

Real-World Comparative Weight Loss of GLP-1 Receptor Agonists in the All of Us Research Program: A Retrospective Cohort Study

Inyoung Hwang¹, Sang Won Lee^{1,2}, Yun Kim^{3,4}

¹Department of Clinical Pharmacology and Therapeutics, Hanyang University Seoul Hospital, Seoul, 04763, Republic of Korea; ²Department of Pharmacology, Hanyang University College of Medicine, Seoul, 04763, Republic of Korea; ³College of Pharmacy, Daegu Catholic University, Gyeongsan, Gyeongbuk, 38430, Republic of Korea; ⁴Convergence Research Institute for Biomedical Sciences, Daegu Catholic University, Gyeongsan, Gyeongbuk, 38430, Republic of Korea

Correspondence: Sang Won Lee, Department of Pharmacology, Hanyang University College of Medicine, 222 Wangsimni-ro, Seongdong-gu, Seoul, 04763, Republic of Korea, Tel +82 02-2220-0657, Email berserk92@hanyang.ac.kr; Yun Kim, College of Pharmacy, Daegu Catholic University, 13-13 Hayang-ro, Hayang-eup, Gyeongsan, Gyeongbuk, 38430, Republic of Korea, Tel +82-53-850-3609, Email ykim71@cu.ac.kr

Purpose: To assess the real-world associations of five glucagon-like peptide-1 receptor agonists (GLP-1RAs) with weight loss in a diverse US population of adults with obesity, with and without type 2 diabetes, and whether the relative ordering of weight loss observed in trials is mirrored in routine practice.

Patients and Methods: We conducted a retrospective, new-user, active-comparator cohort study using the National Institutes of Health *All of Us* Research Program. We included 14,046 adults with obesity (body mass index [BMI] ≥ 30 kg/m²) initiating exenatide, liraglutide, dulaglutide, semaglutide, or tirzepatide. Groups were compared after multivariable adjustment for baseline confounders. Participants were followed from the index date until the outcome, death, or loss to follow-up; percent weight and BMI change was assessed at 12 months (± 45 days). Cox proportional hazards models estimated the relative likelihood over time of achieving weight-loss thresholds ($\geq 5\%$, $\geq 10\%$, $\geq 15\%$) and ≥ 1 BMI-category reduction; linear regression assessed the magnitude of weight change, with liraglutide as reference.

Results: Tirzepatide showed the highest likelihood of $\geq 15\%$ weight loss (adjusted hazard ratio [aHR] 5.57; 95% CI, 3.66–8.49) and the greatest mean weight loss at 12 months (-13.0%), followed by semaglutide (aHR 1.77; 95% CI, 1.56–2.01; -5.9%); both were significantly greater than the older agents in adjusted analyses. Exenatide, dulaglutide, and liraglutide showed comparable, modest weight loss (approximately -4.0%). This ordering was consistent regardless of type 2 diabetes status, although weight loss was greater without diabetes for semaglutide but similar for tirzepatide.

Conclusion: In this large, diverse real-world cohort, tirzepatide and semaglutide were associated with significantly greater weight loss than older agents. These findings are consistent with the relative ordering reported in trials, though the observational design cannot confirm it. Residual confounding cannot be excluded, and the small tirzepatide sample ($n=386$) and limited follow-up warrant cautious interpretation.

Plain Language Summary:

Why was the study done? Obesity is a long-term condition, and many people need more than lifestyle changes to manage it. Newer medicines called GLP-1 receptor agonists can reduce appetite and have worked well in clinical trials. But trials enroll selected volunteers who may differ from people in everyday care. We compared how five of these medicines work for weight loss in a large, diverse, real-world group.

What did the researchers do and find? We studied health records from 14,046 adults with obesity in the National Institutes of Health *All of Us* Research Program. We compared how much weight patients lost in one year after starting exenatide, liraglutide, dulaglutide, semaglutide, or tirzepatide. On average, tirzepatide users lost about 13.0% and semaglutide users about 5.9%, while patients taking the three older medicines lost around 4%. Tirzepatide users were also the most likely to achieve large weight loss of 15% or more.

What do these results mean? These findings suggest the two newest medicines, tirzepatide and semaglutide, are linked to greater weight loss than older options, even in everyday care. Greater weight loss is generally linked to better health, so these differences

could matter for patients, though our study cannot prove the drugs caused them. Because we used everyday medical records rather than a clinical trial, the results show links rather than proof of cause and effect. Tirzepatide is new, so few patients had taken it; larger studies are needed to confirm these early findings.

Keywords: obesity management, incretin mimetics, pharmacoepidemiology, body mass index, type 2 diabetes mellitus

Introduction

Obesity is a chronic disease requiring long-term management, yet lifestyle interventions alone often fail to sustain weight loss.^{1–4} Consequently, pharmacotherapy has become an essential adjunct, a therapeutic landscape recently transformed by the development of glucagon-like peptide-1 receptor agonists (GLP-1RAs).^{5,6} These agents mimic endogenous GLP-1 to regulate appetite and glucose homeostasis, offering significantly improved efficacy and safety profiles compared to earlier obesity pharmacotherapies such as orlistat and sympathomimetic agents.^{7,8}

The efficacy of the GLP-1RA class is well-established in landmark clinical trials. For instance, the SCALE trial demonstrated 8.0% mean weight loss with liraglutide, while the STEP 1 and SURMOUNT-1 trials reported even greater reductions with semaglutide (14.9%) and the dual glucose-dependent insulinotropic polypeptide (GIP)/GLP-1 receptor agonist tirzepatide (up to 20.9%), respectively.^{9–11} Emerging evidence further indicates that tirzepatide is associated with substantial weight reduction and favorable cardiometabolic effects in individuals with obesity without diabetes, further expanding the therapeutic potential of incretin-based pharmacotherapies.¹² While only specific agents are FDA-approved for chronic weight management, others like exenatide and dulaglutide are also utilized in clinical settings.¹³ These agents vary significantly in their molecular structure, half-life, and receptor engagement profiles, which may influence their comparative clinical performance.

Although several head-to-head randomized trials have directly compared individual pairs of GLP-1RAs, large-scale real-world comparisons across the entire class in demographically and clinically diverse populations remain limited.¹³ Real-world data are therefore needed to evaluate whether the relative ordering of weight loss is reproduced under routine prescribing, adherence, and titration patterns, particularly in populations (including those varying in age, sex, race and ethnicity, socioeconomic status, comorbidity burden, and type 2 diabetes status) that are often underrepresented in clinical trials. Therefore, this study aims to assess the real-world associations of initiating exenatide, liraglutide, dulaglutide, semaglutide, or tirzepatide with the relative likelihood over time of achieving prespecified weight-loss thresholds and with the magnitude of weight change at 12 months among adults with obesity using data from the National Institutes of Health's *All of Us* Research Program.

Material and Methods

Data Source

We used data from the National Institutes of Health (NIH) *All of Us* Research Program. The program, which opened for enrollment in May 2018, aims to enroll a diverse group of at least one million individuals across the United States to accelerate precision medicine research.¹⁴ As of the Curated Data Repository (CDR) version 8, the data included 633,547 participants.¹⁵ CDR v8 (released February 2025) includes prospectively collected program-level data from enrollment in May 2018 through 1 October 2023, while retrospective electronic health records (EHR) ingested from site-based data partners extend back further. The program aggregates diverse data types, including participant-provided information (PPI) from health questionnaires, EHR, physical measurements, the use of digital health technology (currently including Fitbit data), and biospecimens (including genomic data such as Whole Genome Sequence and array data).¹⁶ For the EHR data, *All of Us* utilizes the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM), which standardizes information across multiple domains.¹⁷

This study integrated data from multiple OMOP domains corresponding to the SQL extraction strategy: demographics and socioeconomic status were derived from the *Person* and *Survey* tables; clinical laboratory values and vital signs were extracted from the *Measurement* table; medication exposure history was obtained from the *Drug Exposure* table; and

medical history was retrieved from the *Condition Occurrence* table (Table S1). Additionally, the *Visit Occurrence* table was utilized to assess healthcare system interaction, and the *Death* table was used to identify mortality events for censoring. Unlike insurance claims databases, *All of Us* does not capture continuous insurance enrollment; therefore, data availability relies on the presence of longitudinal EHR and survey records.

All analyses were performed within the secure, cloud-based *All of Us* Researcher Workbench, a platform that provides registered researchers with access to de-identified, individual-level participant data and a suite of analytical tools. The use of de-identified data ensures the protection of participant privacy and confidentiality in accordance with the program's security protocols. The study was conducted in compliance with the *All of Us* Research Program's Data and Research Use Policy. The study was reviewed by the Institutional Review Board (IRB) of Hanyang University Seoul Hospital, and was deemed exempt from IRB approval (IRB No.: HYUH-2025-10-005). The IRB granted a waiver of informed consent because the study used previously collected, de-identified secondary data from the *All of Us* Research Program and presented no more than minimal risk to participants. Participant confidentiality was preserved through the *All of Us* Researcher Workbench security protocols, and the study was conducted in accordance with the principles of the Declaration of Helsinki.

