methods for Weight Missingness at Fixed Time Points

Primary, on-treatment analyses considered the population still on-treatment at *t* (step 3 of **eTable 1**). For patients still on treatment but without weight measurements at *t*, possible weight values were imputed using multiple imputation via chained equations.

The following process was taken to impute weight and conduct statistical analyses for primary, on-treatment analyses:

1. Subset to patients still on treatment at *t* (step 3 in **eTable 1**)

2. For each *t*, identify whether a weight value was available for each patient *i* (step 4 in **eTable 1**). If 1 or more weight values are available, use the value closest to *t*. In this analysis, percentage change in weight values at *t*, rather than raw weight values at *t*, are used.
3. For each *t*, if a weight (as weight change) value is unavailable, use multiple imputation via chained equations to impute *m* =10 potential or plausible values of weight change for patient *i*. Imputation assumes missing weights are missing at random (MAR), which allows for the missingness *t* to be conditional on all observed covariates and on outcomes observed at other time points. Imputation therefore used all available information, including baseline covariates and outcomes observed at other time points. As an output of this step, a total of *m* complete datasets are available for each *t*, each representing one potential complete realization of the data.

4. At each *t*, for each *m* dataset, conduct propensity score matching (same process detailed in section 1.2) to balance treatment groups with respect to baseline covariates. As an output of this step, a total of *m* complete, matched samples are available.

5. At each *t*, for each *m* complete, matched sample, estimate the difference in body weight change between tirzepatide and semaglutide using generalized linear models. As an output of this step, for each time point *t*, *m* estimates of difference are available.

6. For each *t*, pool estimates across *m* to create a single estimate of treatment effect and variance at *t* that incorporates the total variance across imputations.

For ITT analyses, the same process was repeated with 2 deviations. First, weight values within the acceptable range of *t* were included regardless of treatment status at *t* (step 3 in **eTable 2)**. Second, indicators of treatment status at *t* (discontinued, switched, or still on treatment) and days since discontinuation or switching (0 for patients still on treatment) were included in the imputation, given the true weight value at *t* is likely conditional on whether the patient was still receiving treatment.

2 eResults

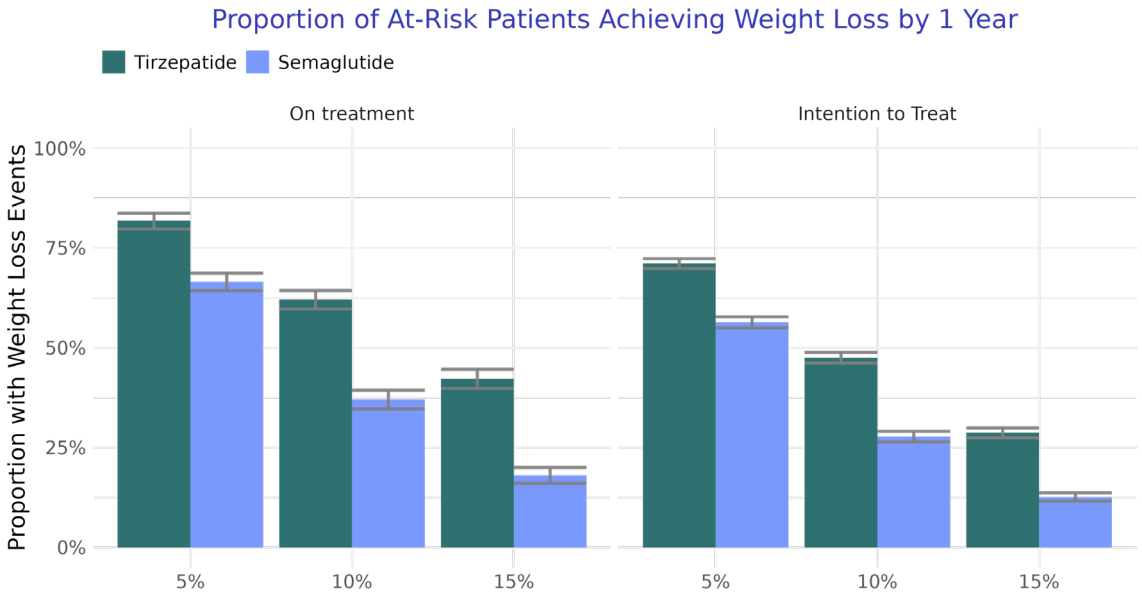
2.1 Sensitivity Analyses

Sensitivity analyses included use of inverse probability of treatment weighting (IPTW), complete case analysis (e.g., no imputation), and modified intention-to-treat (ITT) analysis.