Study Design

This study was conducted as a retrospective, new-user, active-comparator cohort study. Eligible participants were identified from the *All of Us* database by applying the following inclusion and exclusion criteria sequentially. An initial cohort of patients with at least one prescription record for one of the five FDA-approved GLP-1RAs (exenatide, liraglutide, dulaglutide, semaglutide, or tirzepatide) was identified from the *All of Us* database, and the date of the first prescription was defined as the index date. Specifically, this is the earliest record for any of the five GLP-1RAs in the *All of Us* OMOP *Drug Exposure* table, which captures EHR-recorded prescription orders. To create a cohort of new users, patients were excluded if they had a prescription for any GLP-1RA within the 365-day period preceding the index date, minimizing prevalent-user bias and reducing contamination from earlier short-term exposures. To ensure active engagement with the healthcare system and valid capture of baseline data, eligible participants were required to have at least one recorded medical encounter within the 365 days preceding the index date. The study cohort was further refined to establish a population of new users with obesity. Patients were excluded if they met any of the following conditions: prescription for more than one type of GLP-1RA on the index date; age less than 18 years or greater than 75 years at the index date; an index date before January 1, 2006 (the earliest period of GLP-1RA availability captured in the data), or after September 30, 2023; or no evidence of obesity, defined as either a diagnosis code for obesity or a recorded BMI of ≥ 30 kg/m² during the 365-day baseline period. Selected patients were followed up for weight and BMI change from the index date to the first occurrence of study outcome, death, loss to follow-up (defined by the last recorded measurement), or the administrative study end. A primary intention-to-treat (ITT) analysis design was adopted, analyzing all patients according to their first assigned GLP-1RA regardless of subsequent therapy changes, while sensitivity analysis explored as-treated scenarios. We adopted intention-to-treat as the primary analysis to align with the new-user, active-comparator design and to mitigate informative censoring. Discontinuation and switching are themselves influenced by treatment effects, tolerability, and preference; we therefore prespecified the as-treated analysis as a sensitivity analysis to characterize on-treatment effects.

Exposures

Patients meeting all eligibility criteria were assigned to one of five mutually exclusive exposure cohorts based on their index GLP-1RA prescription. Tirzepatide is pharmacologically distinct from the other four agents as a dual GIP/GLP-1 receptor co-agonist, whereas exenatide, liraglutide, dulaglutide, and semaglutide are mono-agonists at the GLP-1 receptor. For each of the five GLP-1RAs, exposure was operationalized as any *Drug Exposure* record whose concept ID was a descendant of the corresponding parent ingredient concept in the OMOP standardized vocabulary (parent concept IDs listed in Table S1). For all comparative effectiveness analyses, liraglutide was selected as the active reference group. Liraglutide was the first GLP-1RA to receive an FDA indication specifically for chronic weight management and has been widely used in clinical practice for this purpose.¹⁸ It exhibits an intermediate efficacy profile for weight loss within the GLP-1RA class, making it a suitable and clinically relevant benchmark against which both older and newer agents can be compared.

Outcomes

The study assessed four distinct, co-primary time-to-event outcomes. These were defined as the time from the index date to the first recorded achievement of the following clinically significant thresholds: 1) $\geq 5\%$ weight reduction from baseline, 2) $\geq 10\%$ weight reduction from baseline, 3) $\geq 15\%$ weight reduction from baseline, and 4) reduction of at least one BMI category (eg, from Obesity Class II to Obesity Class I).¹⁹ Baseline weight and BMI were defined as the most recent valid measurement recorded within 90 days prior to or on the index date. The $\geq 5\%$, $\geq 10\%$, and $\geq 15\%$ thresholds were prespecified to reflect a graded, dose-dependent relationship between weight loss and clinical benefit: a $\geq 5\%$ reduction is the conventional threshold for clinically meaningful improvement in cardiometabolic risk, while $\geq 10\%$ and $\geq 15\%$ confer progressively greater benefit and are often required for improvement in specific weight-related complications.^{11,20}

The secondary outcomes quantified the magnitude of weight loss, defined as the percent change in body weight and BMI from baseline to 12 months. To account for the variability in real-world data capture, the 12-month value was defined as the measurement closest to day 365 within a window of ± 45 days. When multiple measurements were recorded on the same day, the daily mean was used.

For the time-to-event outcomes, an event was recorded at the first qualifying measurement that crossed the relevant threshold relative to baseline; loss to follow-up was defined by the last available weight (or BMI) measurement for the outcome being modeled. BMI categories for the BMI-reduction outcome were defined per the CDC standard (Underweight <18.5 ; Normal $18.5\text{--}24.9$; Overweight $25.0\text{--}29.9$; Class I obesity $30.0\text{--}34.9$; Class II $35.0\text{--}39.9$; Class III ≥ 40 kg/m²).

Statistical Analysis

All statistical analyses were performed using Python (version 3.10.16) within the secure *All of Us* Researcher Workbench (Registered Tier Dataset v8). The primary libraries used for data manipulation, analysis, and visualization included pandas (version 2.0.3), NumPy (version 1.24.4), SciPy (version 1.11.4), lifelines (version 0.30.0) for survival analysis, statsmodels (version 0.14.4) for linear regression, Matplotlib (version 3.7.3), and Seaborn (version 0.12.2). A two-sided $P < 0.05$ was considered statistically significant for all analyses.

Baseline characteristics of the study population were summarized and stratified by the index GLP-1RA. Continuous variables were presented as means and standard deviations (SD), while categorical variables were presented as frequencies and percentages. To protect participant privacy and in accordance with the *All of Us* Research Program's Data and Statistics Dissemination Policy, participant counts representing fewer than 20 individuals were suppressed, and categories with small subgroups were collapsed where necessary. Standardized mean differences (SMDs) were computed for all baseline covariates across exposure cohorts to quantify imbalance, with $|\text{SMD}| > 0.10$ considered indicative of meaningful imbalance. To minimize bias from missing baseline data, we employed Multiple Imputation by Chained Equations (MICE). Five imputed datasets were generated under the missing-at-random assumption, with each incomplete variable imputed from all other variables by Bayesian ridge regression, drawing each imputation from the posterior predictive distribution; categorical variables were ordinal-encoded, imputed on that scale, and mapped back to the nearest category. The imputation model included the index GLP-1RA, age, sex, race and ethnicity, annual income, education, smoking status, alcohol use, baseline body mass index, hemoglobin A1c, the number of outpatient visits in the prior year, the Charlson Comorbidity Index, type 2 diabetes and other comorbidities (atherosclerotic cardiovascular disease, heart failure, chronic kidney disease, hypertension, and dyslipidemia), and baseline glucose-lowering and anti-obesity medications; and estimates from the five datasets were combined using Rubin's Rules to derive valid standard errors and confidence intervals.

The cumulative incidence of achieving each of the four outcomes was visualized using Kaplan-Meier curves for each treatment group. To compare the GLP-1RAs on these outcomes, we employed multivariable Cox proportional hazards models to estimate both unadjusted hazard ratios (HRs) and adjusted hazard ratios (aHRs) with their corresponding 95% confidence intervals (CIs). The proportional-hazards assumption was assessed for each Cox model using scaled Schoenfeld residuals on a per-covariate basis and combined globally using a Cauchy-combined p-

value across all covariates. Because the Schoenfeld test is known to be overpowered in large samples and to reject the null even for practically negligible departures from proportionality, we also visualized smoothed scaled Schoenfeld residuals for the drug-exposure terms over time to assess the magnitude of any time-varying effects, and computed restricted mean survival time (RMST) differences over 365 days as a PH-free effect estimate. To formally test whether the magnitude of GLP-1RA-associated weight change differs by type 2 diabetes status, we additionally fit the main Cox proportional hazards and combined linear-regression models with a drug $\times$ T2DM interaction term, and assessed the joint significance of the four interaction coefficients using a Wald χ^2 -test (df = 4). Liraglutide served as the reference group in all models. The adjusted models included the following baseline covariates: age, sex, race and ethnicity, socioeconomic status (annual income and education level), smoking status, alcohol consumption, BMI, baseline hemoglobin A1c, the number of outpatient visits in the prior year, and the Charlson Comorbidity Index (CCI) score. Clinical history covariates included type 2 diabetes mellitus (T2DM), atherosclerotic cardiovascular disease (ASCVD), heart failure, chronic kidney disease (CKD), hypertension, and dyslipidemia. Concomitant baseline use of metformin, SGLT2 inhibitors, sulfonylureas, thiazolidinediones, DPP-4 inhibitors, insulin, orlistat, and phentermine/topiramate was also included.

To model the magnitude of weight and BMI change at 1 year, we constructed separate linear regression models for the percent change from baseline in body weight and BMI at 12 months. These models compared the adjusted mean percent change for each GLP-1RA to the liraglutide reference group and were adjusted for the same set of baseline covariates used in the Cox proportional hazards models. Percent changes in weight and BMI at 1 year were analyzed on a complete-case basis, restricted to participants with a valid measurement within the day-365 $\pm$ 45-day window; robustness to informative dropout was assessed using inverse-probability-of-censoring weighting and tipping-point analyses.