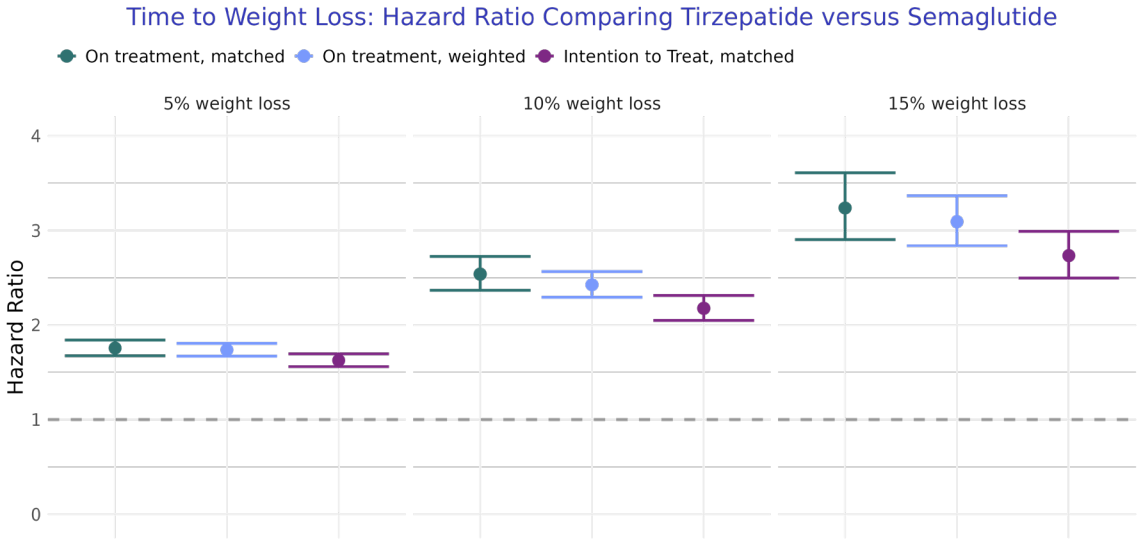
2.1.1 Hazard of Achieving ≥5%, ≥10%, and ≥15%Weight Loss

In modified intention-to-treat analyses where patients were administratively censored only (discontinuation and switching were ignored), fewer patients achieved weight reductions 5%, 10%, and 15% (**eFigure 3**). Treatment effects were somewhat attenuated (**eFigure 4**).

Among a matched subset of patients with any follow-up weight on treatment available, 87.8% (95% CI: 86.0%, 89.3%) on tirzepatide vs. 71.8% (69.7%, 73.8%) on semaglutide achieved ≥5% weight loss, 66% (63.6%, 68.2%) vs. 41.1% (38.7%, 43.4%) achieved ≥10% weight loss, and 44.5% (42%, 46.8%) vs. 19.7% (17.7%, 21.7%) achieved ≥15% weight loss within 365 days. After adjustment for residual confounding, the comparative hazards of achieving the specified weight loss for tirzepatide vs. semaglutide were 1.81 (1.73, 1.9), 2.41 (2.25, 2.59), 3.02 (2.71, 3.36) for ≥5%, ≥10%, ≥15% weight loss, respectively.



eFigure 3: Proportion of at-risk patients achieving weight loss targets by one year for on treatment and intention to treat analyses