A subgroup analysis was performed by stratifying the entire cohort based on the presence or absence of a T2DM diagnosis at baseline. The primary and secondary analyses were repeated within each subgroup. To assess the robustness of our findings, we conducted several sensitivity analyses: (1) an as-treated analysis, censoring patients at the time they switched to or added a different GLP-1RA or discontinued treatment (defined as a gap of >60 days between consecutive prescription orders); (2) restricting the cohort to patients with baseline BMI ≥ 35 kg/m² and ≥ 40 kg/m²; (3) repeating the primary analysis using exenatide and dulaglutide as the reference comparators; and (4) adjusting the primary Cox proportional-hazards models for the index calendar year as a continuous covariate to address secular trends in prescribing and outcome ascertainment ([Table S12](#)).

Results

Patient Characteristics

From an initial population of 21,089 patients, 14,046 eligible new users with obesity were identified ([Figure 1](#)). Semaglutide was the most frequently initiated agent (40.3%), followed by dulaglutide (26.4%), liraglutide (24.9%), exenatide (5.7%), and tirzepatide (2.7%).

Baseline characteristics are presented in [Table 1](#). The cohort (mean age 53.9 ± 12.0 years; 71.4% female) was diverse, including 22.0% Non-Hispanic Black and 16.1% Hispanic participants. Mean baseline BMI was 39.3 ± 8.0 kg/m². Significant differences were observed across cohorts; patients initiating exenatide and dulaglutide had a substantially higher prevalence of T2DM (>84%) and higher baseline HbA1c levels compared to those initiating semaglutide and tirzepatide (T2DM prevalence ~51%). Users of newer agents also tended to be younger with a lower Charlson Comorbidity Index score. Standardized mean differences (|SMD|) versus the liraglutide reference are reported in [Table S2](#). Per-arm completeness for the body-weight measurement at the day-365 ($\pm$ 45 days) analytic window for the secondary outcomes was 47.2% (375/794) for exenatide, 40.0% (1,400/3,501) for liraglutide, 33.9% (1,256/3,709) for dulaglutide, 26.1% (1,476/5,656) for semaglutide, and was suppressed per the *All of Us* Data and Statistics Dissemination Policy for tirzepatide (n < 20 / 386). Completeness for the BMI measurement was similar ([Table S3](#)).

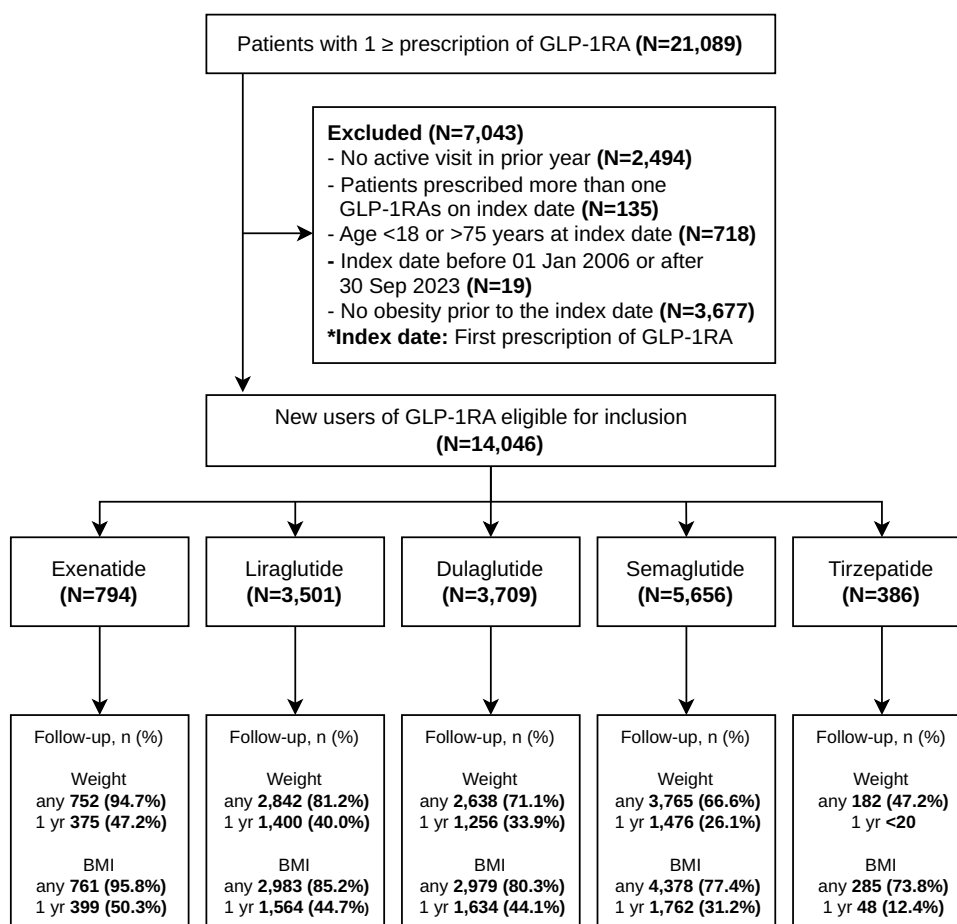


Figure 1 Flow diagram of study participant selection from the All of Us Research Program cohort. Sequential inclusion and exclusion criteria are applied from the initial pool of GLP-1 receptor agonist users to the final analytic cohort of 14,046 new users with obesity, shown by index agent. The final tier reports follow-up data availability for the secondary outcomes, showing, for each exposure group, the number of participants contributing any post-index weight measurement and the number with a measurement within the day-365 (± 45 days) analytic window; per-arm counts for both weight and body mass index are tabulated in [Table S3](#).

GLP-1RA Initiation Trends Over Time

Prescribing patterns shifted substantially over the study period ([Figure S1](#)). New initiations were sparse before 2014; thereafter, while exenatide and liraglutide use declined, semaglutide initiations rose rapidly, accounting for over 40% of new prescriptions by 2022. Tirzepatide demonstrated immediate uptake, comprising over 20% of new initiations in 2023.

Achievement of Weight Loss and BMI Category Reduction

The comparative effectiveness of the five GLP-1RAs was assessed using time-to-event analyses. The cumulative incidence of achieving $\geq 5\%$, $\geq 10\%$, and $\geq 15\%$ weight loss was highest and occurred most rapidly for patients initiating tirzepatide, followed closely by semaglutide ([Figure 2](#)). The curves for liraglutide, dulaglutide, and exenatide were clustered together, indicating substantially lower and slower rates of achieving these clinically significant weight loss milestones. For each co-primary outcome, the overall (k-sample) Log rank test comparing the five exposure strata was significant (χ^2 [4 df] = 247.5, 249.1, 189.1, and 224.6 for $\geq 5\%$, $\geq 10\%$, $\geq 15\%$ weight reduction and ≥ 1 BMI-category reduction, respectively; all $P < 0.001$). In unadjusted pairwise Log rank tests with Benjamini–Hochberg correction, tirzepatide and semaglutide each differed from every other agent for all four outcomes (all BH-adjusted $P < 0.001$), whereas liraglutide and dulaglutide did not differ for any weight-reduction threshold (BH-adjusted P 0.57–0.87; [Tables S4](#) and [S5](#)). At day 365, the Kaplan-Meier cumulative incidence of $\geq 15\%$ weight loss was 7.7% (95% CI 6.6–9.0%) with liraglutide, 7.1% (6.0–8.4%) with dulaglutide, 6.1% (4.3–8.6%) with exenatide, 15.0% (13.6–16.6%) with semaglutide,

Table I Baseline Demographic and Clinical Characteristics of GLP-I Receptor Agonist New Users

Characteristics	Overall (N=14046)	Exenatide (N=794)	Liraglutide (N=3501)	Dulaglutide (N=3709)	Semaglutide (N=5656)	Tirzepatide (N=386)
Age, years	53.9 (± 12.0)	53.7 (± 10.7)	52.6 (± 11.8)	55.5 (± 11.5)	53.8 (± 12.5)	52.6 (± 11.8)
Sex						
Female	10024 (71.4%)	536 (67.5%)	2498 (71.4%)	2494 (67.2%)	4206 (74.4%)	290 (75.1%)
Male or other sex	4022 (28.6%)	258 (32.5%)	1003 (28.6%)	1215 (32.8%)	1450 (25.6%)	96 (24.9%)
Race and ethnicity						
Non-Hispanic White	7359 (52.4%)	398 (50.1%)	1898 (54.2%)	1670 (45.0%)	3123 (55.2%)	270 (69.9%)
Non-Hispanic Black	3091 (22.0%)	169 (21.3%)	719 (20.5%)	1007 (27.2%)	1144 (20.2%)	52 (13.5%)
Hispanic	2267 (16.1%)	151 (19.0%)	553 (15.8%)	676 (18.2%)	857 (15.2%)	30 (7.8%)
Other	1329 (9.5%)	76 (9.6%)	331 (9.5%)	356 (9.6%)	532 (9.4%)	34 (8.8%)
Education: any college on record	4654 (33.1%)	271 (34.1%)	1108 (31.6%)	1242 (33.5%)	1899 (33.6%)	134 (34.7%)
Annual income, \$						
0–24,999	3780 (26.9%)	257 (32.4%)	923 (26.4%)	1303 (35.1%)	1247 (22.0%)	50 (13.0%)
25,000–49,999	2411 (17.2%)	124 (15.6%)	565 (16.1%)	637 (17.2%)	1006 (17.8%)	79 (20.5%)
50,000–99,999	2981 (21.2%)	160 (20.2%)	741 (21.2%)	642 (17.3%)	1332 (23.6%)	106 (27.5%)
Over 100,000	2612 (18.6%)	88 (11.1%)	701 (20.0%)	444 (12.0%)	1262 (22.3%)	117 (30.3%)
Unknown	2262 (16.1%)	165 (20.8%)	571 (16.3%)	683 (18.4%)	809 (14.3%)	34 (8.8%)
Smoking						
Current	1568 (11.2%)	88 (11.1%)	366 (10.5%)	539 (14.5%)	546 (9.7%)	29 (7.5%)
Never or unknown	8524 (60.7%)	457 (57.6%)	2117 (60.5%)	2183 (58.9%)	3531 (62.4%)	236 (61.1%)
Past	3954 (28.2%)	249 (31.4%)	1018 (29.1%)	987 (26.6%)	1579 (27.9%)	121 (31.3%)
Drinking						
Current	9540 (67.9%)	452 (56.9%)	2366 (67.6%)	2324 (62.7%)	4100 (72.5%)	298 (77.2%)
Never or unknown	1677 (11.9%)	118 (14.9%)	427 (12.2%)	549 (14.8%)	553 (9.8%)	30 (7.8%)
Past	2829 (20.1%)	224 (28.2%)	708 (20.2%)	836 (22.5%)	1003 (17.7%)	58 (15.0%)
Body mass index, kg/m²	39.3 (± 8.0)	39.9 (± 7.9)	39.9 (± 8.3)	39.0 (± 8.0)	39.2 (± 7.9)	39.3 (± 7.9)
Weight, kg	110.2 (± 25.8)	113.1 (± 25.9)	112.3 (± 27.6)	110.0 (± 25.8)	108.9 (± 24.5)	110.3 (± 25.7)
Height, cm	167.2 (± 9.7)	167.5 (± 10.1)	167.2 (± 9.8)	167.6 (± 10.1)	166.9 (± 9.4)	167.1 (± 9.0)
Waist Circumference, cm	116.4 (± 16.1)	118.4 (± 14.9)	115.5 (± 16.3)	117.4 (± 15.6)	115.7 (± 16.2)	120.4 (± 18.7)
Systolic blood pressure, mmHg	128.8 (± 16.3)	130.1 (± 16.3)	128.6 (± 16.3)	129.7 (± 17.1)	128.2 (± 15.9)	126.8 (± 15.1)
Diastolic blood pressure, mmHg	77.8 (± 10.1)	76.7 (± 10.2)	77.6 (± 10.1)	77.7 (± 10.5)	78.2 (± 9.9)	78.0 (± 9.1)
Comorbidities						
Type 2 diabetes mellitus	9167 (65.3%)	712 (89.7%)	2219 (63.4%)	3143 (84.7%)	2893 (51.1%)	200 (51.8%)
Hypertension	10413 (74.1%)	639 (80.5%)	2527 (72.2%)	3052 (82.3%)	3939 (69.6%)	256 (66.3%)
Dyslipidemia	9789 (69.7%)	598 (75.3%)	2341 (66.9%)	2814 (75.9%)	3778 (66.8%)	258 (66.8%)
ASCVD	1867 (13.3%)	103 (13.0%)	443 (12.7%)	625 (16.9%)	668 (11.8%)	28 (7.3%)
Chronic Kidney Disease	2152 (15.3%)	95 (12.0%)	489 (14.0%)	760 (20.5%)	770 (13.6%)	38 (9.8%)
Heart Failure	1900 (13.5%)	89 (11.2%)	486 (13.9%)	648 (17.5%)	650 (11.5%)	27 (7.0%)
Charlson Comorbidity Index	3.8 (± 3.2)	3.9 (± 2.8)	3.6 (± 3.1)	4.8 (± 3.3)	3.3 (± 3.2)	2.9 (± 2.9)

(Continued)

Table I (Continued).

Characteristics	Overall (N=14046)	Exenatide (N=794)	Liraglutide (N=3501)	Dulaglutide (N=3709)	Semaglutide (N=5656)	Tirzepatide (N=386)
Comedications						
Metformin	8337 (59.4%)	568 (71.5%)	2018 (57.6%)	2685 (72.4%)	2861 (50.6%)	205 (53.1%)
SGLT2 Inhibitor	1572 (11.2%)	51 (6.4%)	239 (6.8%)	564 (15.2%)	667 (11.8%)	51 (13.2%)
Sulfonylurea	3451 (24.6%)	344 (43.3%)	907 (25.9%)	1256 (33.9%)	894 (15.8%)	50 (13.0%)
Thiazolidinedione	—	159 (20.0%)	420 (12.0%)	465 (12.5%)	217 (3.8%)	<20
DPP-4 Inhibitor	1804 (12.8%)	137 (17.3%)	437 (12.5%)	635 (17.1%)	560 (9.9%)	35 (9.1%)
Insulin	5665 (40.3%)	388 (48.9%)	1440 (41.1%)	2084 (56.2%)	1658 (29.3%)	95 (24.6%)
Orlistat	212 (1.5%)	<20	64 (1.8%)	46 (1.2%)	85 (1.5%)	<20
Phentermine/Topiramate	149 (1.1%)	<20	51 (1.5%)	22 (0.6%)	73 (1.3%)	<20
Clinical laboratory tests						
Hemoglobin A1c, %	7.5 (± 2.1)	8.6 (± 1.9)	7.4 (± 2.1)	8.2 (± 2.1)	7.0 (± 1.8)	6.9 (± 2.0)
Glucose, mg/dL	152.0 (± 80.7)	174.3 (± 83.3)	151.3 (± 82.6)	177.3 (± 88.6)	134.3 (± 68.7)	133.6 (± 69.9)
eGFR, mL/min/1.73m ²	85.1 (± 22.8)	88.5 (± 22.4)	86.4 (± 22.8)	82.4 (± 24.4)	85.3 (± 21.6)	87.2 (± 21.5)
Creatinine, mg/dL	1.0 (± 0.6)	0.9 (± 0.3)	1.0 (± 0.7)	1.0 (± 0.7)	0.9 (± 0.5)	1.0 (± 0.9)
Total Cholesterol, mg/dL	173.9 (± 44.9)	171.9 (± 40.6)	173.5 (± 44.9)	169.4 (± 47.1)	176.9 (± 43.9)	178.9 (± 43.9)
HDL Cholesterol, mg/dL	47.2 (± 14.2)	44.0 (± 13.6)	47.3 (± 15.0)	44.9 (± 13.4)	48.8 (± 14.0)	49.0 (± 13.7)
LDL Cholesterol, mg/dL	95.6 (± 37.8)	92.8 (± 34.8)	94.1 (± 36.6)	89.4 (± 41.3)	99.7 (± 36.7)	101.7 (± 32.9)
Triglycerides, mg/dL	173.8 (± 135.5)	191.1 (± 120.4)	170.0 (± 132.8)	193.9 (± 165.9)	161.5 (± 110.3)	173.1 (± 186.9)
ALT, U/L	29.3 (± 26.6)	32.5 (± 24.5)	29.2 (± 24.9)	30.1 (± 30.6)	28.5 (± 25.8)	26.7 (± 18.4)
AST, U/L	24.8 (± 20.0)	27.3 (± 22.2)	24.6 (± 17.1)	25.2 (± 21.5)	24.5 (± 20.9)	21.7 (± 11.8)
Alkaline Phosphatase, U/L	86.7 (± 33.2)	89.3 (± 38.0)	85.9 (± 32.7)	92.1 (± 38.6)	83.9 (± 28.9)	82.6 (± 27.6)
Total Bilirubin, mg/dL	0.5 (± 0.3)	0.5 (± 0.3)	0.5 (± 0.3)	0.5 (± 0.3)	0.5 (± 0.3)	0.5 (± 0.3)
Platelets, ×10 ³ /μL	264.5 (± 75.9)	261.8 (± 78.0)	264.1 (± 75.7)	259.5 (± 77.5)	267.1 (± 74.5)	280.2 (± 77.1)
Outpatient visits in prior year	23.5 (± 37.9)	22.5 (± 33.2)	29.2 (± 44.4)	25.7 (± 42.6)	19.2 (± 30.5)	15.0 (± 19.7)

Notes: Cells with fewer than 20 participants are suppressed to protect patient confidentiality. The overall total for the Thiazolidinedione row has additionally been withheld (—) to prevent back-calculation of the suppressed count.