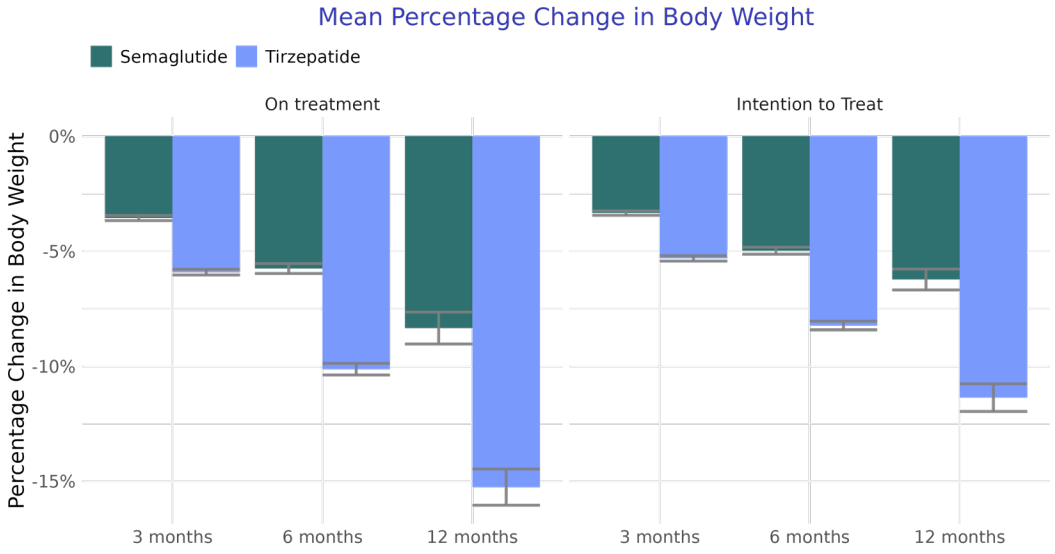


eFigure 4: Hazard ratio comparing tirzepatide vs semaglutide for achieving weight loss targets under different analytic approaches

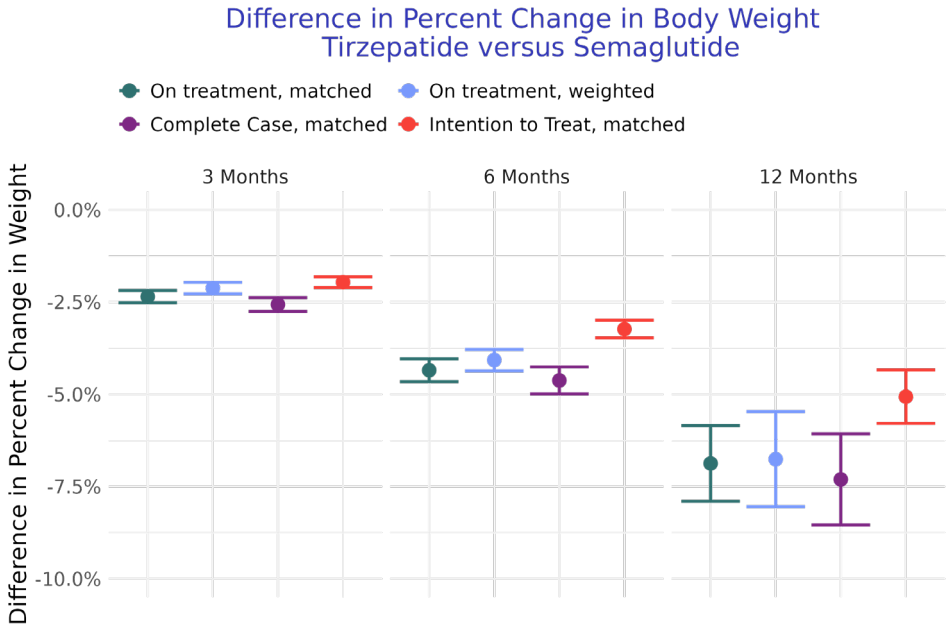
2.1.2 Change in Body Weight

Modified intention-to-treat analyses of percentage changes in body weight included patients not yet administratively censored at the timepoint. Both pre- and post- discontinuation weights were included. Reductions in weight were smaller at all timepoints under this analysis (eFigure 5). Treatment effects were somewhat attenuated (eFigure 6).

Among those with follow-up weights available at the timepoint of interest (complete cases), the mean on-treatment weight change was -6.5% (-6.6%, -6.3%) for tirzepatide and -4.1% (-4.2%, -3.9%) for semaglutide at 3 months, -10.9% (-11.2%, -10.7%) for tirzepatide and -6.4% (-6.7%, -6.1%) for semaglutide at 6 months, and -16.6% (-17.5%, -15.7%) for tirzepatide and -9.2% (-10.2%, -8.3%) for semaglutide at 12 months. The absolute difference in weight loss, after adjustment for residual confounding, was -2.6% (-2.7%, -2.4%) at 3 months, -4.6% (-5%, -4.3%) at 6 months, and -7.3% (-8.5%, -6.1%) at 12 months.



eFigure 5: Mean change in body weight for tirzepatide and semaglutide groups under on treatment and modified intention to treat analyses.