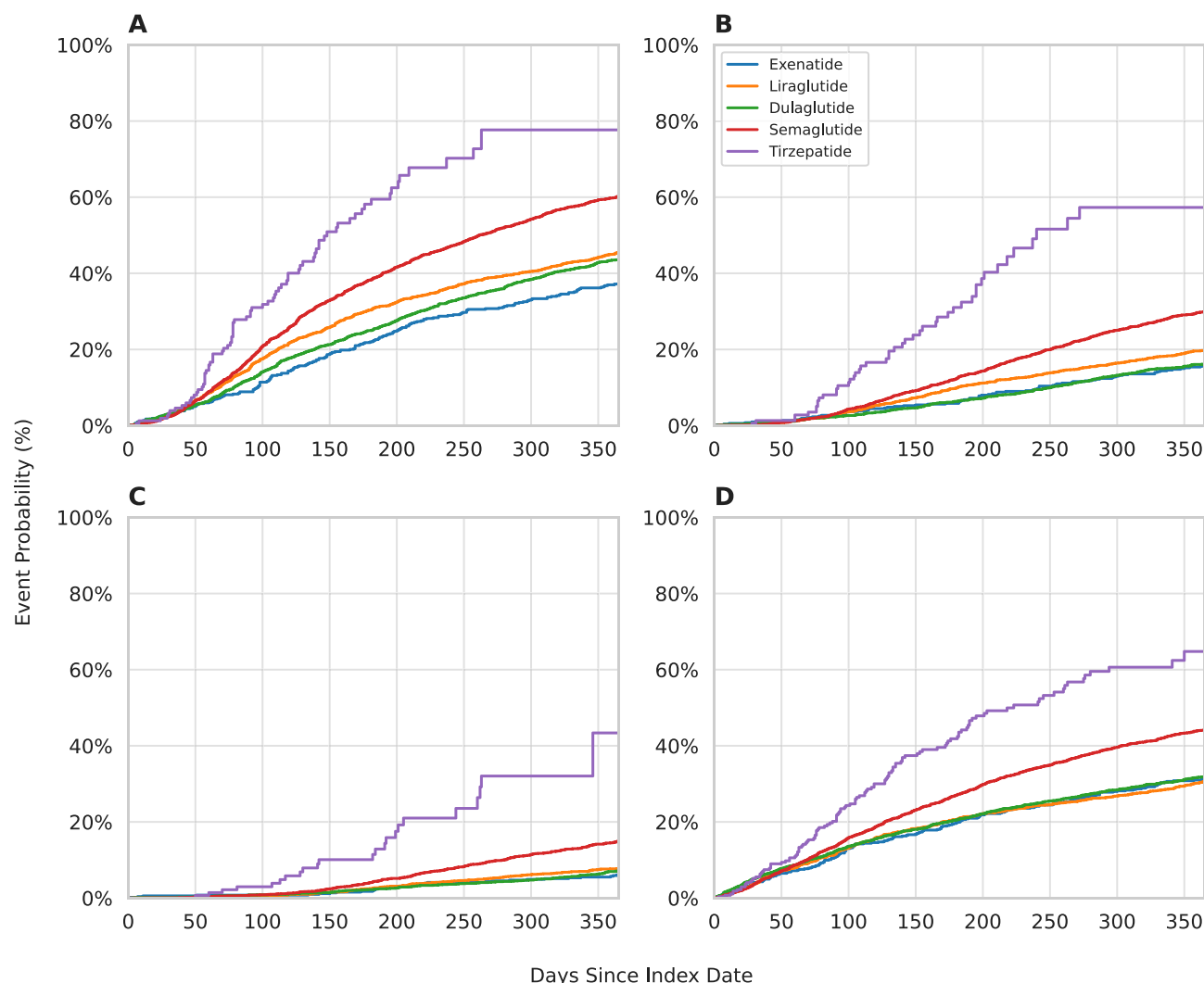


Figure 2 Kaplan-Meier curves for the cumulative incidence of achieving (A) $\geq 5\%$ weight reduction, (B) $\geq 10\%$ weight reduction, (C) $\geq 15\%$ weight reduction, and (D) reduction of at least one BMI category from baseline.

and 43.4% (24.4–68.6%) with tirzepatide. Corresponding 12-month cumulative incidences for all four outcomes are tabulated in [Table S6](#).

After adjustment for baseline covariates, the between-agent differences persisted and widened with increasing stringency of the weight-loss threshold ([Figure 3](#)). For achieving $\geq 5\%$ weight loss, both tirzepatide (aHR 2.53; 95% CI, 2.01–3.18) and semaglutide (aHR 1.44; 95% CI, 1.34–1.55) were associated with a significantly higher likelihood of $\geq 5\%$ weight loss than liraglutide; dulaglutide showed a modest increase of borderline significance (aHR 1.09; 95% CI, 1.00–1.17), and exenatide did not differ (aHR 0.91; 95% CI, 0.81–1.02). For $\geq 10\%$ weight loss, tirzepatide was more than four times as likely to achieve the threshold (aHR 4.39; 95% CI, 3.26–5.91) and semaglutide remained associated with a significantly higher likelihood (aHR 1.57; 95% CI, 1.43–1.73). The difference was most pronounced for $\geq 15\%$ weight loss, where tirzepatide showed a more than five-fold higher likelihood than liraglutide (aHR 5.57; 95% CI, 3.66–8.49) and semaglutide remained associated with a significantly higher likelihood (aHR 1.77; 95% CI, 1.56–2.01).

A similar pattern was observed for achieving a reduction in BMI category, with tirzepatide (aHR 2.58; 95% CI, 2.14–3.11) and semaglutide (aHR 1.45; 95% CI, 1.35–1.57) associated with a higher adjusted likelihood of BMI-category reduction than liraglutide.

Schoenfeld tests detected statistically significant departures from proportional hazards in each Cox model ([Table S7](#)), most prominently for baseline BMI. Smoothed residual plots for the drug-exposure terms ([Figure S2](#)) showed small-

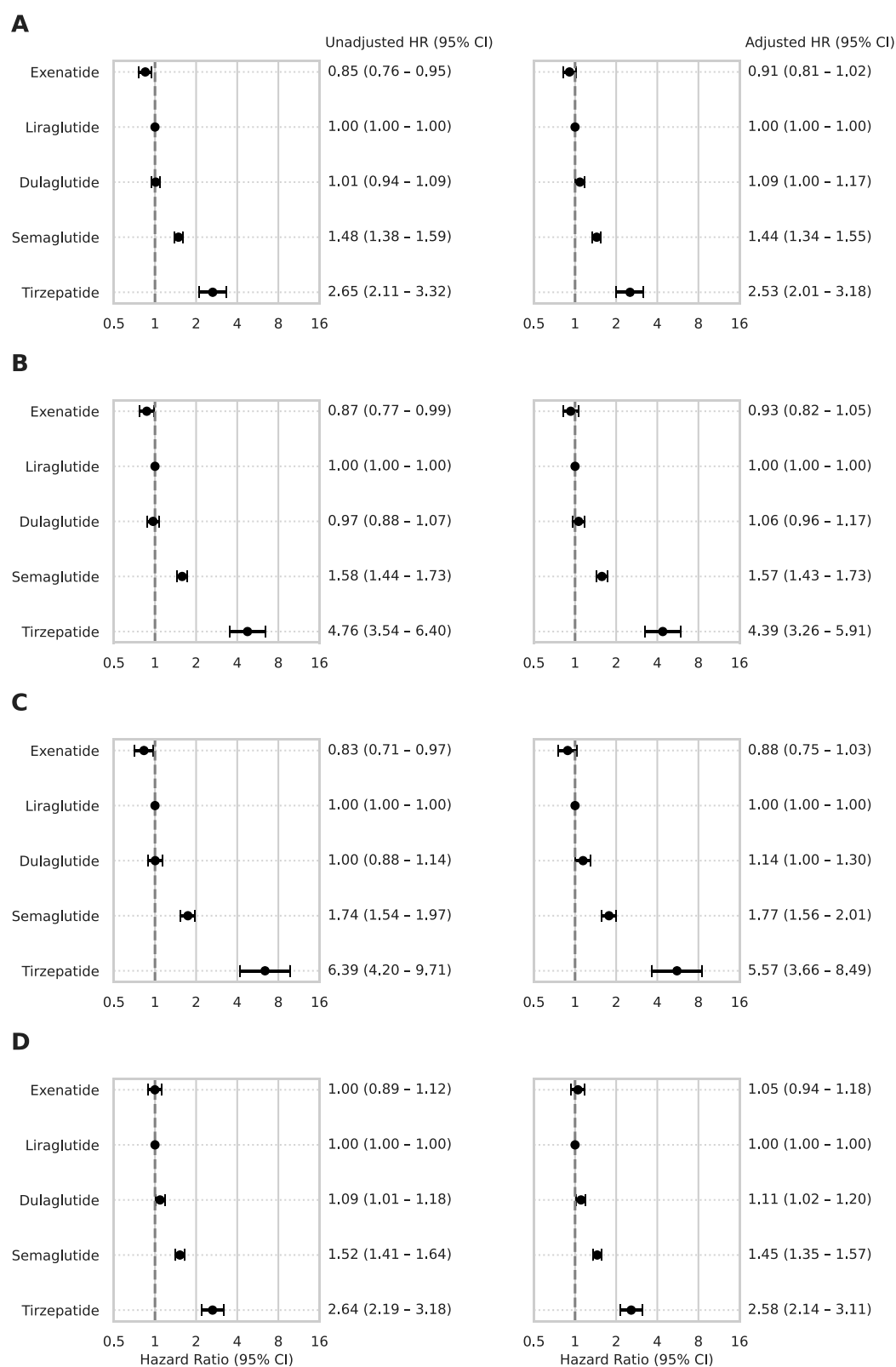


Figure 3 Unadjusted and adjusted hazard ratios for achieving (A) $\geq 5\%$ weight reduction, (B) $\geq 10\%$ weight reduction, (C) $\geq 15\%$ weight reduction, and (D) reduction of at least one BMI category from baseline.

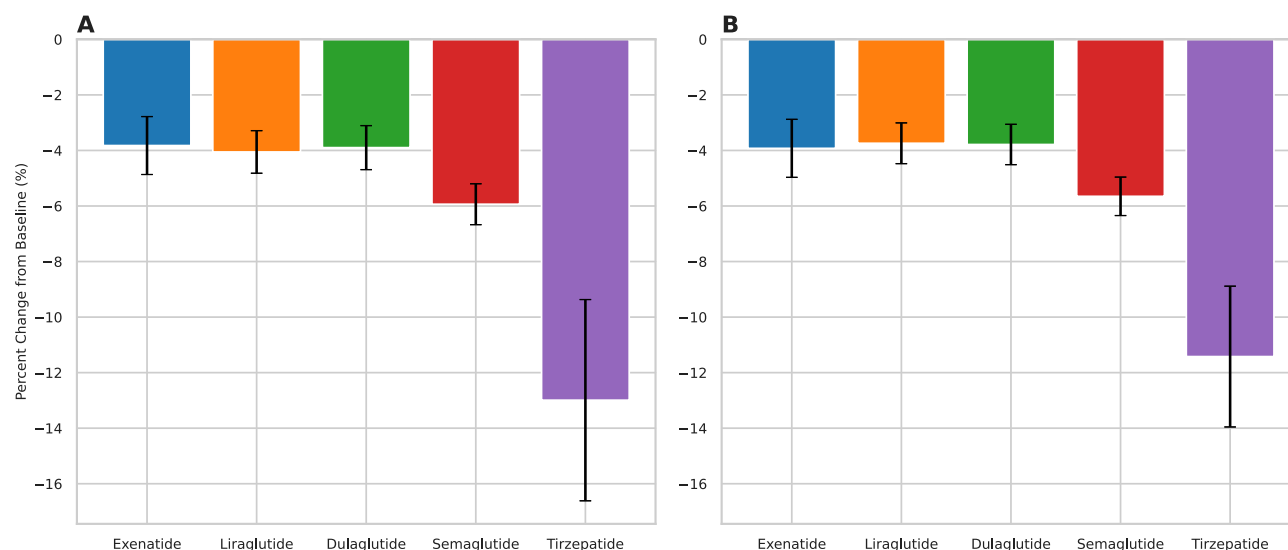


Figure 4 Adjusted mean percent change from baseline in (A) body weight and (B) body mass index at 12 months.

magnitude departures, and covariate-balanced (IPTW-adjusted) restricted mean survival time (RMST) at 365 days (Table S8) reproduced the adjusted-Cox ranking: tirzepatide reached each threshold 31 to 77 days earlier than liraglutide ($\geq 15\%$ weight loss: -31.3 days, 95% CI -53.9 to -13.3), followed by semaglutide (-7.5 to -23.2 days; all $P < 0.001$). After covariate balancing, exenatide and dulaglutide differed little from liraglutide, consistent with their near-null adjusted hazard ratios.

Across these time-to-event analyses, the tirzepatide estimates should be interpreted with caution: they derive from a small cohort ($n=386$) with limited available follow-up and are therefore sensitive to baseline imbalance (Table S2) and residual confounding.

Percent Change in Body Weight and BMI

Adjusted mean percent weight loss at 12 months was largest with tirzepatide (-13.0% ; 95% CI, -16.6% to -9.4%) and semaglutide (-5.9% ; 95% CI, -6.7% to -5.2%), with comparable modest reductions for exenatide, liraglutide, and dulaglutide (approximately -4.0%) (Figure 4). To formally compare agents in a single model, we fit a combined linear-regression model with drug as a five-level categorical exposure (Table S9). Relative to liraglutide, both tirzepatide (adjusted mean difference, -8.55% ; 95% CI, -12.16 to -4.94) and semaglutide (-1.84% ; 95% CI, -2.46 to -1.23) showed significantly greater weight loss (both BH-adjusted $P < 0.001$), whereas exenatide (0.15% ; 95% CI, -0.80 to 1.09) and dulaglutide (0.10% ; 95% CI, -0.55 to 0.75) did not differ. An analogous pattern held for percent BMI change, with tirzepatide (-7.45% ; 95% CI, -9.96 to -4.93) and semaglutide (-1.88% ; 95% CI, -2.49 to -1.28) again significantly lower than liraglutide (both BH-adjusted $P < 0.001$). After Benjamini–Hochberg correction, tirzepatide also showed significantly greater percent weight loss than every other agent (eg, vs semaglutide, -6.71% ; 95% CI, -10.31 to -3.10), semaglutide exceeded each of the three older agents (differences $\approx -1.8\%$ to -2.0% ; all BH-adjusted $P < 0.001$), and the older agents did not differ significantly from one another (all BH-adjusted $P \geq 0.85$).

Subgroup Analysis by Type 2 Diabetes Status

Across diabetes strata, the relative ordering of agents was preserved, but the speed of reaching the weight-loss thresholds differed by diabetes status, consistent with the well-recognized attenuation of GLP-1RA-associated weight loss in adults with type 2 diabetes. To examine effect modification by type 2 diabetes status, we stratified the cohort, refitted the analyses (Figures S3–S8), and tested formal drug $\times$ T2DM interaction terms (Table S10). In the main Cox models, the interaction was statistically significant for the $\geq 5\%$, $\geq 10\%$, and $\geq 15\%$ weight-reduction outcomes (joint Wald $\chi^2 = 15.84$, 14.40, and 12.18 on 4 df; $P = 0.003$, 0.006, and 0.016, respectively) but not for ≥ 1 BMI-category reduction ($\chi^2 = 3.03$;

$P = 0.553$). In the combined linear-regression models, the interaction did not reach significance for either percent weight change (Wald $\chi^2 = 5.45$; $P = 0.244$) or percent BMI change (Wald $\chi^2 = 5.98$; $P = 0.201$). Among patients with T2DM ($N = 9,167$), tirzepatide was associated with the highest adjusted likelihood of $\geq 15\%$ weight loss (aHR 5.73; 95% CI, 2.91–11.27), followed by semaglutide (aHR 1.62; 95% CI, 1.38–1.91); among those without T2DM ($N = 4,879$), the corresponding aHRs were 5.73 (95% CI, 3.30–9.93) for tirzepatide and 1.94 (95% CI, 1.57–2.39) for semaglutide. The attenuation by diabetes was clearest for semaglutide (adjusted mean weight loss -5.1% [95% CI, -5.9 to -4.2] with vs -7.6% [95% CI, -9.2 to -6.0] without T2DM) and was also evident for liraglutide (-3.5% [95% CI, -4.4 to -2.6] with vs -5.2% [95% CI, -6.9 to -3.5] without T2DM), whereas tirzepatide was essentially unaffected (-13.6% [95% CI, -18.4 to -8.7] vs -13.5% [95% CI, -19.4 to -7.7]).

Sensitivity Analyses

The results of the sensitivity analyses were consistent with the primary findings. In the as-treated analysis, the association with the newer agents was even stronger; the adjusted likelihood of achieving $\geq 15\%$ weight loss with tirzepatide rose to an aHR of 6.64 (95% CI, 4.02–10.97) compared to liraglutide. The relative ordering of the agents remained robust when restricting the cohort to higher baseline BMI categories (≥ 35 kg/m² and ≥ 40 kg/m²) and when using exenatide or dulaglutide as the active comparator (Table S11). Adjustment for the index calendar year as a continuous covariate attenuated the tirzepatide and semaglutide aHRs toward the null but did not alter the principal findings: both remained significantly more likely than liraglutide to achieve all four co-primary outcomes, with the tirzepatide > semaglutide ordering preserved (Table S12). Inverse-probability-of-censoring-weighted (IPCW) re-estimation of the secondary outcome yielded adjusted mean differences vs liraglutide of -7.71% (95% CI -9.25 to -6.18) for tirzepatide and -1.94% (95% CI -2.59 to -1.30) for semaglutide (Table S13), and a tipping-point analysis in which non-completers' counterfactual percent weight change was shifted by Δ from -5% to $+10\%$ confirmed that the tirzepatide and semaglutide advantages remained statistically significant across the full grid (Table S14).