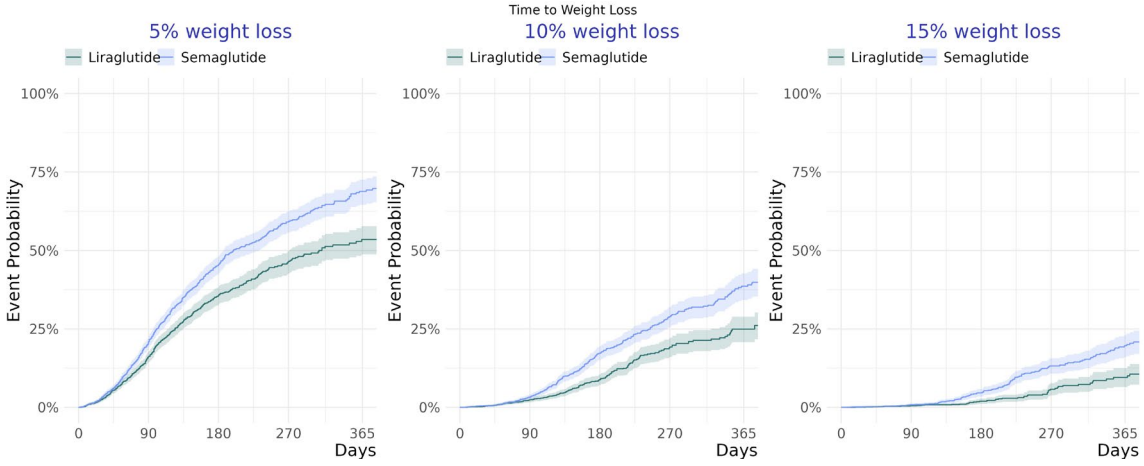


eFigure 6: Difference in percent change in body weight comparing tirzepatide to semaglutide under different analytic approaches.

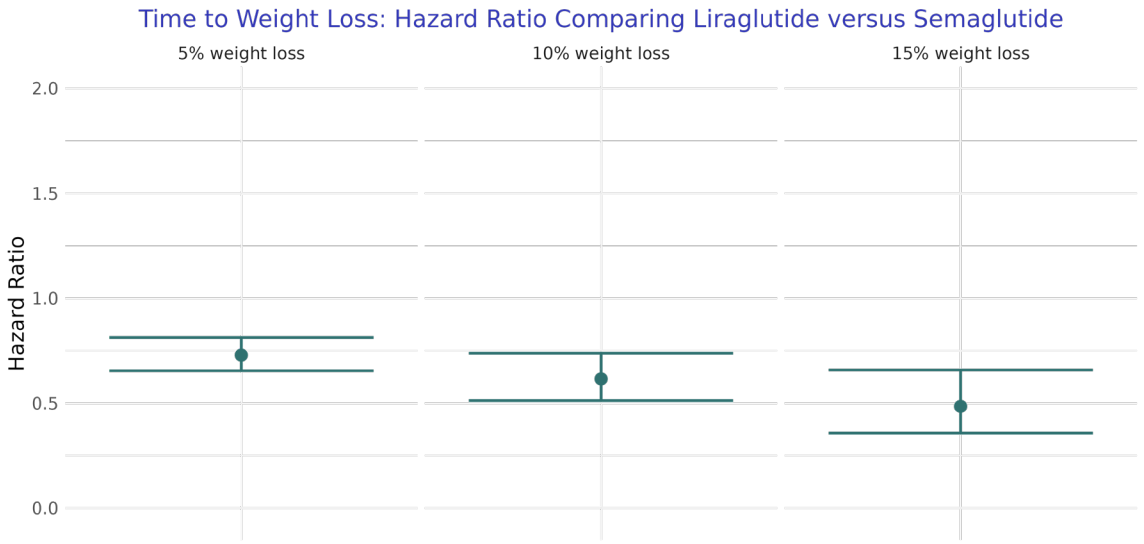
2.2 Validation Analysis (Liraglutide)

In a validation analysis, patients initiating liraglutide labeled for T2D (brand name: Victoza) were 1:1 propensity score matched to patients initiating semaglutide. The analysis steps described in the methods section were repeated for this group.

2.2.1 Hazard of Achieving ≥5%, ≥10%, and ≥15%Weight Loss

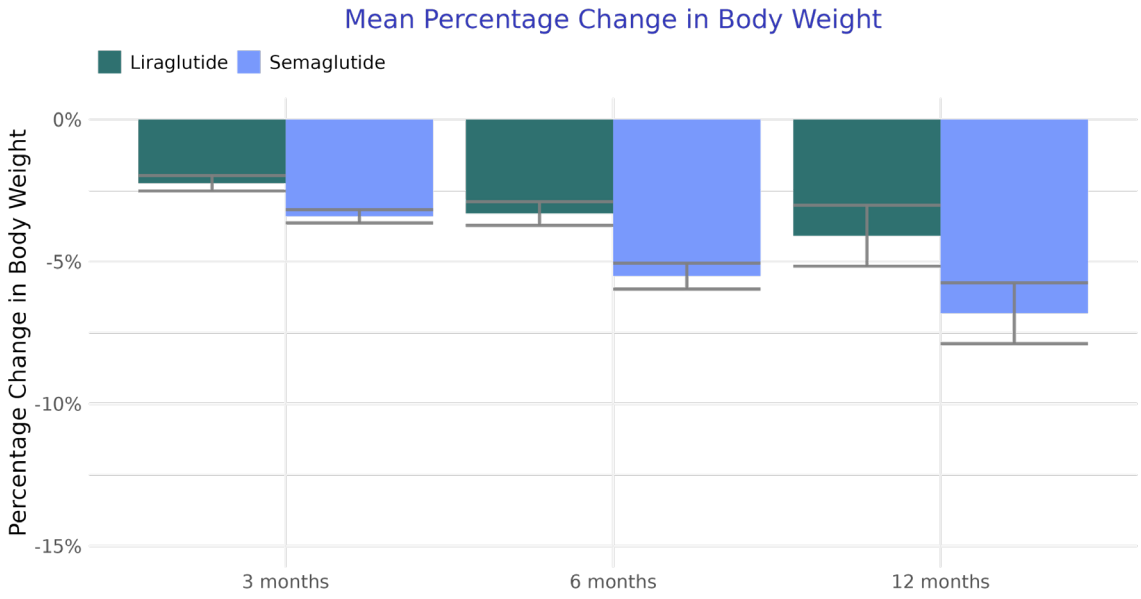


eFigure 7: Event probabilities of weight loss, accounting for censoring, for patients on liraglutide and semaglutide