Discussion

In this large-scale, real-world cohort study of 14,046 adults with obesity drawn from the diverse *All of Us* Research Program, we observed a consistent ordering in the magnitude of weight loss of five commonly prescribed GLP-1RAs. After adjustment for baseline demographic and clinical characteristics, the dual GIP/GLP-1 receptor agonist tirzepatide and semaglutide were associated with a higher relative likelihood over time of achieving clinically meaningful weight loss and with a greater reduction in both body weight and BMI compared to the older agents exenatide, liraglutide, and dulaglutide. The magnitude of weight loss observed for tirzepatide was the largest, particularly for achieving high-threshold weight loss of 15% or more, though this estimate is imprecise given the small tirzepatide cohort (see Limitations). This level of weight reduction is associated with critical health benefits such as reductions in all-cause and cardiovascular mortality, improvements in inflammatory markers, and significant reversal of hepatic steatosis.²⁰

Our findings from this real-world setting are consistent with the hierarchy of efficacy established in head-to-head randomized clinical trials (RCTs). The greater weight loss we observed with semaglutide than liraglutide aligns with the STEP 8 trial, which directly compared these two agents for weight management in individuals with obesity.²¹ Similarly, the larger weight loss observed in the tirzepatide group than the semaglutide group mirrors the SURMOUNT-5 trial and is supported by real-world evidence studies that have consistently shown greater weight loss with the dual-agonist therapy.^{22,23} The comparable and modest weight loss effects observed for exenatide, liraglutide and dulaglutide in our study are also in line with previous evidence from clinical trials such as LEAD-6 and AWARD-6.^{24,25} Furthermore, a comparison of our findings in the non-T2DM subgroup with the results of pivotal registration trials reveals important insights into the translation of efficacy into real-world outcomes. In our study, patients without type 2 diabetes who initiated liraglutide, semaglutide, and tirzepatide achieved an adjusted mean weight loss of -5.2% (95% CI, -6.9 to -3.5), -7.6% (95% CI, -9.2 to -6.0), and -13.5% (95% CI, -19.4 to -7.7), respectively (Figure S8). These results are more modest than those reported in the landmark SCALE (8.0% with liraglutide), STEP 1 (14.9% with semaglutide), and SURMOUNT-1 (up to 20.9% with tirzepatide) trials. Of note, those landmark trials enrolled predominantly adults with obesity but without type 2 diabetes mellitus, whereas approximately two-thirds of our cohort had type 2 diabetes, a well-

recognized attenuator of GLP-1RA-associated weight loss. Direct numerical comparison with these trials should therefore be interpreted in light of population differences rather than as a head-to-head benchmark. This attenuation in effect size, known as the efficacy-effectiveness gap, is a common observation in real-world studies, and most plausibly reflects factors such as suboptimal prescribing, slower or incomplete dose titration, and imperfect patient adherence.²⁶ However, despite the smaller magnitude, the ordering we observed (tirzepatide > semaglutide > liraglutide) is consistent with the efficacy ordering reported in controlled trials, although an observational design cannot formally validate it. Because this is an EHR-based, retrospective cohort, any inference about between-agent differences should be interpreted as associational rather than causal.

This study possesses several notable strengths. Its use of the *All of Us* cohort provides demographic and racial diversity, ensuring the findings are more generalizable to the multi-ethnic United States population compared to results from less diverse RCTs.¹⁴ Specifically, our analysis included substantial representation of Hispanic (16.1%) and Black or African American (22.0%) populations, addressing a critical gap in health equity research as these groups are often disproportionately affected by obesity yet underrepresented in clinical trials.^{27,28} Finally, the new-user, active-comparator design helps to mitigate certain biases inherent in observational research, such as immortal time bias.²⁹

The study is subject to limitations inherent to its observational design. Despite extensive multivariable adjustment, residual confounding cannot be excluded. Baseline characteristics were substantially imbalanced across cohorts: standardized mean differences versus the liraglutide reference exceeded the 0.10 threshold for several covariates, most notably type 2 diabetes ($|SMD| \approx 0.65$ for exenatide and 0.50 for dulaglutide), HbA1c, and comorbidity burden ([Table S2](#)), with initiators of the newer agents tending to be younger and to have a lower Charlson Comorbidity Index. All measured imbalances were entered as covariates in the multivariable models, missing covariates were handled by multiple imputation, and the new-initiator, active-comparator design further limited confounding by restricting the comparison to patients judged candidates for GLP-1 receptor agonist therapy. Under alternative model specifications and cohort restrictions ([Table S11](#)), the relative ordering of agents (tirzepatide > semaglutide > the older agents) was preserved throughout. Nonetheless, unmeasured or incompletely captured factors such as diet, physical activity, weight-loss motivation, socioeconomic context, and prescribing preferences cannot be adjusted for and may have contributed to the observed between-agent differences; because these plausibly favor recipients of the newer agents, they could inflate the apparent advantage of semaglutide and tirzepatide.

Most importantly, the tirzepatide cohort was small ($n=386$), with only limited follow-up available given its recent market approval. Events-per-variable in this stratum fell below the conventional threshold of 10 for all four co-primary outcomes (EPV 2.0, 1.2, 0.6, and 3.2 for $\geq 5\%$, $\geq 10\%$, $\geq 15\%$ weight reduction and BMI-category reduction, respectively; [Table S15](#)), limiting the precision of tirzepatide-specific effect estimates and the stability of model fits. Nonetheless, the tirzepatide ranking is corroborated by three PH-free effect estimates: 12-month Kaplan-Meier cumulative incidence ([Table S6](#)), inverse-probability-of-censoring-weighted estimates ([Table S13](#)), and 365-day restricted mean survival time differences ([Table S8](#)), suggesting the qualitative ordering is not solely an artifact of the proportional-hazards assumption, although these estimates rely on the same small number of tirzepatide events and remain imprecise.

Confounding by indication and channeling bias cannot be excluded. The substantial baseline imbalance in type 2 diabetes prevalence (approximately 84–90% in exenatide and dulaglutide initiators versus approximately 51% in semaglutide and tirzepatide initiators), together with corresponding differences in HbA1c and comorbidity burden, reflects differential prescribing across indications and time periods: exenatide and dulaglutide, neither approved for chronic weight management, were preferentially initiated in sicker, predominantly diabetic patients, whereas semaglutide and tirzepatide were more often initiated in healthier, lower-comorbidity patients, including those treated for weight management. The likely direction of this bias is to favor the newer agents, since channeling of healthier patients toward semaglutide and tirzepatide would tend to inflate their apparent advantage while making the older agents appear less effective than they are. We therefore repeated all analyses within diabetes strata, formally tested a drug $\times$ type 2 diabetes interaction, and performed alternative-comparator analyses using exenatide or dulaglutide as the reference. Although the speed of reaching the weight-loss thresholds differed by diabetes status (significant interaction for the weight-threshold time-to-event outcomes), the relative ordering of agents (tirzepatide > semaglutide > the older agents) was preserved

within both diabetes strata (Figures S3–S8 and Table S10) and when the diabetes-indicated agents were used as the comparator (Table S11). Residual confounding by unmeasured indication severity nonetheless cannot be fully excluded.

The high overall prevalence of type 2 diabetes (approximately two-thirds of the cohort) is a further limitation: as a recognized attenuator of GLP-1RA-associated weight loss, it likely renders the absolute reductions observed here conservative relative to obesity populations without diabetes and limits generalizability to that setting.

Finally, calendar-time and secular trends in GLP-1RA prescribing are a concern: tirzepatide and semaglutide were predominantly initiated in 2022–2023, when increases in patient demand, broader media attention, and evolving titration practices may have differed from those of earlier agents. We examined this directly by adjusting for the index calendar year as a continuous covariate (Table S12); tirzepatide and semaglutide remained significantly more likely than liraglutide to achieve all four co-primary outcomes and the tirzepatide > semaglutide ordering was preserved, although the smaller exenatide- and dulaglutide-versus-liraglutide contrasts, which lay near the null, were more sensitive to this adjustment.

Our EHR-based data source lacks granular information on medication adherence, specific dosing and titration schedules, and reasons for discontinuation, all of which are key determinants of treatment effectiveness. Furthermore, body-weight and BMI values were ascertained from routine EHR records rather than standardized study measurements, and may therefore vary in measurement timing, instrumentation, and conditions. The per-agent linear-regression models (Figure 4) do not by themselves test between-agent differences for the secondary outcome; formal pairwise comparisons are therefore drawn from a combined linear-regression model (Table S9), and the per-agent adjusted means are presented descriptively. The cohort also lacked an untreated (no-pharmacotherapy) comparator to illustrate the natural weight trajectory; such a comparison would be confounded by indication and immortal-time bias and is better addressed in future target-trial-emulation studies. The 12-month follow-up window combined with the ± 45 -day analytic interval also means that participants who discontinued or were lost to follow-up before that window did not contribute to the secondary outcome, and unmeasured discontinuation may have differed across agents. Finally, the follow-up time for tirzepatide was limited due to its recent market approval, which precluded a longer-term comparative analysis for this agent.³⁰ More broadly, the 12-month analytic horizon limits inference about the long-term durability and maintenance of weight loss for all agents; our estimates reflect weight change over the first year, and longer-term comparative durability, including the potential for weight regain, remains to be established.

Conclusion

In this large, diverse, real-world cohort of US adults with obesity, initiating treatment with the dual GIP/GLP-1 receptor agonist tirzepatide or the GLP-1RA semaglutide was associated with a significantly greater magnitude of weight change and a higher relative likelihood over time of achieving clinically meaningful weight loss compared to older agents. These findings provide real-world evidence consistent with the relative weight-loss ordering reported in pivotal clinical trials. These differences may also be clinically meaningful, as the greater absolute weight reduction achieved with tirzepatide and semaglutide is associated with improved metabolic and cardiovascular outcomes. However, given the observational design, baseline imbalance across exposure groups, the small sample size and short follow-up for tirzepatide, and the absence of granular adherence and dosing data, these results should be interpreted as associations rather than causal differences. Future research using head-to-head comparative designs and advanced causal-inference methods is warranted to validate these findings and to better establish comparative effectiveness in real-world settings.