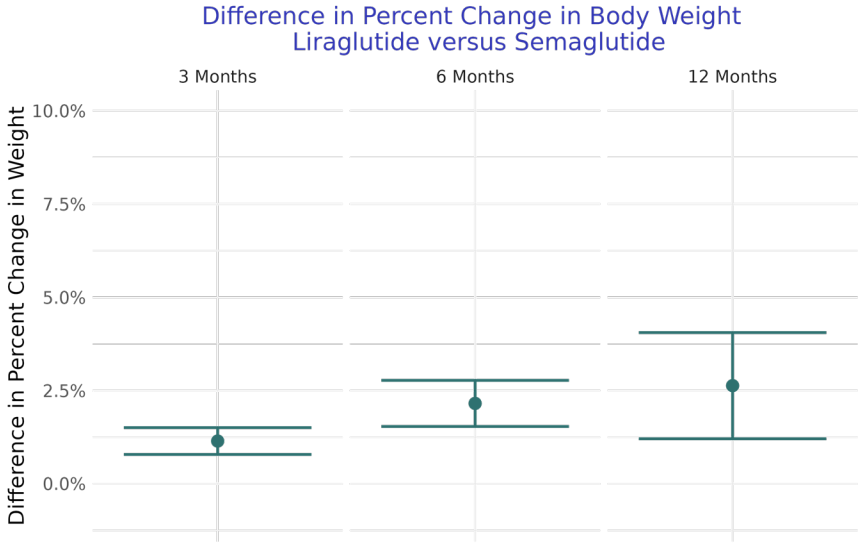


eFigure 8: Hazard ratios comparing liraglutide to semaglutide

2.2.2 Change in Body Weight



eFigure 9: Mean change in body weight for liraglutide and semaglutide

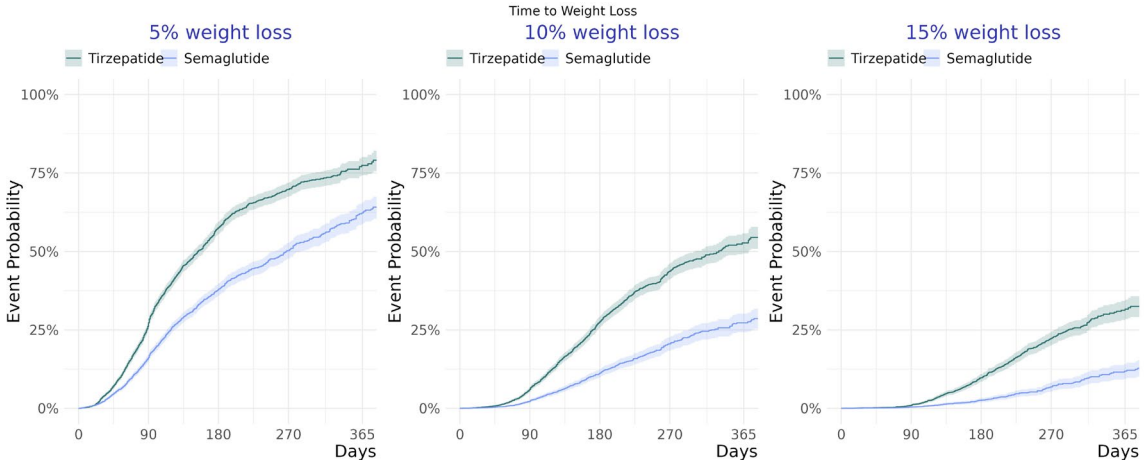


eFigure 10: Difference in percent change in body weight comparing liraglutide to semaglutide

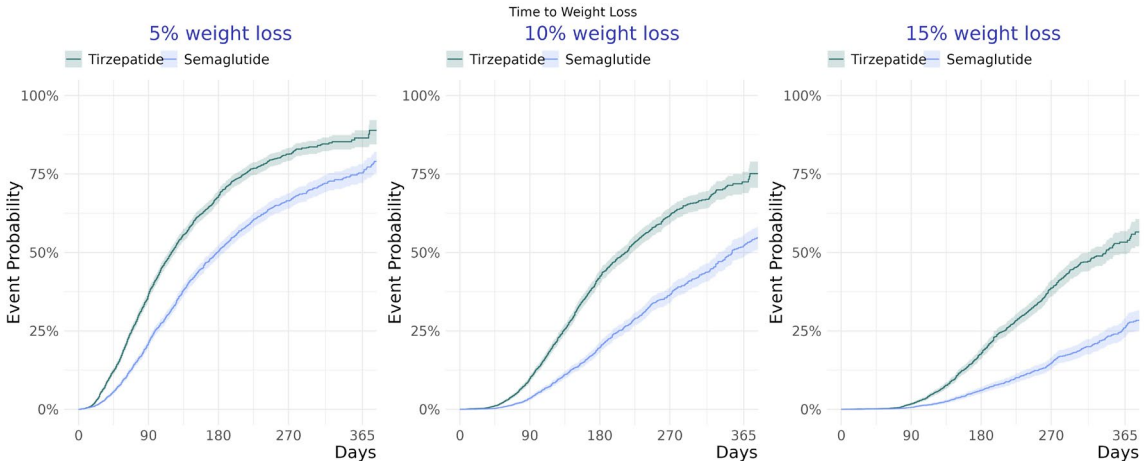
2.3 Subgroup Analyses

The following analyses were conducted on subgroups stratified by T2D (e.g., on-label use) and no T2D (e.g., off-label use).