Data Sharing Statement

The data used in this study are available to registered researchers via the *All of Us* Researcher Workbench (<https://www.researchallofus.org/>). The specific code and analytic pipelines used for this study are available from the corresponding authors upon reasonable request.

Acknowledgments

We gratefully acknowledge *All of Us* participants for their contributions, without whom this research would not have been possible. We also thank the National Institutes of Health's *All of Us* Research Program for making available the participant data examined in this study. We acknowledge the use of Google Gemini 3.0 (Google LLC) solely for the purpose of language editing and grammatical refinement of this manuscript. The authors reviewed and edited the content as needed and take full responsibility for the scientific content and conclusions.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This work was supported by research funds from the Gyeongsangbuk-do RISE (Regional Innovation System & Education) Project (2025-RISE-15-107), and National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2026-25498828).

Disclosure

The authors report no conflicts of interest in this work.

References

1. Busetto L, Dicker D, Fruhbeck G, et al. A new framework for the diagnosis, staging and management of obesity in adults. *Nat Med.* **2024**;30(9):2395–2399. doi:10.1038/s41591-024-03095-3
2. Webb VL, Wadden TA. Intensive Lifestyle Intervention for Obesity: principles, Practices, and Results. *Gastroenterology.* **2017**;152(7):1752–1764. doi:10.1053/j.gastro.2017.01.045
3. Dandanell S, Skovborg C, Praest CB, et al. Maintaining a clinical weight loss after intensive lifestyle intervention is the key to cardiometabolic health. *Obes Res Clin Pract.* **2017**;11(4):489–498. doi:10.1016/j.orcp.2016.09.009
4. Hall KD, Kahan S. Maintenance of Lost Weight and Long-Term Management of Obesity. *Med Clin North Am.* **2018**;102(1):183–197. doi:10.1016/j.mcna.2017.08.012
5. Grunvald E, Shah R, Hernaez R, et al. AGA Clinical Practice Guideline on Pharmacological Interventions for Adults With Obesity. *Gastroenterology.* **2022**;163(5):1198–1225. doi:10.1053/j.gastro.2022.08.045
6. Muller TD, Bluher M, Tschop MH, DiMarchi RD. Anti-obesity drug discovery: advances and challenges. *Nat Rev Drug Discov.* **2022**;21(3):201–223. doi:10.1038/s41573-021-00337-8
7. Meier JJ. GLP-1 receptor agonists for individualized treatment of type 2 diabetes mellitus. *Nat Rev Endocrinol.* **2012**;8(12):728–742. doi:10.1038/nrendo.2012.140
8. Ostergaard L, Frandsen CS, Madsbad S. Treatment potential of the GLP-1 receptor agonists in type 2 diabetes mellitus: a review. *Expert Rev Clin Pharmacol.* **2016**;9(2):241–265. doi:10.1586/17512433.2016.1121808
9. Pi-Sunyer X, Astrup A, Fujioka K, et al. A Randomized, Controlled Trial of 3.0 mg of Liraglutide in Weight Management. *N Engl J Med.* **2015**;373(1):11–22. doi:10.1056/NEJMoa1411892
10. 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. doi:10.1056/NEJMoa2032183
11. Jastreboff AM, Aronne LJ, Ahmad NN, et al. Tirzepatide Once Weekly for the Treatment of Obesity. *N Engl J Med.* **2022**;387(3):205–216. doi:10.1056/NEJMoa2206038
12. Sallam M, Snygg J, Ghandour SE, Sallam M. Efficacy and Safety of Tirzepatide for Weight Management in Non-Diabetic Obese Individuals: a Narrative Review. *Obesities.* **2025**;5(2):26.
13. Thomsen RW, Mailhac A, Lohde JB, Pottegård A. Real-world evidence on the utilization, clinical and comparative effectiveness, and adverse effects of newer GLP-1RA-based weight-loss therapies. *Diabetes Obesity Metab.* **2025**;27(S2):66–88. doi:10.1111/dom.16364
14. Investigators AURP. The “All of Us” Research Program. *New Engl J Med.* **2019**;381(7):668–676. doi:10.1056/NEJMs1809937
15. Curated Data Repository (CDR) version 8 Release Notes. *All of Us Research Program.* **2025**. Available from: <https://support.researchallofus.org/hc/en-us/articles/30294451486356-Curated-Data-Repository-CDR-version-8-Release-Notes>. Accessed October 13, 2025.
16. Mayo KR, Basford MA, Carroll RJ, et al. The All of Us Data and Research Center: creating a Secure, Scalable, and Sustainable Ecosystem for Biomedical Research. *Annu Rev Biomed Data Sci.* **2023**;6:443–464. doi:10.1146/annurev-biodatasci-122120-104825
17. Observational Medical Outcomes Partnership (OMOP) common data model (CDM). *All of Us Research Program.* **2024**. Available from: <https://support.researchallofus.org/hc/en-us/articles/360039685271-Observational-Medical-Outcomes-Partnership-OMOP-common-data-model-CDM>. Accessed October 13, 2025.

18. Crane J, McGowan B. The GLP-1 agonist, liraglutide, as a pharmacotherapy for obesity. *Therap Adv Chronic Dis.* 2016;7(2):92–107. doi:10.1177/2040622315620180
19. Prevention CfDCA. About Adult BMI. 2024. Available from: <https://www.cdc.gov/bmi/adult-calculator/bmi-categories.html>. Accessed October 13, 2025.
20. Ryan DH, Yockey SR. Weight Loss and Improvement in Comorbidity: differences at 5%, 10%, 15%, and Over. *Curr Obes Rep.* 2017;6(2):187–194. doi:10.1007/s13679-017-0262-y
21. Rubino DM, Greenway FL, Khalid U, et al. Effect of Weekly Subcutaneous Semaglutide vs Daily Liraglutide on Body Weight in Adults With Overweight or Obesity Without Diabetes: the STEP 8 Randomized Clinical Trial. *JAMA.* 2022;327(2):138–150. doi:10.1001/jama.2021.23619
22. Aronne LJ, Horn DB, le Roux CW, et al. Tirzepatide as Compared with Semaglutide for the Treatment of Obesity. *New Engl J Med.* 2025;393(1):26–36. doi:10.1056/NEJMoa2416394
23. Rodriguez PJ, Goodwin Cartwright BM, Gratzl S, et al. Semaglutide vs Tirzepatide for Weight Loss in Adults With Overweight or Obesity. *JAMA Intern Med.* 2024;184(9):1056–1064. doi:10.1001/jamainternmed.2024.2525
24. Buse JB, Rosenstock J, Sesti G, et al. Liraglutide once a day versus exenatide twice a day for type 2 diabetes: a 26-week randomised, parallel-group, multinational, open-label trial (LEAD-6). *Lancet.* 2009;374(9683):39–47. doi:10.1016/S0140-6736(09)60659-0
25. Dungan KM, Povedano ST, Forst T, et al. Once-weekly dulaglutide versus once-daily liraglutide in metformin-treated patients with type 2 diabetes (AWARD-6): a randomised, open-label, Phase 3, non-inferiority trial. *Lancet.* 2014;384(9951):1349–1357. doi:10.1016/S0140-6736(14)60976-4
26. Eichler HG, Abadie E, Breckenridge A, et al. Bridging the efficacy-effectiveness gap: a regulator's perspective on addressing variability of drug response. *Nat Rev Drug Discov.* 2011;10(7):495–506. doi:10.1038/nrd3501
27. Jeyakumar Y, Richardson L, Sarma S, Retnakaran R, Kramer CK. Representation of racialised and ethnically diverse populations in multicentre randomised controlled trials of GLP-1 medicines for obesity: a systematic review and meta-analysis of gaps. *BMJ Global Health.* 2024;9(11):e017177. doi:10.1136/bmjgh-2024-017177
28. Washington TB, Johnson VR, Kendrick K, et al. Disparities in Access and Quality of Obesity Care. *Gastroenterol Clin North Am.* 2023;52(2):429–441. doi:10.1016/j.gtc.2023.02.003
29. Yoshida K, Solomon DH, Kim SC. Active-comparator design and new-user design in observational studies. *Nat Rev Rheumatol.* 2015;11(7):437–441. doi:10.1038/nrrheum.2015.30
30. Syed YY. Tirzepatide: first approval. *Drugs.* 2022;82(11):1213–1220.

Drug Design, Development and Therapy

Publish your work in this journal

Drug Design, Development and Therapy is an international, peer-reviewed open-access journal that spans the spectrum of drug design and development through to clinical applications. Clinical outcomes, patient safety, and programs for the development and effective, safe, and sustained use of medicines are a feature of the journal, which has also been accepted for indexing on PubMed Central. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/drug-design-development-and-therapy-journal>

Dovepress
Taylor & Francis Group