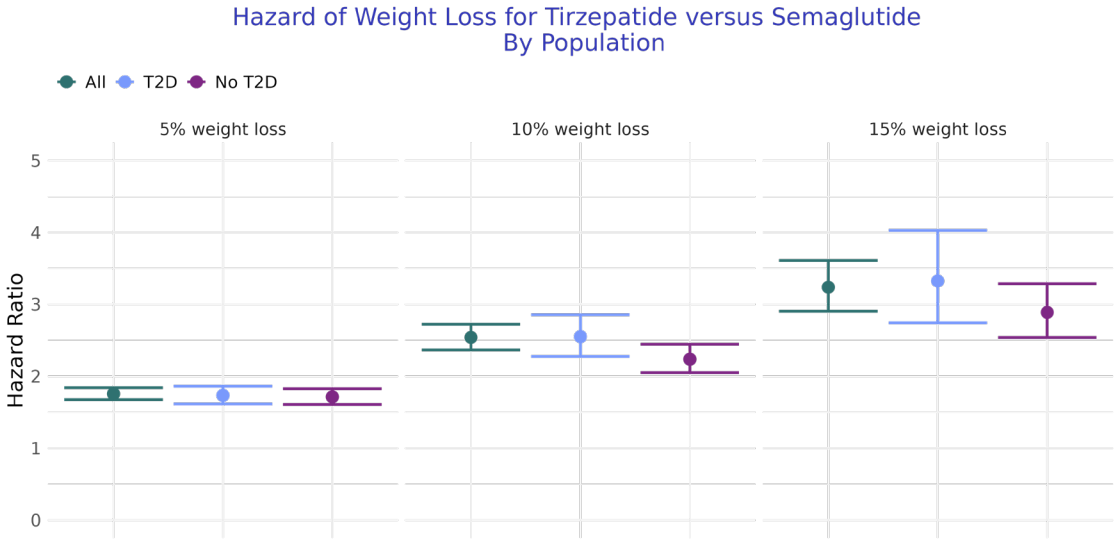
2.3.1 Hazard of Achieving ≥5%, ≥10%, and ≥15%Weight Loss



eFigure 11: Weight loss event probabilities, accounting for censoring, for patients with T2D.

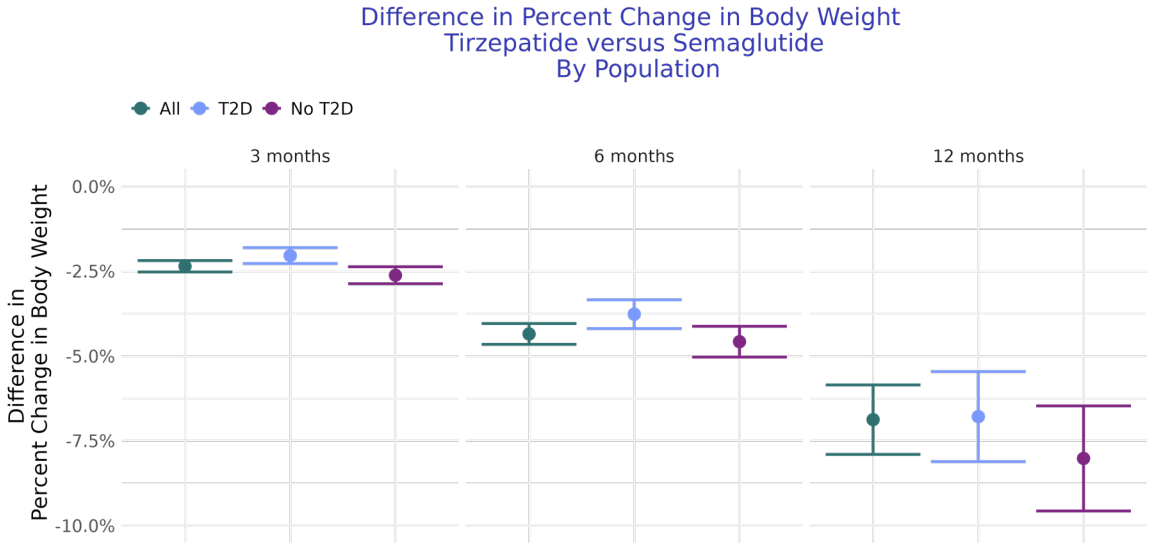


eFigure 12: Weight loss event probabilities, accounting for censoring, for patients without T2D.



eFigure 13: Hazard ratio comparing tirzepatide to semaglutide for different populations.

2.2.2 Change in Body weight



eFigure 14: Difference in percent change in body weight comparing tirzepatide to semaglutide for different populations.

2.3 Gastrointestinal Adverse Events

eTable 4: Gastrointestinal Adverse Event Rates Per 1000 person-years

Adverse Event	Tirzepatide	Semaglutide	Harzard Ratio (CI)
Bowel Obstruction	6.26	5.54	1.12 (0.63-1.97)
Cholecystitis	6.50	5.06	1.25 (0.7-2.21)
Cholelithiasis	11.89	12.66	0.94 (0.63-1.39)
Gastroenteritis	19.75	20.07	1 (0.73-1.37)
Gastroparesis	3.61	4.81	0.73 (0.38-1.42)
Pancreatitis	3.84	3.60	1.04 (0.52-2.11)